



wwPDB EM Validation Summary Report ⓘ

Oct 7, 2024 – 02:26 PM JST

PDB ID : 8WA3
EMDB ID : EMD-37390
Title : Cryo-EM structure of peptide free and Gs-coupled GIPR
Authors : Cong, Z.T.; Zhao, F.H.; Li, Y.; Luo, G.; Zhou, Q.T.; Yang, D.H.; Wang, M.W.
Deposited on : 2023-09-06
Resolution : 2.86 Å(reported)

This is a wwPDB EM Validation Summary Report for a publicly released PDB 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.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

EMDB validation analysis : **FAILED**
MolProbity : 4.02b-467
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ : **FAILED**
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.39

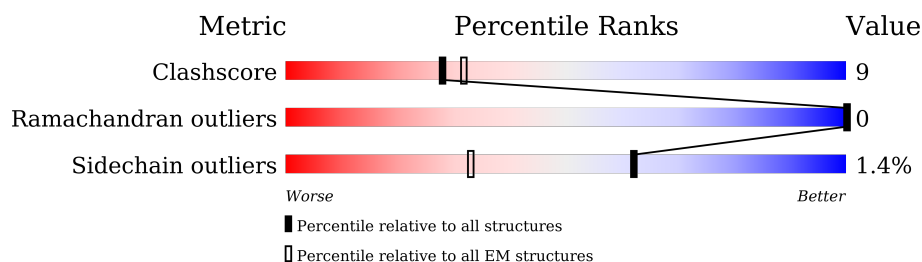
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 2.86 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	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. 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	R	573	
2	A	394	
3	B	371	
4	G	71	
5	N	140	

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 8163 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 Gastric inhibitory polypeptide receptor,Fusion protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	R	274	Total	C	N	O	S	0	0
			2222	1459	394	358	11		

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
R	345	PHE	THR	engineered mutation	UNP P48546
R	422	GLY	-	linker	UNP P48546
R	423	SER	-	linker	UNP P48546
R	424	SER	-	linker	UNP P48546
R	425	GLY	-	linker	UNP P48546
R	426	GLY	-	linker	UNP P48546
R	427	GLY	-	linker	UNP P48546
R	428	GLY	-	linker	UNP P48546
R	429	SER	-	linker	UNP P48546
R	430	GLY	-	linker	UNP P48546
R	431	GLY	-	linker	UNP P48546
R	432	GLY	-	linker	UNP P48546
R	433	GLY	-	linker	UNP P48546
R	434	SER	-	linker	UNP P48546
R	435	SER	-	linker	UNP P48546
R	436	GLY	-	linker	UNP P48546

- Molecule 2 is a protein called Guanine nucleotide-binding protein G(s) subunit alpha isoforms short.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	A	232	Total	C	N	O	S	0	0
			1922	1210	348	357	7		

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	54	ASN	SER	engineered mutation	UNP P04896
A	226	ALA	GLY	engineered mutation	UNP P04896
A	268	ALA	GLU	engineered mutation	UNP P04896
A	271	LYS	ASN	engineered mutation	UNP P04896
A	274	ASP	LYS	engineered mutation	UNP P04896
A	280	LYS	ARG	engineered mutation	UNP P04896
A	284	ASP	THR	engineered mutation	UNP P04896
A	285	THR	ILE	engineered mutation	UNP P04896
A	366	SER	ALA	engineered mutation	UNP P04896

- Molecule 3 is a protein called Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-1,O-antigen polymerase.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	B	341	Total	C	N	O	S	0	0
			2616	1612	470	513	21		

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-4	MET	-	initiating methionine	UNP P54311
B	-3	GLY	-	expression tag	UNP P54311
B	-2	SER	-	expression tag	UNP P54311
B	-1	LEU	-	expression tag	UNP P54311
B	0	LEU	-	expression tag	UNP P54311
B	1	GLN	-	expression tag	UNP P54311
B	341	GLY	-	linker	UNP P54311
B	342	SER	-	linker	UNP P54311
B	343	SER	-	linker	UNP P54311
B	344	GLY	-	linker	UNP P54311
B	345	GLY	-	linker	UNP P54311
B	346	GLY	-	linker	UNP P54311
B	347	GLY	-	linker	UNP P54311
B	348	SER	-	linker	UNP P54311
B	349	GLY	-	linker	UNP P54311
B	350	GLY	-	linker	UNP P54311
B	351	GLY	-	linker	UNP P54311
B	352	GLY	-	linker	UNP P54311
B	353	SER	-	linker	UNP P54311
B	354	SER	-	linker	UNP P54311
B	355	GLY	-	linker	UNP P54311
B	356	VAL	-	linker	UNP P54311
B	357	SER	-	linker	UNP P54311

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	358	GLY	-	linker	UNP P54311

- Molecule 4 is a protein called Guanine nucleotide-binding protein G(I)/G(S)/G(O) subunit gamma-2.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	G	57	Total	C	N	O	S	0	0
			436	273	77	83	3		

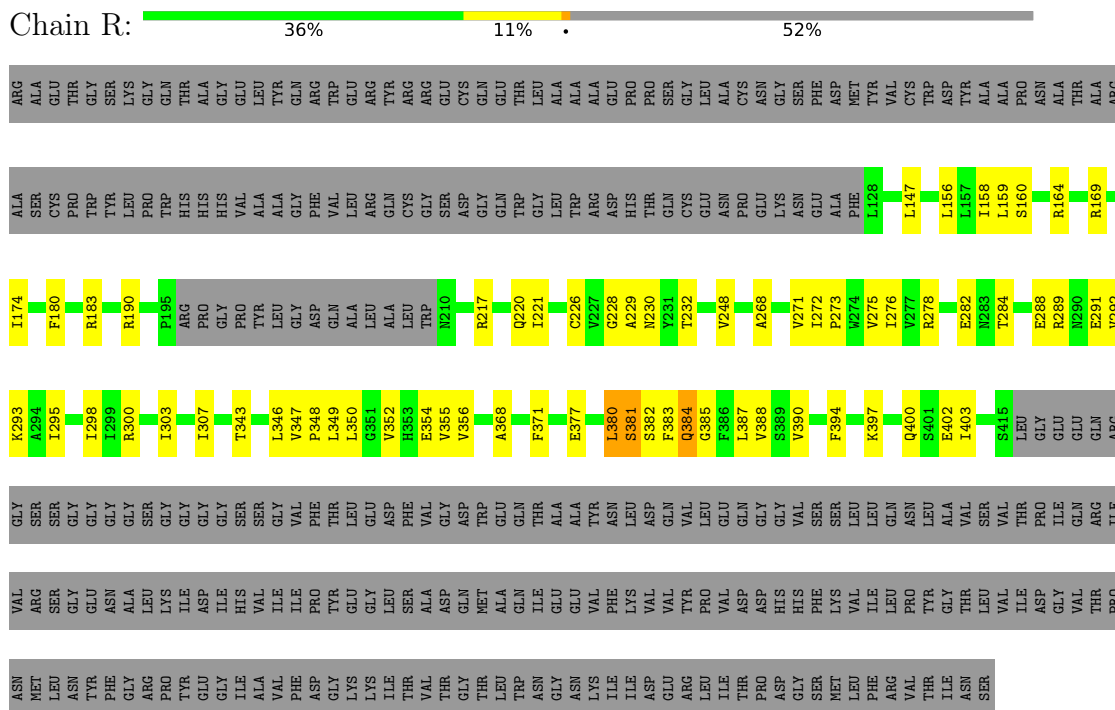
- Molecule 5 is a protein called Nanobody-35.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	N	127	Total	C	N	O	S	0	0
			967	602	169	190	6		

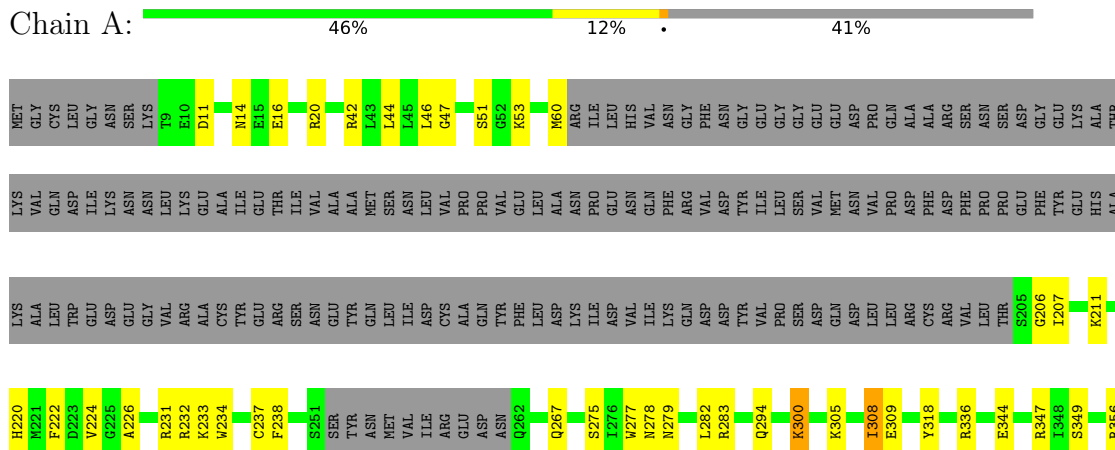
3 Residue-property plots

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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: Gastric inhibitory polypeptide receptor, Fusion protein



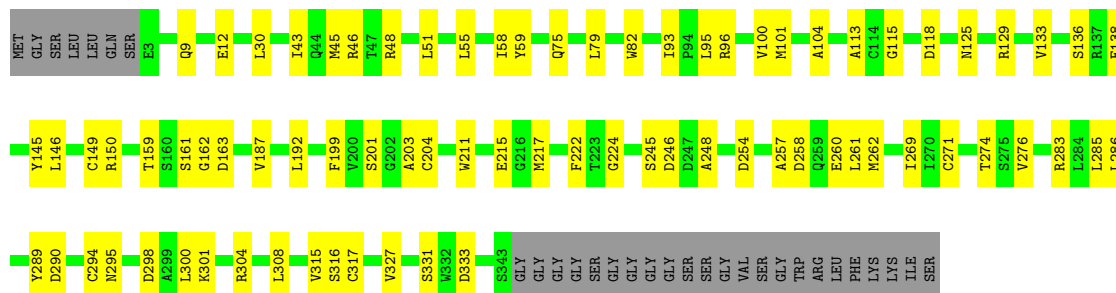
- Molecule 2: Guanine nucleotide-binding protein G(s) subunit alpha isoforms short





- Molecule 3: Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-1, O-antigen polymerase

Chain B: 71% 21% 8%



- Molecule 4: Guanine nucleotide-binding protein G(I)/G(S)/G(O) subunit gamma-2

Chain G: 69% 11% 20%



- Molecule 5: Nanobody-35

Chain N: 66% 25% 9%



4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	807839	Depositor
Resolution determination method	OTHER	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	80	Depositor
Minimum defocus (nm)	1200	Depositor
Maximum defocus (nm)	2200	Depositor
Magnification	Not provided	
Image detector	GATAN K3 (6k 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	$\# Z > 5$	RMSZ	$\# Z > 5$
1	R	0.26	0/2272	0.48	1/3085 (0.0%)
2	A	0.26	0/1957	0.41	0/2629
3	B	0.25	0/2663	0.45	0/3610
4	G	0.24	0/442	0.45	0/597
5	N	0.27	0/987	0.49	0/1337
All	All	0.26	0/8321	0.46	1/11258 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	R	147	LEU	CA-CB-CG	5.11	127.05	115.30

There are no chirality outliers.

There are no planarity outliers.

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	R	2222	0	2300	51	0
2	A	1922	0	1901	34	0
3	B	2616	0	2518	44	0
4	G	436	0	448	5	0
5	N	967	0	933	28	0
All	All	8163	0	8100	152	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 9.

The worst 5 of 152 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)
3:B:75:GLN:NE2	3:B:100:VAL:O	2.18	0.77
1:R:354:GLU:HG3	1:R:380:LEU:HB2	1.65	0.76
3:B:294:CYS:HB3	3:B:308:LEU:HB2	1.70	0.72
3:B:115:GLY:HA3	3:B:146:LEU:HD23	1.73	0.71
3:B:254:ASP:HB2	3:B:261:LEU:HD11	1.74	0.70

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	R	270/573 (47%)	252 (93%)	18 (7%)	0	100	100
2	A	226/394 (57%)	215 (95%)	11 (5%)	0	100	100
3	B	339/371 (91%)	328 (97%)	11 (3%)	0	100	100
4	G	55/71 (78%)	55 (100%)	0	0	100	100
5	N	125/140 (89%)	113 (90%)	12 (10%)	0	100	100
All	All	1015/1549 (66%)	963 (95%)	52 (5%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains [i](#)

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	R	234/483 (48%)	229 (98%)	5 (2%)	48	72
2	A	208/351 (59%)	202 (97%)	6 (3%)	37	63
3	B	283/302 (94%)	282 (100%)	1 (0%)	89	95
4	G	46/58 (79%)	46 (100%)	0	100	100
5	N	105/116 (90%)	105 (100%)	0	100	100
All	All	876/1310 (67%)	864 (99%)	12 (1%)	62	82

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	A	300	LYS
2	A	305	LYS
3	B	129	ARG
2	A	308	ILE
1	R	381	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
2	A	390	GLN

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 oligosaccharides in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.