



wwPDB/EMDatabank EM Map/Model Validation Summary Report ⓘ

Oct 14, 2019 – 11:15 AM EDT

PDB ID : 6PYH
EMDB ID: : EMD-20524
Title : Cryo-EM structure of full-length IGF1R-IGF1 complex. Only the extracellular region of the complex is resolved.
Authors : Li, J.; Choi, E.; Yu, H.T.; Bai, X.C.
Deposited on : 2019-07-29
Resolution : 4.30 Å(reported)

This is a wwPDB/EMDatabank EM Map/Model Validation Summary Report
for a publicly released PDB/EMDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/EMValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

MolProbity : 4.02b-467
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : 2.4

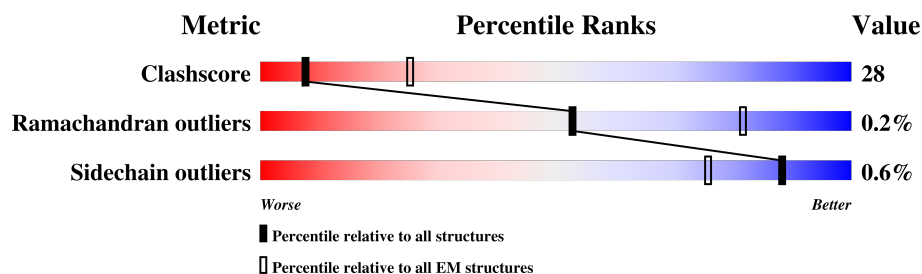
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 4.30 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	136327	1886
Ramachandran outliers	132723	1663
Sidechain outliers	132532	1531

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	1273	
1	D	1273	
2	B	70	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 13187 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Insulin-like growth factor 1 receptor.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	793	Total	C	N	O	S	0	0
			6360	4021	1104	1187	48		
1	D	798	Total	C	N	O	S	0	0
			6411	4060	1111	1193	47		

There are 26 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	951	ALA	TYR	conflict	UNP Q60751
A	1107	ASN	ASP	conflict	UNP Q60751
A	1263	GLY	-	expression tag	UNP Q60751
A	1264	THR	-	expression tag	UNP Q60751
A	1265	SER	-	expression tag	UNP Q60751
A	1266	SER	-	expression tag	UNP Q60751
A	1267	GLY	-	expression tag	UNP Q60751
A	1268	LEU	-	expression tag	UNP Q60751
A	1269	GLU	-	expression tag	UNP Q60751
A	1270	VAL	-	expression tag	UNP Q60751
A	1271	LEU	-	expression tag	UNP Q60751
A	1272	PHE	-	expression tag	UNP Q60751
A	1273	GLN	-	expression tag	UNP Q60751
D	951	ALA	TYR	conflict	UNP Q60751
D	1107	ASN	ASP	conflict	UNP Q60751
D	1263	GLY	-	expression tag	UNP Q60751
D	1264	THR	-	expression tag	UNP Q60751
D	1265	SER	-	expression tag	UNP Q60751
D	1266	SER	-	expression tag	UNP Q60751
D	1267	GLY	-	expression tag	UNP Q60751
D	1268	LEU	-	expression tag	UNP Q60751
D	1269	GLU	-	expression tag	UNP Q60751
D	1270	VAL	-	expression tag	UNP Q60751
D	1271	LEU	-	expression tag	UNP Q60751
D	1272	PHE	-	expression tag	UNP Q60751
D	1273	GLN	-	expression tag	UNP Q60751

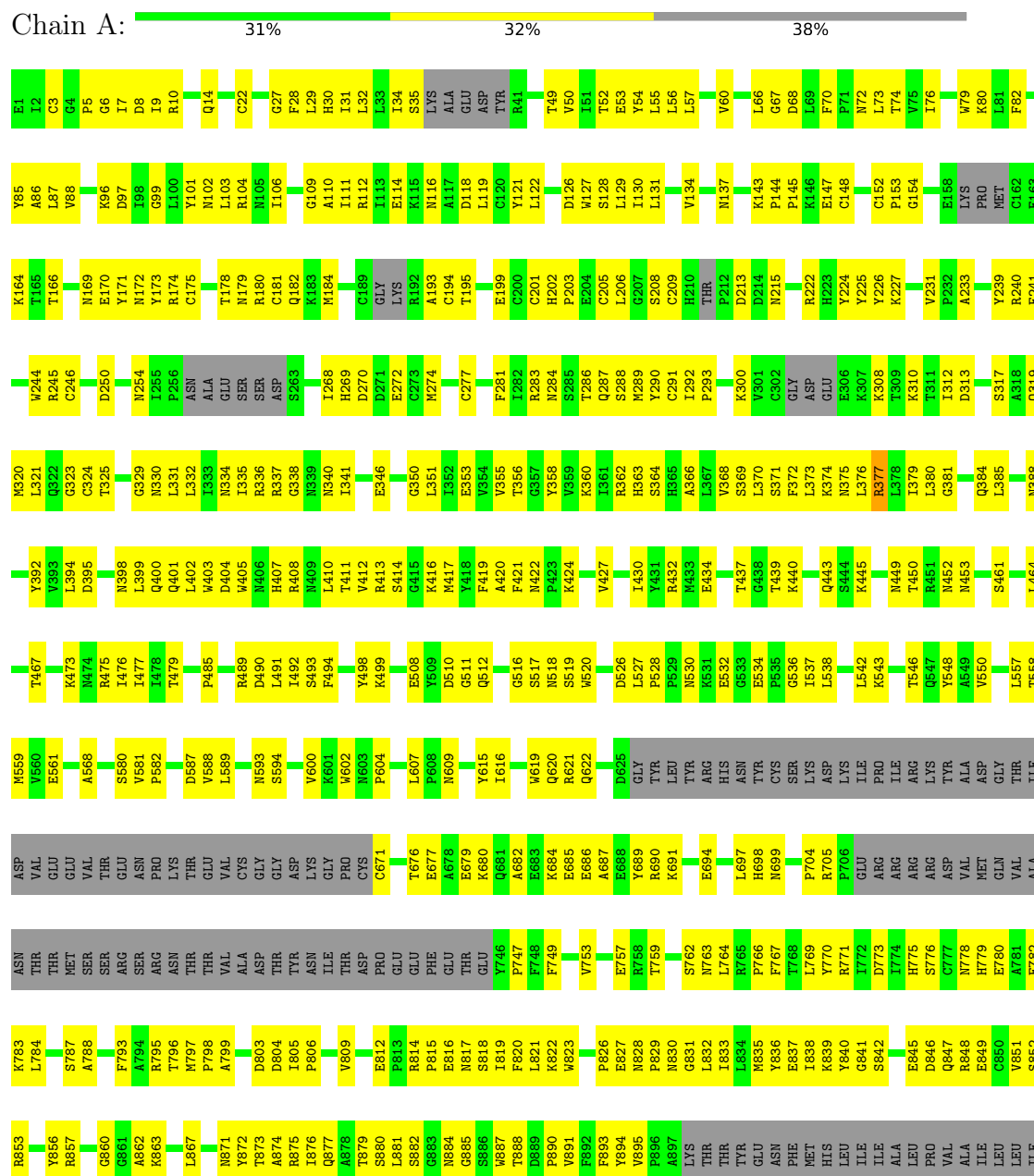
- Molecule 2 is a protein called Insulin-like growth factor I.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	57	Total	C	N	O	S	0	0
			416	261	70	78	7		

3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

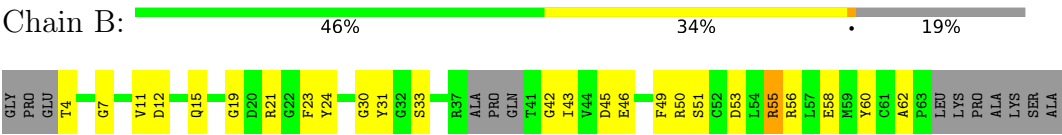
- Molecule 1: Insulin-like growth factor 1 receptor






PRO GLY THR	LEU	GLY	SER	THR	GLN	GLY	LYS	ALA	TRP	GLU	ASN	N755
						ASP	GLY	THR	GLY	ASP	N756	
SER	LEU	GLY	SER	THR	GLN	LEU	ILE	VAL	VAL	PHE	THR	E757
						PRO	LEU	GLN	MET	VAL	R758	
SER	LEU	GLY	SER	THR	GLN	ASP	GLN	LYS	ALA	HIS	ALA	N830
						ASP	VAL	GLU	GLY	ALA	G831	
LEU	LEU	GLY	SER	THR	GLN	ASN	MET	PHE	ARG	ILE	ILE	L832
						ASN	VAL	GLU	GLY	GLU	L832	
VAL	VAL	VAL	VAL	VAL	VAL	CYS	GLY	ASN	LYS	ILE	ALA	M835
						PRO	TRP	GLU	CYS	ILE	ALA	Y836
PHE	PHE	PHE	PHE	PHE	PHE	ASP	ILE	HIS	THR	LEU	LEU	E837
						LEU	ALA	HIS	MET	PRO	ASN	L764
GLN	GLN	GLN	GLN	GLN	GLN	PRO	ASP	VAL	ASN	VAL	VAL	I838
						PHE	GLY	VAL	ARG	ALA	ARG	K839
PHE	PHE	PHE	PHE	PHE	PHE	GLU	GLY	VAL	ARG	ALA	ALA	Y840
						GLU	SER	MET	ARG	ILE	ILE	F767
GLN	GLN	GLN	GLN	GLN	GLN	LEU	ALA	LEU	LEU	LEU	LEU	T768
						LEU	LEU	LEU	LEU	LEU	LEU	L769
PHE	PHE	PHE	PHE	PHE	PHE	ASP	LYS	GLY	GLY	ILE	ILE	V770
						ARG	GLY	VAL	GLN	GLY	VAL	R771
GLN	GLN	GLN	GLN	GLN	GLN	TRP	PHE	ASN	SER	GLY	GLY	D773
						GLN	THR	LYS	PHE	GLY	GLY	E849
GLN	GLN	GLN	GLN	GLN	GLN	ASN	VAL	GLN	PHE	GLY	GLY	G850
						ASN	VAL	GLN	GLY	GLY	GLY	V851
GLN	GLN	GLN	GLN	GLN	GLN	THR	THR	GLY	MET	VAL	VAL	S852
						THR	THR	GLY	MET	VAL	VAL	I853
GLN	GLN	GLN	GLN	GLN	GLN	HIS	HIS	VAL	VAL	ILE	ILE	C777
						ASN	ASN	GLN	VAL	ILE	ILE	N778
GLN	GLN	GLN	GLN	GLN	GLN	PRO	HIS	PRO	TYR	MET	MET	H775
						PRO	ASN	THR	GLY	LEU	LEU	E855
GLN	GLN	GLN	GLN	GLN	GLN	LYS	ASN	THR	GLY	TYR	TYR	H779
						ASN	ASN	GLY	GLY	GLY	GLY	Y856
GLN	GLN	GLN	GLN	GLN	GLN	MET	VAL	VAL	VAL	VAL	VAL	R857
						ARG	TRP	LEU	VAL	VAL	VAL	A781
GLN	GLN	GLN	GLN	GLN	GLN	SER	ALA	ILE	ALA	PHE	PHE	E782
						PRO	SER	ALA	ALA	ALA	ALA	K858
GLN	GLN	GLN	GLN	GLN	GLN	SER	ALA	MET	LYS	HIS	HIS	K863
						THR	PHE	GLY	GLY	ARG	ARG	L864
GLN	GLN	GLN	GLN	GLN	GLN	LEU	ASN	LEU	VAL	LYS	LYS	N865
						VAL	VAL	VAL	VAL	VAL	VAL	R866
GLN	GLN	GLN	GLN	GLN	GLN	ILE	CYS	MET	THR	ASN	ASN	A788
						ILE	LEU	MET	THR	ASN	ASN	F787
GLN	GLN	GLN	GLN	GLN	GLN	ILE	VAL	ARG	ASP	ASN	ASN	G867
						GLY	VAL	GLY	GLU	GLY	GLY	A793
GLN	GLN	GLN	GLN	GLN	GLN	SER	ILE	ASP	PRO	ARG	ARG	R868
						ILE	ILE	GLY	GLY	GLY	GLY	R795
GLN	GLN	GLN	GLN	GLN	GLN	LYS	THR	PHE	THR	GLY	GLY	T796
						ASN	THR	ASN	ASN	ASN	ASN	M797
GLN	GLN	GLN	GLN	GLN	GLN	ASP	LEU	SER	ARG	ASN	ASN	F798
						GLY	ALA	VAL	VAL	GLY	GLY	A799
GLN	GLN	GLN	GLN	GLN	GLN	MET	GLY	LYS	VAL	VAL	VAL	R875
						GLY	GLY	GLY	ALA	ALA	ALA	I876
GLN	GLN	GLN	GLN	GLN	GLN	GLY	GLN	ARG	ILE	LEU	LEU	Q877
						VAL	VAL	VAL	VAL	VAL	VAL	A878
GLN	GLN	GLN	GLN	GLN	GLN	PRO	GLY	SER	LYS	TYR	TYR	E800
						ASN	ASN	ASN	ASN	ASN	ASN	R805
GLN	GLN	GLN	GLN	GLN	GLN	PHE	GLY	ARG	VAL	SER	SER	P806
						GLN	GLY	PRO	VAL	VAL	VAL	E810
GLN	GLN	GLN	GLN	GLN	GLN	GLY	LEU	GLY	ASN	ASN	ASN	T810
						VAL	VAL	VAL	VAL	VAL	VAL	W811
GLN	GLN	GLN	GLN	GLN	GLN	LEU	THR	VAL	ALA	ALA	ALA	E812
						GLY	GLY	GLY	GLY	GLY	GLY	P813
GLN	GLN	GLN	GLN	GLN	GLN	PHE	GLY	GLN	SER	ALA	ALA	R814
						GLY	GLY	ASN	MET	PHE	PHE	E815
GLN	GLN	GLN	GLN	GLN	GLN	TYR	TYR	ASN	ARG	SER	SER	P815
						LEU	LEU	ASN	ARG	ALA	ALA	E816
GLN	GLN	GLN	GLN	GLN	GLN	GLY	THR	VAL	GLY	ALA	ALA	N817
						GLY	GLY	VAL	ILE	ILE	ILE	S818
GLN	GLN	GLN	GLN	GLN	GLN	ASP	ASP	GLY	ILE	ASP	ASP	R818
						GLY	GLY	GLY	GLY	GLY	GLY	E819
GLN	GLN	GLN	GLN	GLN	GLN	VAL	TYR	ILE	GLY	VAL	VAL	F820
						VAL	TYR	VAL	VAL	VAL	VAL	L821
GLN	GLN	GLN	GLN	GLN	GLN	LYS	TYR	PRO	PHE	TYR	TYR	R822
						ASN	ASN	ASN	ASN	ASN	ASN	E823
GLN	GLN	GLN	GLN	GLN	GLN	PRO	ARG	PRO	LEU	VAL	VAL	L821
						PRO	GLY	GLY	GLY	GLY	GLY	R822
GLN	GLN	GLN	GLN	GLN	GLN	GLY	GLY	SER	GLY	ASP	ASP	W823
						GLY	GLY	LEU	GLY	GLY	GLY	E823

● Molecule 2: Insulin-like growth factor I



4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	51573	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	50	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	GATAN K2 IS (4k x 4k)	Depositor

5 Model quality [i](#)

5.1 Standard geometry [i](#)

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z >2	RMSZ	# Z >2
1	A	0.45	0/6509	0.54	0/8828
1	D	0.41	0/6561	0.52	0/8897
2	B	0.43	0/422	0.53	0/568
All	All	0.43	0/13492	0.53	0/18293

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	D	0	2
All	All	0	3

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	816	GLU	Peptide
1	D	815	PRO	Peptide
1	D	816	GLU	Peptide

5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	6360	0	6171	343	0
1	D	6411	0	6201	363	0
2	B	416	0	357	34	0
All	All	13187	0	12729	719	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 28.

The worst 5 of 719 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:872:TYR:H	1:A:894:TYR:HA	1.27	1.00
1:D:238:THR:HA	1:D:247:VAL:O	1.63	0.99
1:D:838:ILE:HB	1:D:849:GLU:HB3	1.52	0.90
1:A:164:LYS:HB3	1:A:173:TYR:HD1	1.41	0.85
1:D:769:LEU:HB2	1:D:795:ARG:HH21	1.44	0.83

There are no symmetry-related clashes.

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	773/1273 (61%)	631 (82%)	142 (18%)	0	100	100
1	D	776/1273 (61%)	629 (81%)	144 (19%)	3 (0%)	36	76
2	B	53/70 (76%)	48 (91%)	5 (9%)	0	100	100
All	All	1602/2616 (61%)	1308 (82%)	291 (18%)	3 (0%)	53	84

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	817	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	816	GLU
1	D	815	PRO

5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	709/1132 (63%)	706 (100%)	3 (0%)	92	95
1	D	708/1132 (62%)	703 (99%)	5 (1%)	85	92
2	B	39/57 (68%)	38 (97%)	1 (3%)	49	73
All	All	1456/2321 (63%)	1447 (99%)	9 (1%)	88	93

5 of 9 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	174	ARG
1	D	777	CYS
1	D	408	ARG
1	A	473	LYS
1	D	181	CYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 30 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	775	HIS
1	D	275	GLN
1	D	763	ASN
1	D	11	ASN
1	D	365	HIS

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	1
1	D	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	D	798:PRO	C	799:ALA	N	4.80
1	A	798:PRO	C	799:ALA	N	2.94