wwPDB EM Validation Summary Report (i)

Oct 7, 2024 – 02:07 PM JST

PDB ID	:	8HQE
EMDB ID	:	EMD-34948
Title	:	Cryo-EM structure of the apo-GPR132-Gi
Authors	:	Wang, J.L.; Ding, J.H.; Sun, J.P.; Yu, X.
Deposited on	:	2022-12-13
Resolution	:	2.97 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 (i) 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 (i)) were used in the production of this report:

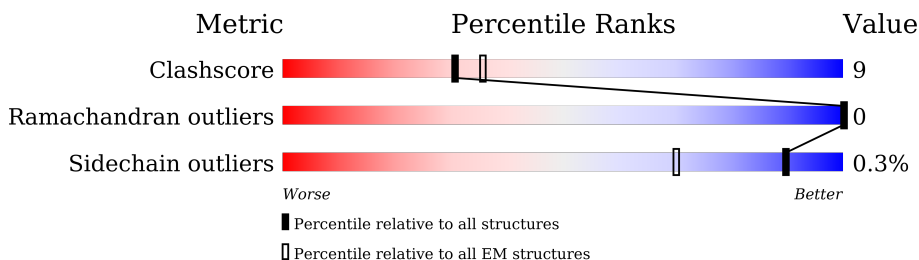|		
EMDB validation analysis	:	FAILED
MolProbity	:	4.02b-467
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ	:	FAILED
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

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 2.97 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		9
Ramachandran outliers		0
Sidechain outliers		0.3%

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	B	358	74% 21% 5%
2	C	58	81% 14% 5%
3	H	250	66% 26% 7%
4	R	510	50% 10% 40%
5	A	354	52% 9% 40%

2   Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 8903 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 Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-1.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	B	339	Total 2588	C 1596	N 461	O 511	S 20	0	0

There are 19 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-17	MET	-	initiating methionine	UNP P62873
B	-16	HIS	-	expression tag	UNP P62873
B	-15	HIS	-	expression tag	UNP P62873
B	-14	HIS	-	expression tag	UNP P62873
B	-13	HIS	-	expression tag	UNP P62873
B	-12	HIS	-	expression tag	UNP P62873
B	-11	HIS	-	expression tag	UNP P62873
B	-10	LEU	-	expression tag	UNP P62873
B	-9	GLU	-	expression tag	UNP P62873
B	-8	VAL	-	expression tag	UNP P62873
B	-7	LEU	-	expression tag	UNP P62873
B	-6	PHE	-	expression tag	UNP P62873
B	-5	GLN	-	expression tag	UNP P62873
B	-4	GLY	-	expression tag	UNP P62873
B	-3	PRO	-	expression tag	UNP P62873
B	-2	GLY	-	expression tag	UNP P62873
B	-1	SER	-	expression tag	UNP P62873
B	0	SER	-	expression tag	UNP P62873
B	1	GLN	-	expression tag	UNP P62873

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

Mol	Chain	Residues	Atoms					AltConf	Trace
2	C	55	Total 414	C 259	N 74	O 79	S 2	0	0

WORLDWIDE
wwPDB
PROTEIN DATA BANK

- Molecule 3 is a protein called scFv16.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	H	232	Total 1774	C 1127	N 291	O 346	S 10	0	0

- Molecule 4 is a protein called Soluble cytochrome b562,Probable G-protein coupled receptor 132.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	R	307	Total 2440	C 1594	N 412	O 415	S 19	0	0

There are 27 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
R	-129	MET	-	initiating methionine	UNP P0ABE7
R	-128	LYS	-	expression tag	UNP P0ABE7
R	-127	THR	-	expression tag	UNP P0ABE7
R	-126	ILE	-	expression tag	UNP P0ABE7
R	-125	ILE	-	expression tag	UNP P0ABE7
R	-124	ALA	-	expression tag	UNP P0ABE7
R	-123	LEU	-	expression tag	UNP P0ABE7
R	-122	SER	-	expression tag	UNP P0ABE7
R	-121	TYR	-	expression tag	UNP P0ABE7
R	-120	ILE	-	expression tag	UNP P0ABE7
R	-119	PHE	-	expression tag	UNP P0ABE7
R	-118	CYS	-	expression tag	UNP P0ABE7
R	-117	LEU	-	expression tag	UNP P0ABE7
R	-116	VAL	-	expression tag	UNP P0ABE7
R	-115	PHE	-	expression tag	UNP P0ABE7
R	-114	ALA	-	expression tag	UNP P0ABE7
R	-113	ASP	-	expression tag	UNP P0ABE7
R	-112	TYR	-	expression tag	UNP P0ABE7
R	-111	LYS	-	expression tag	UNP P0ABE7
R	-110	ASP	-	expression tag	UNP P0ABE7
R	-109	ASP	-	expression tag	UNP P0ABE7
R	-108	ASP	-	expression tag	UNP P0ABE7
R	-107	ASP	-	expression tag	UNP P0ABE7
R	-106	LYS	-	expression tag	UNP P0ABE7
R	-99	TRP	MET	conflict	UNP P0ABE7
R	-4	ILE	HIS	conflict	UNP P0ABE7
R	0	LEU	ARG	conflict	UNP P0ABE7

- Molecule 5 is a protein called Guanine nucleotide-binding protein G(i) subunit alpha-1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	A	214	1687	1073	279	323	12	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	203	ALA	GLY	conflict	UNP P63096
A	326	SER	ALA	conflict	UNP P63096

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

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: Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-1
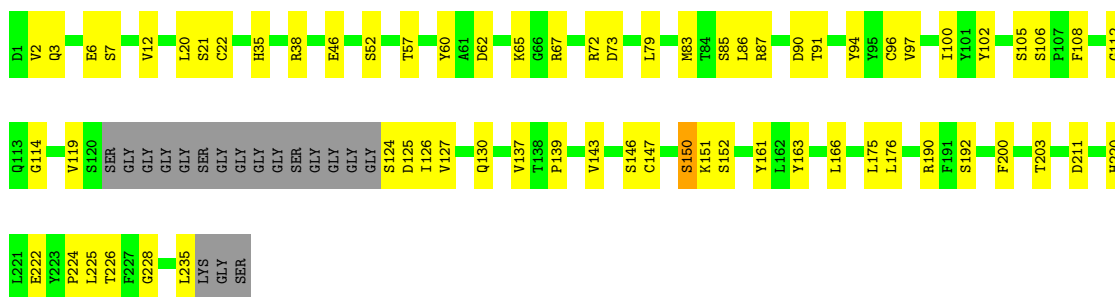
Chain B:
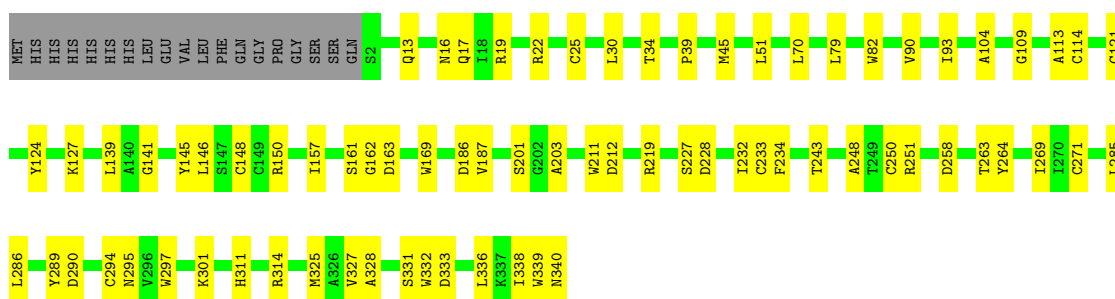
• Molecule 2: Guanine nucleotide-binding protein G(I)/G(S)/G(O) subunit gamma-2
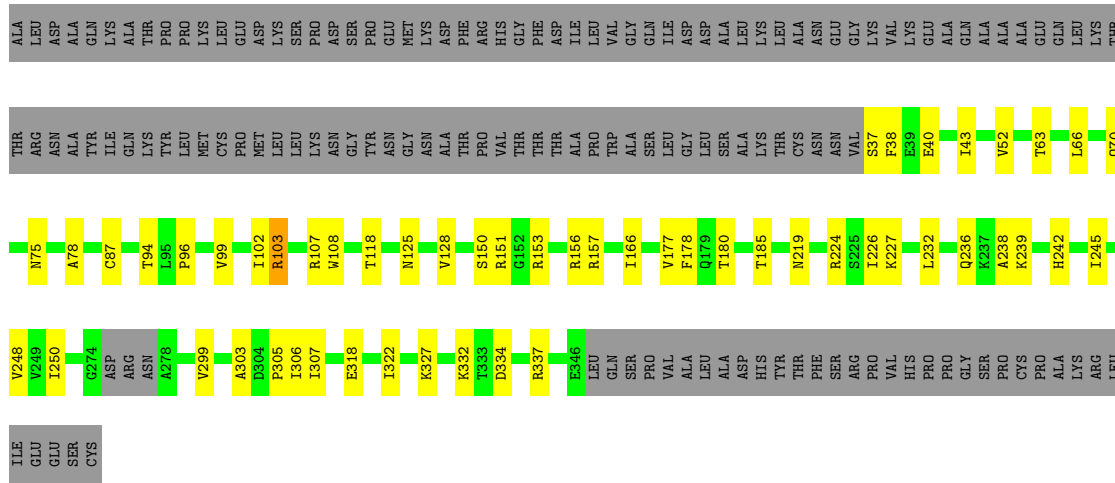
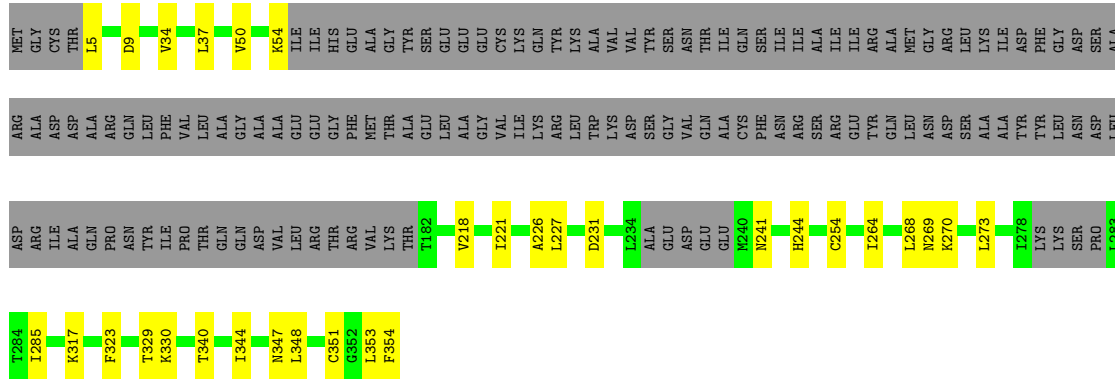Chain C:

• Molecule 3: scFv16

Chain H:

• Molecule 4: Soluble cytochrome b562,Probable G-protein coupled receptor 132

Chain R:

ALA LEU ASP ALA GLN LYS ALA THR PRO PRO LYS LEU GLU ASP LYS SER ASP PRO ASP SER PRO GLU MET LYS ASP PHE ARG HIS GLY PHE ASP ILE LEU VAL GLY GLN ILE ASP ASP ALA LEU LYS LEU ALA ASN GLU GLY LYS VAL LYS GLU ALA GLN ALA ALA ALA GLU GLN LEU LYS THR

THR ARG ASN ALA TYR ILE GLN LYS TYR LEU MET CYS PRO MET LEU LEU LYS ASN GLY TYR ASN ASN ALA THR PRO VAL THR THR THR ALA PRO TRP ALA SER LEU GLY LEU SER ALA LYS THR CYS ASN ASN VAL **S37** **F38** **E39** **E40** **I43** **V52** **T63** **L66** **Q70**

N75 **A78** **C87** **T94** **L95** **P96** **V99** **I102** **R103** **R107** **W108** **T118** **N125** **V128** **S150** **R151** **G152** **R153** **R156** **R157** **I166** **V177** **F178** **Q179** **T180** **T185** **N219** **R224** **S225** **I226** **K227** **L232** **Q236** **K237** **A238** **K239** **H242** **I245**

V248 **V249** **I250** **G274** ASP ARG ASN **A278** **V299** **A303** **D304** **P305** **I306** **I307** **E318** **I322** **K327** **K332** **T333** **D334** **R337** **E346** LEU GLN SER PRO VAL ALA LEU ALA ASP HIS TYR THR PHE SER ARG PRO VAL HIS PRO PRO GLY SER PRO CYS PRO ALA LYS ARG LEU

ILE GLU GLU SER CYS

• Molecule 5: Guanine nucleotide-binding protein G(i) subunit alpha-1

Chain A: 52%   9%   40%

MET GLY CYS THR **L5** **D9** **V34** **L37** **V50** **K54** ILE ILE HIS GLU ALA GLY TYR SER GLU GLU GLU CYS LYS GLN TYR LYS ALA VAL VAL TYR SER ASN THR ILE GLN SER ILE ILE ALA ILE ILE ARG ALA MET GLY ARG LEU LYS ILE ASP PHE GLY ASP SER ALA

ARG ALA ASP ALA ARG GLN LEU PHE VAL LEU ALA GLY ALA ALA GLU GLU GLY PHE MET THR ALA GLU LEU ALA GLY VAL ILE LYS ARG LEU TRP LYS ASP SER GLY VAL GLN ALA CYS PHE ASN ARG SER ARG GLU TYR GLN LEU ASN ASP SER ALA ALA TYR TYR LEU ASN ASP LEU

ASP ARG ILE ALA GLN PRO ASN TYR ILE PRO THR GLN GLN ASP VAL LEU ARG THR ARG VAL LYS THR **T182** **V218** **I221** **A226** **L227** **D231** **L234** ALA GLU ASP GLU GLU **M240** **N241** **H244** **C254** **I264** **L268** **N269** **K270** **L273** **I278** LYS LYS SER PRO **L283**

T284 **I285** **K317** **F323** **T329** **K330** **T340** **I344** **N347** **L348** **C351** **G352** **L353** **F354**

4 Experimental information ⓘ

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	363528	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 $(e^-/\text{Å}^2)$	60	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	2500	Depositor
Magnification	Not provided	
Image detector	GATAN K2 SUMMIT (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	$\#|Z|>5$	RMSZ	$\#|Z|>5$
1	B	0.26	0/2635	0.60	0/3576
2	C	0.27	0/420	0.53	0/568
3	H	0.29	0/1818	0.59	0/2467
4	R	0.29	0/2494	0.60	0/3389
5	A	0.27	0/1712	0.50	0/2301
All	All	0.28	0/9079	0.58	0/12301

There are no bond length outliers.

There are no bond angle outliers.

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	B	2588	0	2470	49	0
2	C	414	0	414	8	0
3	H	1774	0	1699	40	0
4	R	2440	0	2508	40	0
5	A	1687	0	1642	27	0
All	All	8903	0	8733	153	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 153 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)
4:R:239:LYS:HZ1	5:A:354:PHE:HB2	1.17	1.10
5:A:54:LYS:HA	5:A:54:LYS:HE2	1.36	1.08
4:R:239:LYS:NZ	5:A:354:PHE:HB2	1.69	1.06
3:H:146:SER:HB2	3:H:200:PHE:O	1.73	0.86
5:A:54:LYS:HE2	5:A:54:LYS:CA	2.00	0.85

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	B	337/358 (94%)	324 (96%)	13 (4%)	0	100	100
2	C	53/58 (91%)	49 (92%)	4 (8%)	0	100	100
3	H	228/250 (91%)	209 (92%)	19 (8%)	0	100	100
4	R	303/510 (59%)	290 (96%)	13 (4%)	0	100	100
5	A	206/354 (58%)	203 (98%)	3 (2%)	0	100	100
All	All	1127/1530 (74%)	1075 (95%)	52 (5%)	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	B	278/299 (93%)	278 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	C	42/47 (89%)	42 (100%)	0	100	100
3	H	194/202 (96%)	193 (100%)	1 (0%)	86	94
4	R	265/435 (61%)	263 (99%)	2 (1%)	79	90
5	A	181/306 (59%)	181 (100%)	0	100	100
All	All	960/1289 (74%)	957 (100%)	3 (0%)	90	96

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	H	150	SER
4	R	103	ARG
4	R	178	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 (i)

There are no chain breaks in this entry.