WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB EM Validation Summary Report (i)

Nov 5, 2024 – 10:09 AM JST

PDB ID	:	8HMV
EMDB ID	:	EMD-34903
Title	:	Structure of GPR21-Gs complex
Authors	:	Wong, T.S.; Gao, W.
Deposited on	:	2022-12-05
Resolution	:	2.91 Å(reported)

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

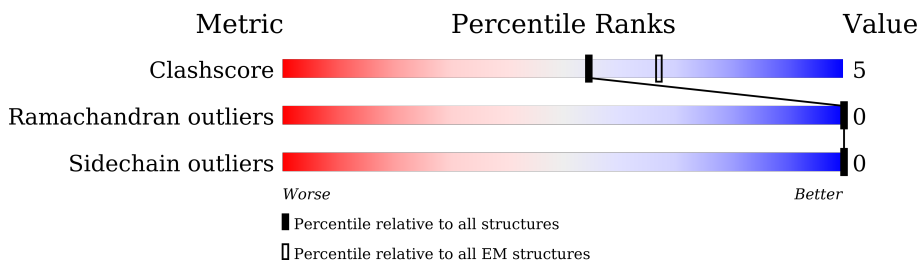|		
EMDB validation analysis	:	FAILED
MolProbity	:	4.02b-467
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ	:	FAILED
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

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 2.91 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	Percentile Ranks	Value
Clashscore		5
Ramachandran outliers		0
Sidechain outliers		0

Worse   *Better*

▮ Percentile relative to all structures
▯ Percentile relative to all EM structures

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 <=5%

Mol	Chain	Length	Quality of chain
1	A	289	84% 11% 5%
2	C	384	52% 7% 41%
3	B	338	86% 14%
4	G	57	95% 5%
5	N	126	88% 12%

2   Entry composition ⓘ

There are 5 unique types of molecules in this entry. The entry contains 8127 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 Probable G-protein coupled receptor 21.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	274	Total 2237	C 1498	N 354	O 370	S 15	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
2	C	227	Total 1892	C 1194	N 344	O 347	S 7	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	205	THR	SER	conflict	UNP P63092
C	226	ALA	GLY	engineered mutation	UNP P63092
C	366	SER	ALA	engineered mutation	UNP P63092

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

Mol	Chain	Residues	Atoms					AltConf	Trace
3	B	338	Total 2601	C 1604	N 467	O 509	S 21	0	0

- 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 436	C 273	N 77	O 83	S 3	0	0

- Molecule 5 is a protein called Nanobody Nb35.

| Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	N	126	961	599	168	188	6	0	0

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.
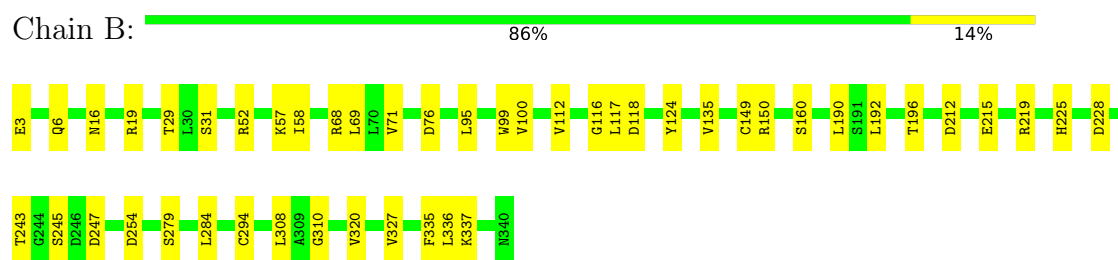
- Molecule 1: Probable G-protein coupled receptor 21

Chain A:
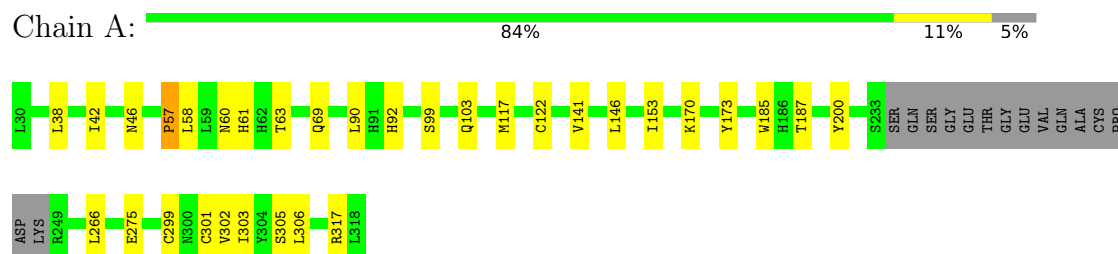
- Molecule 2: Guanine nucleotide-binding protein G(s) subunit alpha isoforms short
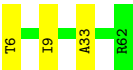
Chain C:
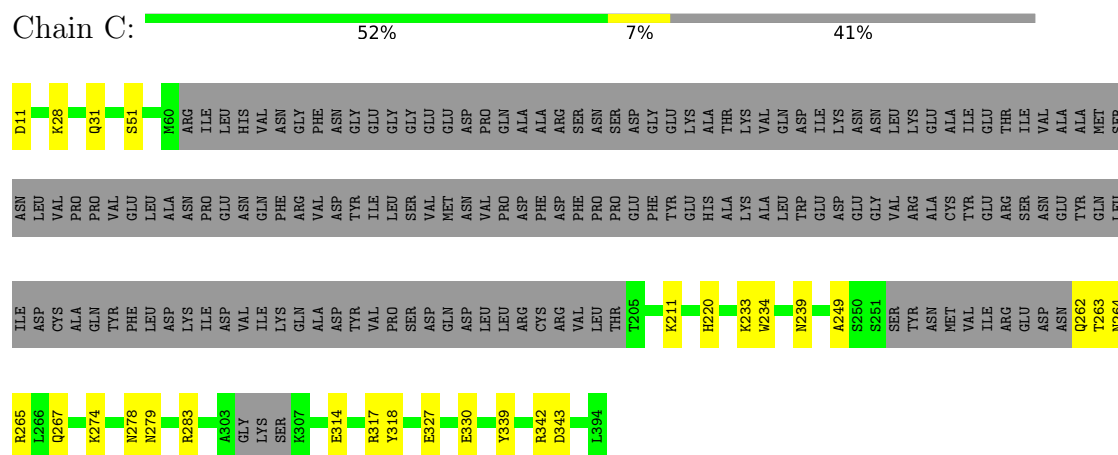
- Molecule 3: Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-1

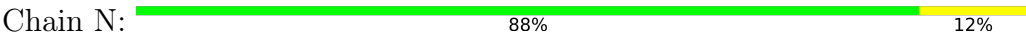Chain B:

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

Chain G:

T6 I9 A33 R62

• Molecule 5: Nanobody Nb35

Chain N:

88%   12%

Q1 L4 V12 L18 F29 W36 S52 Q53 R67 I70 S71 R72 D90 V110 T111 S112 Y117 R118 V126

4   Experimental information ⓘ

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	149012	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN	Depositor
Voltage (kV)	300	Depositor
Electron dose $(e^-/\text{Å}^2)$	46	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	1400	Depositor
Magnification	Not provided	
Image detector	GATAN K3 BIOQUANTUM (6k 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	$\#|Z|>5$	RMSZ	$\#|Z|>5$
1	A	0.32	0/2308	0.45	0/3150
2	C	0.30	0/1926	0.42	0/2587
3	B	0.32	0/2648	0.52	0/3589
4	G	0.27	0/442	0.44	0/597
5	N	0.33	0/981	0.48	0/1329
All	All	0.31	0/8305	0.47	0/11252

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

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	A	57	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.

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2237	0	2255	24	0
2	C	1892	0	1877	18	0
3	B	2601	0	2505	29	0
4	G	436	0	448	2	0
5	N	961	0	928	10	0
All	All	8127	0	8013	79	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 5.

The worst 5 of 79 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:103:GLN:HE22	1:A:185:TRP:HE1	1.40	0.68
2:C:239:ASN:OD1	3:B:57:LYS:NZ	2.23	0.66
3:B:310:GLY:O	3:B:337:LYS:NZ	2.24	0.66
2:C:263:THR:O	2:C:265:ARG:NH1	2.28	0.66
5:N:67:ARG:NH2	5:N:90:ASP:OD2	2.30	0.65

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	270/289 (93%)	257 (95%)	13 (5%)	0	100	100
2	C	219/384 (57%)	216 (99%)	3 (1%)	0	100	100
3	B	336/338 (99%)	325 (97%)	11 (3%)	0	100	100
4	G	55/57 (96%)	54 (98%)	1 (2%)	0	100	100
5	N	124/126 (98%)	119 (96%)	5 (4%)	0	100	100
All	All	1004/1194 (84%)	971 (97%)	33 (3%)	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	252/264 (96%)	252 (100%)	0	100	100
2	C	205/343 (60%)	205 (100%)	0	100	100
3	B	281/281 (100%)	281 (100%)	0	100	100
4	G	46/46 (100%)	46 (100%)	0	100	100
5	N	104/104 (100%)	104 (100%)	0	100	100
All	All	888/1038 (86%)	888 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	103	GLN
3	B	75	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.