wwPDB EM Validation Summary Report (i)

Oct 22, 2024 – 12:35 AM JST

PDB ID	:	8XVI
EMDB ID	:	EMD-38705
Title	:	Cryo-EM structure of ETAR bound with Endothelin1
Authors	:	Hou, J.Y.; Liu, S.H.; Wu, L.J.; Liu, Z.J.; Hua, T.
Deposited on	:	2024-01-15
Resolution	:	3.32 Å(reported)

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 3.32 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		4.0%

Worse *Better*

▌ Percentile relative to all structures
▯ Percentile relative to all EM structures
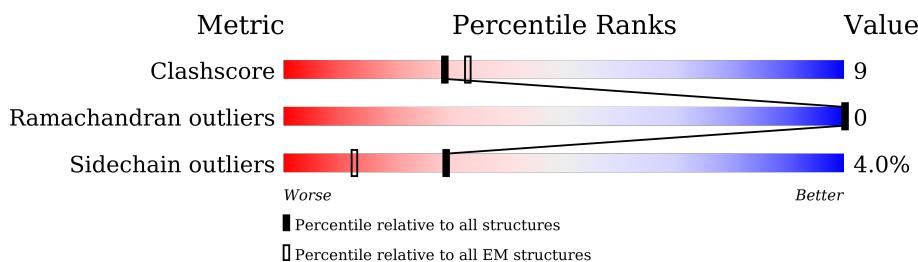
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415

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	261	69% 21% 10%
2	B	346	72% 24% • •
3	G	71	68% 13% 20%
4	N	157	57% 22% • 20%
5	R	683	36% 9% 55%
6	T	21	71% 24% 5%

2 Entry composition ⓘ

There are 6 unique types of molecules in this entry. The entry contains 8587 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 Isoform Gnas-2 of Guanine nucleotide-binding protein G(s) subunit alpha isoforms short.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	234	Total	C	N	O	S	0	0
			1943	1227	349	360	7		

- Molecule 2 is a protein called Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-1.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	338	Total	C	N	O	S	0	0
			2600	1604	467	508	21		

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-5	ILE	-	expression tag	UNP P62873
B	-4	GLY	-	expression tag	UNP P62873
B	-3	ARG	-	expression tag	UNP P62873
B	-2	ALA	-	expression tag	UNP P62873
B	-1	ARG	-	expression tag	UNP P62873
B	0	GLY	-	expression tag	UNP P62873
B	1	PHE	-	expression tag	UNP P62873

- Molecule 3 is a protein called Guanine nucleotide-binding protein G(I)/G(S)/G(O) subunit gamma-2.

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

- Molecule 4 is a protein called Nanobody 35.

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

- Molecule 5 is a protein called Endoglucanase H,Endothelin-1 receptor.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	R	308	2476	1637	398	416	25	0	0

There are 49 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
R	-259	MET	-	initiating methionine	UNP P16218
R	-258	LYS	-	expression tag	UNP P16218
R	-257	THR	-	expression tag	UNP P16218
R	-256	ILE	-	expression tag	UNP P16218
R	-255	ILE	-	expression tag	UNP P16218
R	-254	ALA	-	expression tag	UNP P16218
R	-253	LEU	-	expression tag	UNP P16218
R	-252	SER	-	expression tag	UNP P16218
R	-251	TYR	-	expression tag	UNP P16218
R	-250	ILE	-	expression tag	UNP P16218
R	-249	PHE	-	expression tag	UNP P16218
R	-248	CYS	-	expression tag	UNP P16218
R	-247	LEU	-	expression tag	UNP P16218
R	-246	VAL	-	expression tag	UNP P16218
R	-245	PHE	-	expression tag	UNP P16218
R	-244	ALA	-	expression tag	UNP P16218
R	-243	ASP	-	expression tag	UNP P16218
R	-242	TYR	-	expression tag	UNP P16218
R	-241	LYS	-	expression tag	UNP P16218
R	-240	ASP	-	expression tag	UNP P16218
R	-239	ASP	-	expression tag	UNP P16218
R	-238	ASP	-	expression tag	UNP P16218
R	-237	ASP	-	expression tag	UNP P16218
R	-236	ALA	-	expression tag	UNP P16218
R	-235	GLY	-	expression tag	UNP P16218
R	-234	ARG	-	expression tag	UNP P16218
R	-233	ALA	-	expression tag	UNP P16218
R	-232	MET	-	expression tag	UNP P16218
R	-231	ALA	-	expression tag	UNP P16218
R	-230	SER	-	expression tag	UNP P16218
R	-124	ALA	GLU	conflict	UNP P16218

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
R	406	LEU	-	expression tag	UNP P25101
R	407	GLU	-	expression tag	UNP P25101
R	408	VAL	-	expression tag	UNP P25101
R	409	LEU	-	expression tag	UNP P25101
R	410	PHE	-	expression tag	UNP P25101
R	411	GLN	-	expression tag	UNP P25101
R	412	GLY	-	expression tag	UNP P25101
R	413	PRO	-	expression tag	UNP P25101
R	414	HIS	-	expression tag	UNP P25101
R	415	HIS	-	expression tag	UNP P25101
R	416	HIS	-	expression tag	UNP P25101
R	417	HIS	-	expression tag	UNP P25101
R	418	HIS	-	expression tag	UNP P25101
R	419	HIS	-	expression tag	UNP P25101
R	420	HIS	-	expression tag	UNP P25101
R	421	HIS	-	expression tag	UNP P25101
R	422	HIS	-	expression tag	UNP P25101
R	423	HIS	-	expression tag	UNP P25101

- Molecule 6 is a protein called Endothelin-1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	T	21	171	109	25	32	5	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.
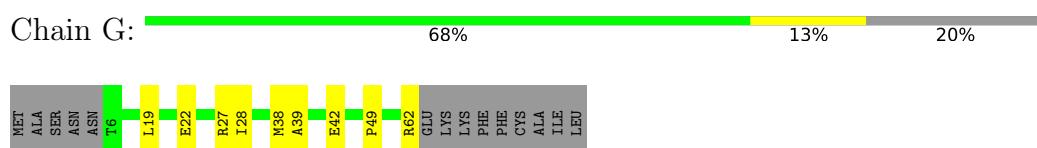
• Molecule 1: Isoform Gnas-2 of Guanine nucleotide-binding protein G(s) subunit alpha isoforms short
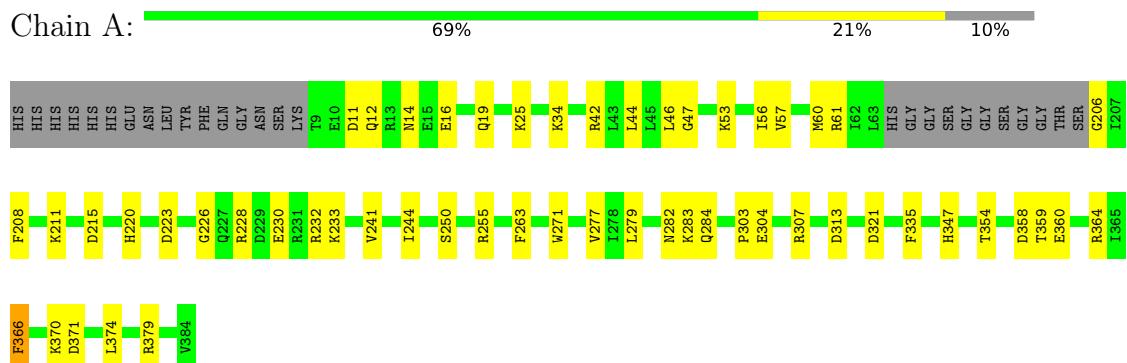
Chain A:

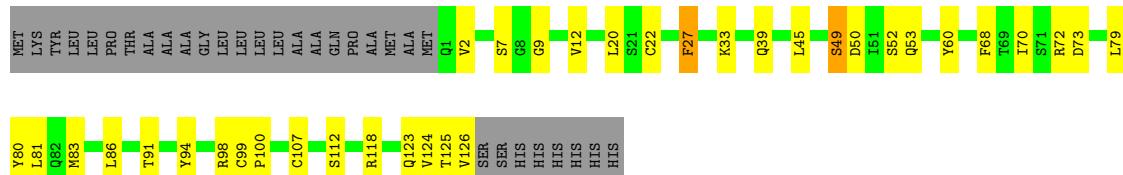• Molecule 2: Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-1

Chain B:

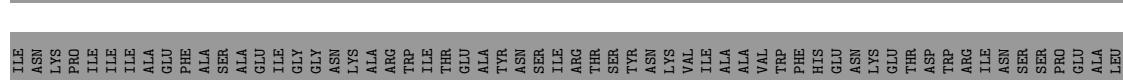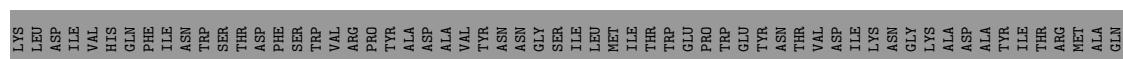• Molecule 3: Guanine nucleotide-binding protein G(I)/G(S)/G(O) subunit gamma-2

Chain G:

• Molecule 4: Nanobody 35

Chain N:

MET LYS TYR LEU PRO THR ALA ALA GLY LEU LEU LEU LEU ALA ALA GLN PRO ALA MET ALA MET Q1 V2 S7 G8 G9 V12 L20 S21 C22 F27 K33 Q39 L45 S49 D50 I51 S52 Q53 Y60 F68 T69 I70 S71 R72 D73 L79

Y80 L81 Q82 M83 L86 T91 Y94 R98 C99 P100 C107 S112 R118 Q123 V124 T125 V126 SER HIS HIS HIS HIS HIS

• Molecule 5: Endoglucanase H,Endothelin-1 receptor

Chain R: 36% 9% 55%

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

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

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

PHE ASN VAL ASN CYS ASP VAL GLY ASN GLY ASN GLY THR SER TYR LEU GLY HIS TYR PRO GLY ASP ALA ASN TYR VAL ASP TYR THR SER ILE ASP GLY TYR ASN THR GLY GLN THR THR GLN SER TRP GLY SER GLN ASP GLN VAL PHE SER ALA ARG ALA TYR GLN LEU ALA SER

ILE ASN LYS PRO ILE ILE ALA GLU GLU PHE ALA SER ALA GLU ILE GLY GLY ASN LYS ALA ARG TRP ILE THR GLU ALA TYR ASN GLY LYS VAL ILE ALA ALA VAL TRP PHE HIS GLU LYS LYS THR ASP TRP ASP ARG ILE ASN SER SER GLU ALA LEU

ALA ALA TYR ARG GLU ALA ILE GLY ALA THR THR HIS GLN PRO THR ASN LEU VAL LEU PRO SER ASN GLY SER MET HIS N67 Y68 F79 S87 N98 I104 M111 R112 N113 N116 D126 V130 D133 L134 P135 I136 N137 V138 L141 D149 F153

F161 P162 F163 L164 Q165 L176 R183 Y184 V187 W190 I198 P219 E220 A221 F224 V225 K237 T238 C239 M240 L241 F247 V254 K255 L259 F260 F264 Y275 M282 L283 N284 ARG ARG ASN GLY SER LEU ARG ILE A293 R301 V308 I314

H323 Y333 D337 K338 N339 R340 L349 Y352 L357 M360 N361 I364 F376 C379 S382 CYS LEU CYS CYS CYS CYS TYR GLN SER LYS SER LEU MET THR SER VAL PRO MET ASN GLY THR SER ILE LEU GLU VAL LEU PHE GLN GLY PRO HIS HIS

HIS HIS HIS HIS HIS HIS

• Molecule 6: Endothelin-1

Chain T: 71% 24% 5%

C1 S5 L6 M7 H16 L17 D18 I19 I20 W21

4 Experimental information ⓘ

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	114666	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 $(e^-/\text{Å}^2)$	60	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	2000	Depositor
Magnification	Not provided	
Image detector	FEI FALCON IV (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	$\#|Z| > 5$	RMSZ	$\#|Z| > 5$
1	A	0.29	0/1979	0.54	1/2663 (0.0%)
2	B	0.26	0/2647	0.56	1/3589 (0.0%)
3	G	0.24	0/442	0.44	0/597
4	N	0.30	0/981	0.57	0/1329
5	R	0.29	0/2537	0.47	0/3440
6	T	0.26	0/175	0.58	0/234
All	All	0.28	0/8761	0.53	2/11852 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	298	ASP	CB-CG-OD2	5.48	123.23	118.30
1	A	371	ASP	CB-CG-OD1	5.15	122.94	118.30

There are no chirality outliers.

There are no planarity outliers.

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	1943	0	1926	36	0
2	B	2600	0	2505	62	0
3	G	436	0	448	7	0
4	N	961	0	928	23	0
5	R	2476	0	2541	39	0

Continued on next page...

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

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	T	171	0	157	5	0
All	All	8587	0	8505	158	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 158 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:N:70:ILE:HG22	4:N:81:LEU:HD13	1.62	0.79
4:N:9:GLY:HA2	4:N:124:VAL:HG12	1.68	0.74
1:A:279:LEU:HD11	1:A:335:PHE:HD2	1.53	0.72
2:B:49:ARG:NH1	2:B:85:TYR:O	2.23	0.71
5:R:111:MET:O	5:R:116:ASN:ND2	2.22	0.71

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	230/261 (88%)	223 (97%)	7 (3%)	0	100	100
2	B	336/346 (97%)	312 (93%)	24 (7%)	0	100	100
3	G	55/71 (78%)	54 (98%)	1 (2%)	0	100	100
4	N	124/157 (79%)	119 (96%)	5 (4%)	0	100	100
5	R	304/683 (44%)	289 (95%)	15 (5%)	0	100	100
6	T	19/21 (90%)	18 (95%)	1 (5%)	0	100	100
All	All	1068/1539 (69%)	1015 (95%)	53 (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	A	211/231 (91%)	205 (97%)	6 (3%)	38	64
2	B	281/286 (98%)	265 (94%)	16 (6%)	17	45
3	G	46/58 (79%)	45 (98%)	1 (2%)	47	70
4	N	104/127 (82%)	98 (94%)	6 (6%)	17	44
5	R	276/591 (47%)	268 (97%)	8 (3%)	37	64
6	T	21/21 (100%)	20 (95%)	1 (5%)	21	50
All	All	939/1314 (72%)	901 (96%)	38 (4%)	29	54

5 of 38 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	N	123	GLN
5	R	361	ASN
5	R	68	TYR
5	R	241	LEU
6	T	7	MET

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

Mol	Chain	Res	Type
2	B	220	GLN
5	R	165	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.