



# wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Nov 14, 2019 – 08:20 AM EST

PDB ID : 6SOF  
EMDB ID: : EMD-10273  
Title : human insulin receptor ectodomain bound by 4 insulin  
Authors : Gutmann, T.; Schaefer, I.B.; Poojari, C.S.; Vattulainen, I.; Strauss, M.;  
Coskun, U.  
Deposited on : 2019-08-29  
Resolution : 4.30 Å(reported)  
Based on PDB ID : 6CEB

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](mailto: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

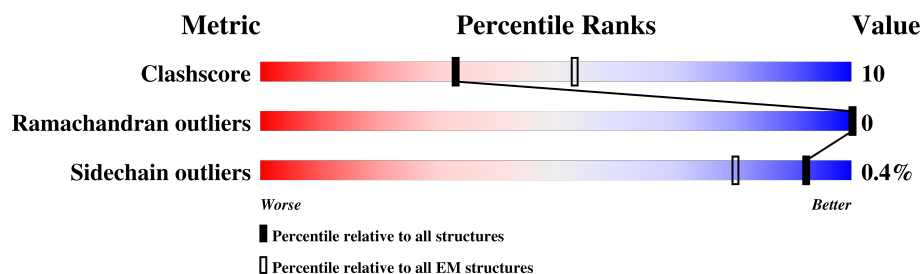
# 1 Overall quality at a glance

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  $\geq 3$ , 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	719	76% 24%
1	C	719	75% 24%
2	B	162	69% 31%
2	D	162	88% 11% .
3	E	21	57% 43%
3	G	21	71% 29%
3	I	21	67% 33%
3	K	21	76% 24%
4	F	30	83% 17%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
4	H	30	 83% 17%
4	J	30	 80% 20%
4	L	30	 87% 13%

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 30333 atoms, of which 14564 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 receptor.

Mol	Chain	Residues	Atoms						AltConf	Trace
1	A	719	Total	C	H	N	O	S	0	0
			11127	3654	5349	990	1088	46		
1	C	719	Total	C	H	N	O	S	0	0
			11118	3654	5337	993	1088	46		

- Molecule 2 is a protein called Insulin receptor.

Mol	Chain	Residues	Atoms						AltConf	Trace
2	B	162	Total	C	H	N	O	S	0	0
			2532	821	1229	232	244	6		
2	D	162	Total	C	H	N	O	S	0	0
			2508	818	1210	232	242	6		

- Molecule 3 is a protein called Insulin.

Mol	Chain	Residues	Atoms						AltConf	Trace
3	G	21	Total	C	H	N	O	S	0	0
			286	96	128	25	33	4		
3	E	21	Total	C	H	N	O	S	0	0
			299	99	136	25	35	4		
3	K	21	Total	C	H	N	O	S	0	0
			292	96	132	25	35	4		
3	I	21	Total	C	H	N	O	S	0	0
			292	96	132	25	35	4		

- Molecule 4 is a protein called Insulin.

Mol	Chain	Residues	Atoms						AltConf	Trace
4	H	30	Total	C	H	N	O	S	0	0
			471	158	229	40	42	2		
4	F	30	Total	C	H	N	O	S	0	0
			471	158	229	40	42	2		

*Continued on next page...*

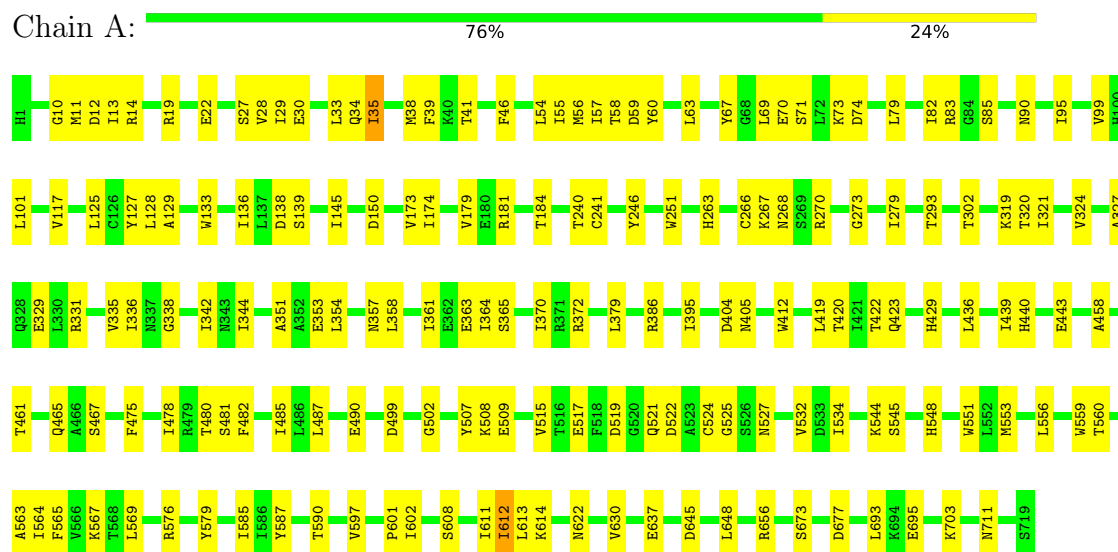
*Continued from previous page...*

Mol	Chain	Residues	Atoms						AltConf	Trace
4	L	30	Total	C	H	N	O	S	0	0
			471	158	229	40	42	2		
4	J	30	Total	C	H	N	O	S	0	0
			466	158	224	40	42	2		

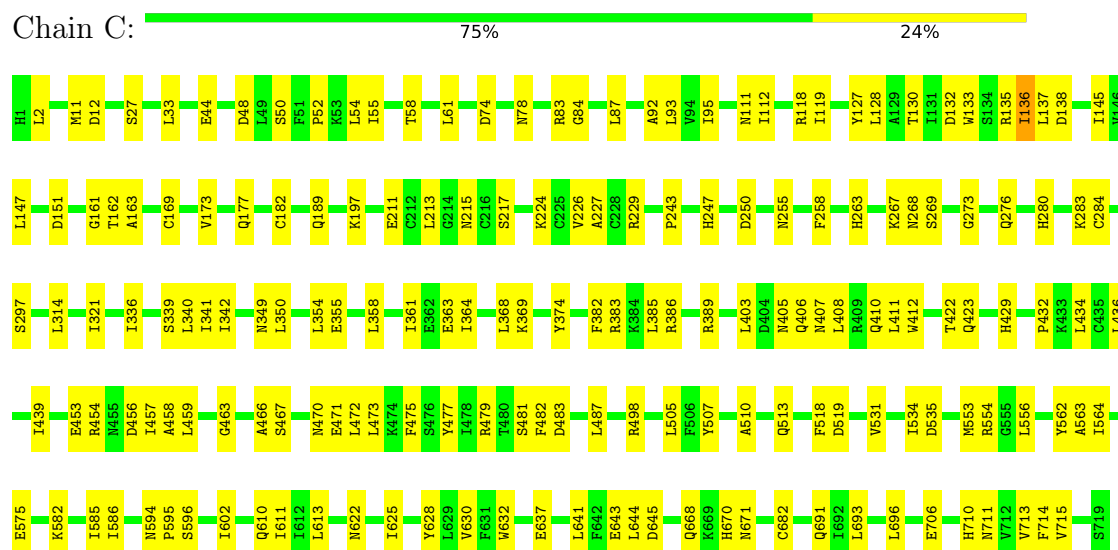
### 3 Residue-property plots

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 receptor

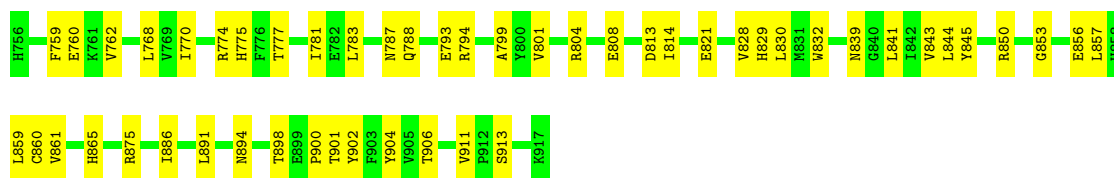


#### • Molecule 1: Insulin receptor



#### • Molecule 2: Insulin receptor

Chain B:  69% 31%



- Molecule 2: Insulin receptor

Chain D:  88% 11%



- Molecule 3: Insulin

Chain G:  71% 29%




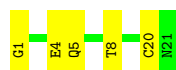
- Molecule 3: Insulin

Chain E:  57% 43%



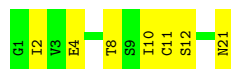
- Molecule 3: Insulin

Chain K:  76% 24%




- Molecule 3: Insulin

Chain I:  67% 33%




- Molecule 4: Insulin

Chain H:  83% 17%




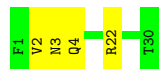
- Molecule 4: Insulin

Chain F:  83% 17%




- Molecule 4: Insulin

Chain L:  87% 13%



- Molecule 4: Insulin

Chain J:  80% 20%



## 4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	326257	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	NONE	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	55	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor

## 5 Model quality

### 5.1 Standard geometry

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	0.34	0/5924	0.62	0/8037
1	C	0.34	0/5927	0.61	0/8040
2	B	0.38	0/1335	0.68	0/1811
2	D	0.34	0/1330	0.66	0/1806
3	E	0.40	0/164	0.60	0/220
3	G	0.34	0/159	0.46	0/215
3	I	0.38	0/161	0.60	0/216
3	K	0.31	0/161	0.54	0/216
4	F	0.35	0/249	0.51	0/335
4	H	0.34	0/249	0.57	0/335
4	J	0.33	0/249	0.59	0/335
4	L	0.34	0/249	0.54	0/335
All	All	0.34	0/16157	0.62	0/21901

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	C	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	243	PRO	Peptide

## 5.2 Too-close contacts ⓘ

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	5778	5349	5572	119	0
1	C	5781	5337	5576	129	0
2	B	1303	1229	1264	40	0
2	D	1298	1210	1255	16	0
3	E	163	136	149	8	0
3	G	158	128	142	6	0
3	I	160	132	140	5	0
3	K	160	132	140	3	0
4	F	242	229	232	5	0
4	H	242	229	234	4	0
4	J	242	224	232	6	0
4	L	242	229	232	4	0
All	All	15769	14564	15168	309	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 10.

The worst 5 of 309 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)
2:D:915:ILE:HG22	2:D:917:LYS:HG3	1.66	0.77
1:C:119:ILE:HD12	1:C:145:ILE:HG12	1.67	0.76
1:A:263:HIS:O	1:A:273:GLY:N	2.20	0.75
1:C:432:PRO:O	1:C:463:GLY:N	2.19	0.75
1:A:129:ALA:O	1:A:181:ARG:NH1	2.21	0.73

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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	717/719 (100%)	589 (82%)	128 (18%)	0	100	100
1	C	717/719 (100%)	594 (83%)	123 (17%)	0	100	100
2	B	160/162 (99%)	131 (82%)	29 (18%)	0	100	100
2	D	160/162 (99%)	122 (76%)	38 (24%)	0	100	100
3	E	19/21 (90%)	17 (90%)	2 (10%)	0	100	100
3	G	19/21 (90%)	18 (95%)	1 (5%)	0	100	100
3	I	19/21 (90%)	16 (84%)	3 (16%)	0	100	100
3	K	19/21 (90%)	18 (95%)	1 (5%)	0	100	100
4	F	28/30 (93%)	26 (93%)	2 (7%)	0	100	100
4	H	28/30 (93%)	25 (89%)	3 (11%)	0	100	100
4	J	28/30 (93%)	24 (86%)	4 (14%)	0	100	100
4	L	28/30 (93%)	19 (68%)	9 (32%)	0	100	100
All	All	1942/1966 (99%)	1599 (82%)	343 (18%)	0	100	100

There are no Ramachandran outliers to report.

### 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	656/658 (100%)	652 (99%)	4 (1%)	87	93
1	C	656/658 (100%)	654 (100%)	2 (0%)	93	96
2	B	142/142 (100%)	142 (100%)	0	100	100
2	D	141/142 (99%)	140 (99%)	1 (1%)	85	92
3	E	20/20 (100%)	20 (100%)	0	100	100
3	G	19/20 (95%)	19 (100%)	0	100	100
3	I	19/20 (95%)	19 (100%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	K	19/20 (95%)	19 (100%)	0	100	100
4	F	26/26 (100%)	26 (100%)	0	100	100
4	H	26/26 (100%)	26 (100%)	0	100	100
4	J	26/26 (100%)	26 (100%)	0	100	100
4	L	26/26 (100%)	26 (100%)	0	100	100
All	All	1776/1784 (100%)	1769 (100%)	7 (0%)	92	95

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	612	ILE
2	D	770	ILE
1	C	136	ILE
1	A	95	ILE
1	C	534	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 27 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	788	GLN
1	C	280	HIS
1	C	594	ASN
2	B	894	ASN
1	A	148	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

## 5.5 Carbohydrates ⓘ

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](#)

There are no chain breaks in this entry.