wwPDB EM Validation Summary Report (i)

Nov 14, 2024 – 10:07 PM JST

PDB ID	:	8XVK
EMDB ID	:	EMD-38707
Title	:	Cryo-EM structure of ETAR bound with Ambrisentan
Authors	:	Hou, J.Y.; Liu, S.H.; Wu, L.J.; Liu, Z.J.; Hua, T.
Deposited on	:	2024-01-15
Resolution	:	3.21 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:

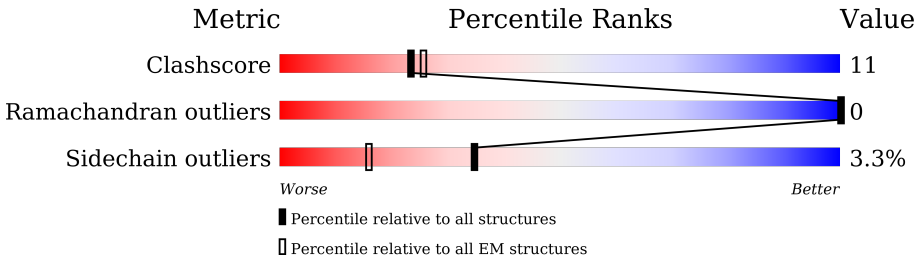
EMDB validation analysis	:	FAILED
Mogul	:	1.8.5 (274361), CSD as541be (2020)
MolProbity	:	4.02b-467
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ	:	FAILED
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 3.21 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 <=5%

Mol	Chain	Length	Quality of chain
1	H	266	61% 23% • 15%
2	K	136	57% 29% • 12%
3	L	235	74% 15% • 10%
4	R	783	38% 12% • 50%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 7388 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 anti-BRIL Fab Heavy chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	H	225	Total	C	N	O	S	0	0
			1682	1069	279	329	5		

- Molecule 2 is a protein called anti-Fab Nanobody.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	K	120	Total	C	N	O	S	0	0
			921	571	162	184	4		

- Molecule 3 is a protein called anti-BRIL Fab Light chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	L	212	Total	C	N	O	S	0	0
			1624	1019	270	330	5		

- Molecule 4 is a protein called Endoglucanase H,Endothelin-1 receptor,Soluble cytochrome b562.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	R	392	Total	C	N	O	S	0	0
			3133	2044	519	546	24		

There are 63 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
R	-276	MET	-	initiating methionine	UNP P16218
R	-275	LYS	-	expression tag	UNP P16218
R	-274	THR	-	expression tag	UNP P16218
R	-273	ILE	-	expression tag	UNP P16218
R	-272	ILE	-	expression tag	UNP P16218
R	-271	ALA	-	expression tag	UNP P16218
R	-270	LEU	-	expression tag	UNP P16218
R	-269	SER	-	expression tag	UNP P16218

Continued on next page...

Continued from previous page...

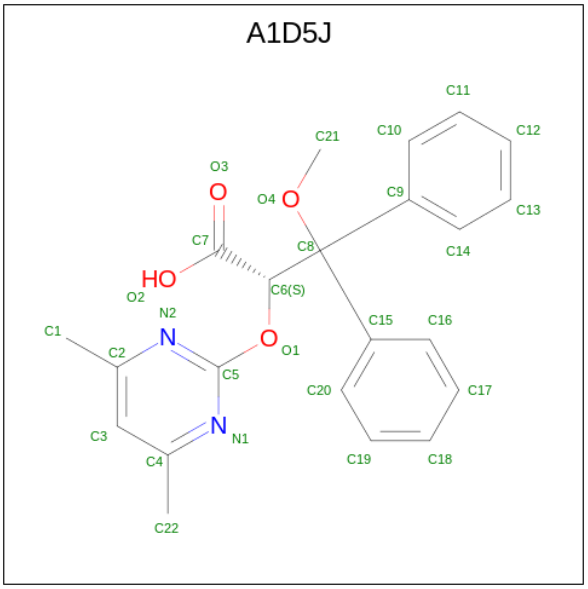
Chain	Residue	Modelled	Actual	Comment	Reference
R	-268	TYR	-	expression tag	UNP P16218
R	-267	ILE	-	expression tag	UNP P16218
R	-266	PHE	-	expression tag	UNP P16218
R	-265	CYS	-	expression tag	UNP P16218
R	-264	LEU	-	expression tag	UNP P16218
R	-263	VAL	-	expression tag	UNP P16218
R	-262	PHE	-	expression tag	UNP P16218
R	-261	ALA	-	expression tag	UNP P16218
R	-260	ASP	-	expression tag	UNP P16218
R	-259	TYR	-	expression tag	UNP P16218
R	-258	LYS	-	expression tag	UNP P16218
R	-257	ASP	-	expression tag	UNP P16218
R	-256	ASP	-	expression tag	UNP P16218
R	-255	ASP	-	expression tag	UNP P16218
R	-254	ASP	-	expression tag	UNP P16218
R	-253	ALA	-	expression tag	UNP P16218
R	-252	HIS	-	expression tag	UNP P16218
R	-251	HIS	-	expression tag	UNP P16218
R	-250	HIS	-	expression tag	UNP P16218
R	-249	HIS	-	expression tag	UNP P16218
R	-248	HIS	-	expression tag	UNP P16218
R	-247	HIS	-	expression tag	UNP P16218
R	-246	HIS	-	expression tag	UNP P16218
R	-245	HIS	-	expression tag	UNP P16218
R	-244	HIS	-	expression tag	UNP P16218
R	-243	HIS	-	expression tag	UNP P16218
R	-242	GLY	-	expression tag	UNP P16218
R	-241	ARG	-	expression tag	UNP P16218
R	-240	ALA	-	expression tag	UNP P16218
R	-239	MET	-	expression tag	UNP P16218
R	-238	ALA	-	expression tag	UNP P16218
R	-237	SER	-	expression tag	UNP P16218
R	-131	ALA	GLU	conflict	UNP P16218
R	43	GLU	-	linker	UNP P16218
R	44	ASN	-	linker	UNP P16218
R	45	LEU	-	linker	UNP P16218
R	46	TYR	-	linker	UNP P16218
R	47	PHE	-	linker	UNP P16218
R	48	GLN	-	linker	UNP P16218
R	49	GLY	-	linker	UNP P16218
R	281	GLU	-	linker	UNP P25101
R	1001	ALA	-	linker	UNP P25101

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
R	1002	ARG	-	linker	UNP P25101
R	1003	ARG	-	linker	UNP P25101
R	1004	GLN	-	linker	UNP P25101
R	1005	LEU	-	linker	UNP P25101
R	1012	TRP	MET	conflict	UNP P0ABE7
R	1107	ILE	HIS	conflict	UNP P0ABE7
R	1111	LEU	-	linker	UNP P0ABE7
R	1112	GLU	-	linker	UNP P0ABE7
R	1113	ARG	-	linker	UNP P0ABE7
R	1114	ALA	-	linker	UNP P0ABE7
R	1115	ARG	-	linker	UNP P0ABE7
R	1116	SER	-	linker	UNP P0ABE7
R	1117	THR	-	linker	UNP P0ABE7

- Molecule 5 is Ambrisentan (three-letter code: A1D5J) (formula: $C_{22}H_{22}N_2O_4$).



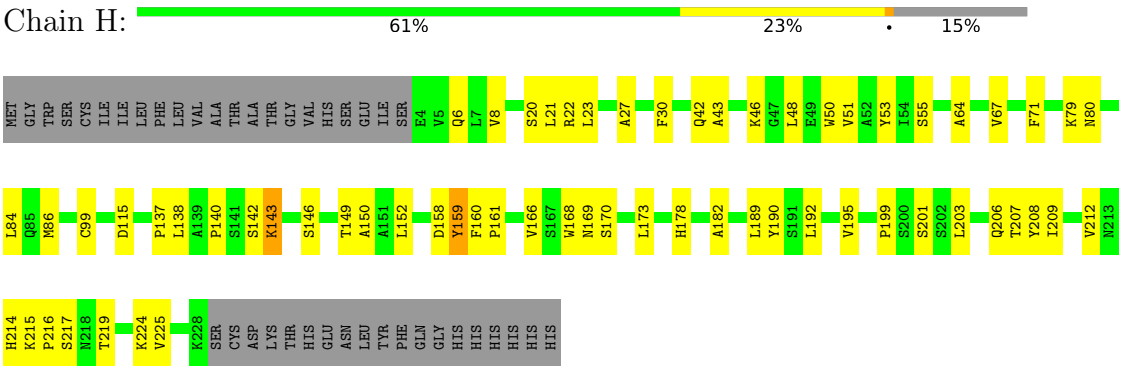
Mol	Chain	Residues	Atoms				AltConf
			Total	C	N	O	
5	R	1	28	22	2	4	0

3 Residue-property plots (i)

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.

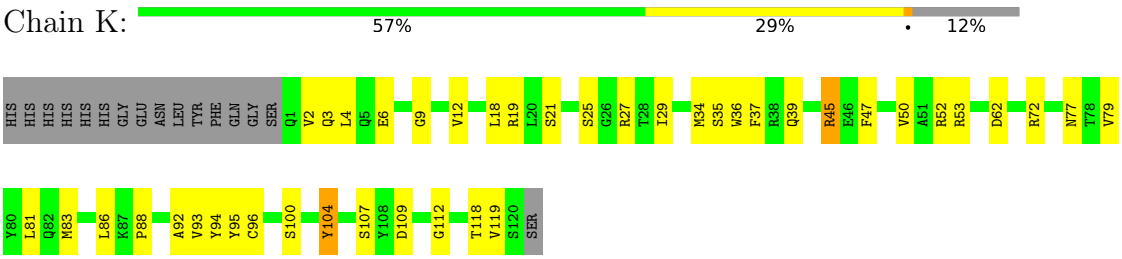
• Molecule 1: anti-BRIL Fab Heavy chain

Chain H:



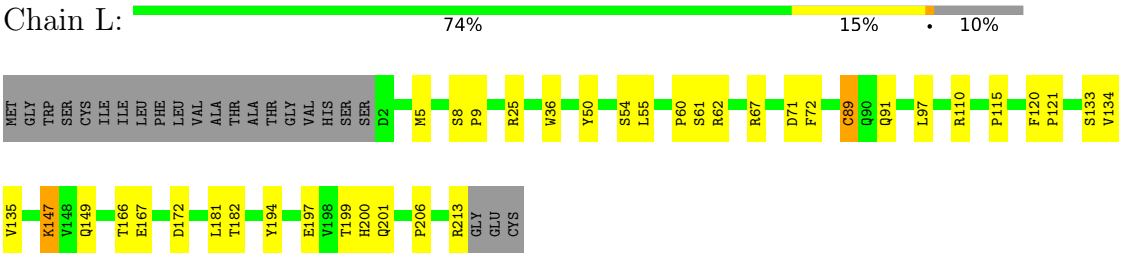
• Molecule 2: anti-Fab Nanobody

Chain K:



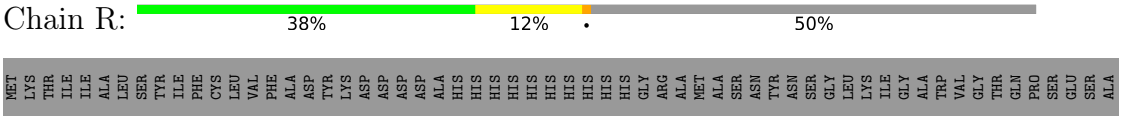
• Molecule 3: anti-BRIL Fab Light chain

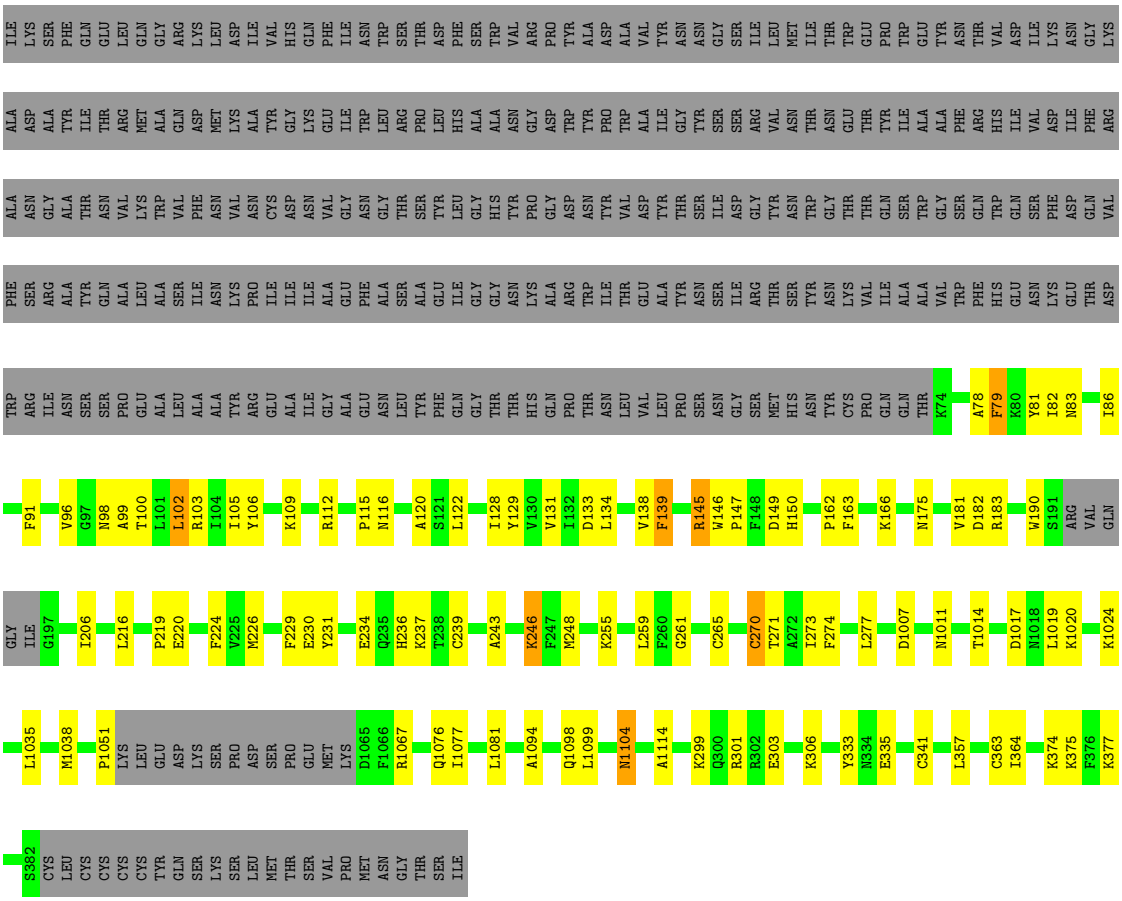
Chain L:



• Molecule 4: Endoglucanase H,Endothelin-1 receptor,Soluble cytochrome b562

Chain R:




ILE LYS SER PHE GLN GLU LEU GLN GLY ARG LYS LEU ASP ILE VAL HIS GLN PHE ILE ASN TRP SER THR ASP PHE SER TRP VAL ARG PRO TYR ALA ASP ALA VAL TYR ASN ASN GLY SER ILE LEU MET ILE 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 PRO TRP ALA ILE GLY TYR SER SER ARG VAL ASN THR ASN GLU THR 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 ASN VAL GLY ASN GLY THR SER TYR LEU GLY HIS TYR PRO GLY ASP ASN TYR SER ILE ASP GLY TYR ASN TRP GLY THR THR GLN SER TRP GLY SER GLN TRP GLN SER PHE ASP GLN VAL
PHE SER ARG ALA TYR GLN ALA LEU ALA SER ILE ASN LYS PRO ILE ILE ILE ALA GLU PHE ALA SER ALA GLU ILE GLY GLY ASN LEU TYR PHE GLN GLY THR THR HIS GLN PRO THR ASN LEU VAL LEU PRO SER ASN GLY SER MET HIS ASN TYR CYS PRO GLN GLN THR K74 A78 F79 K80 Y81 I82 N83 I86
F91 V96 G97 N98 A99 T100 L101 L102 R103 I104 I105 Y106 K109 R112 P115 N116 A120 S121 L122 I128 Y129 V130 V131 I132 D133 L134 V138 F139 R145 W146 P147 F148 D149 H150 P162 F163 K166 N175 V181 D182 R183 W190 S191 ARG VAL GLN
GLY ILE G197 I206 L216 P219 E220 F224 V225 M226 F229 E230 Y231 E234 Q235 H236 K237 T238 C239 A243 K246 F247 M248 K255 L259 F260 G261 C265 C270 T271 K272 I273 F274 L277 D1007 N1011 T1014 D1017 N1018 L1019 K1020 K1024
L1035 M1038 P1051 LYS LEU GLU ASP LYS SER PRO ASP SER PRO GLU MET LYS D1065 F1066 R1067 Q1076 I1077 L1081 A1094 Q1098 L1099 N1104 A1114 K299 Q300 R301 R302 E303 K306 Y333 N334 E335 C341 L357 C363 I364 K374 K375 F376 K377
S382 CYS LEU CYS CYS CYS CYS TYR GLN SER LYS SER LEU MET THR SER VAL PRO MET ASN GLY THR SER ILE

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	568027	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 ($e^-/Å^2$)	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

Bond lengths and bond angles in the following residue types are not validated in this section: A1D5J

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	H	0.27	0/1730	0.50	0/2365
2	K	0.28	0/940	0.57	0/1272
3	L	0.28	0/1659	0.50	0/2256
4	R	0.28	0/3199	0.48	1/4330 (0.0%)
All	All	0.28	0/7528	0.50	1/10223 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	R	219	PRO	CA-N-CD	-5.19	104.23	111.50

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	H	1682	0	1619	44	0
2	K	921	0	862	31	0
3	L	1624	0	1573	24	0
4	R	3133	0	3211	61	0
5	R	28	0	0	0	0
All	All	7388	0	7265	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 11.

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)
3:L:97:LEU:HD12	3:L:97:LEU:O	1.60	1.00
2:K:34:MET:HG2	2:K:79:VAL:HG11	1.63	0.80
1:H:158:ASP:HA	1:H:189:LEU:HD13	1.70	0.73
4:R:115:PRO:HG3	4:R:183:ARG:HH22	1.57	0.68
4:R:226:MET:HG2	4:R:239:CYS:HB3	1.76	0.67

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	H	223/266 (84%)	206 (92%)	17 (8%)	0	100	100
2	K	118/136 (87%)	109 (92%)	9 (8%)	0	100	100
3	L	210/235 (89%)	195 (93%)	15 (7%)	0	100	100
4	R	386/783 (49%)	372 (96%)	14 (4%)	0	100	100
All	All	937/1420 (66%)	882 (94%)	55 (6%)	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	H	185/222 (83%)	179 (97%)	6 (3%)	34	63
2	K	95/110 (86%)	93 (98%)	2 (2%)	48	72
3	L	186/206 (90%)	183 (98%)	3 (2%)	58	78
4	R	340/671 (51%)	324 (95%)	16 (5%)	22	53
All	All	806/1209 (67%)	779 (97%)	27 (3%)	35	62

5 of 27 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	R	139	PHE
4	R	182	ASP
4	R	306	LYS
4	R	163	PHE
4	R	224	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	H	6	GLN
1	H	214	HIS
3	L	201	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

There are no oligosaccharides in this entry.

5.6 Ligand geometry (i)

1 ligand is modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
5	A1D5J	R	1201	-	29,30,30	0.75	0	32,42,42	0.51	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	A1D5J	R	1201	-	-	7/29/29/29	0/3/3/3

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 7 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	R	1201	A1D5J	C8-C6-C7-O2
5	R	1201	A1D5J	C6-C8-O4-C21
5	R	1201	A1D5J	C16-C15-C8-C9
5	R	1201	A1D5J	C20-C15-C8-C6
5	R	1201	A1D5J	C20-C15-C8-C9

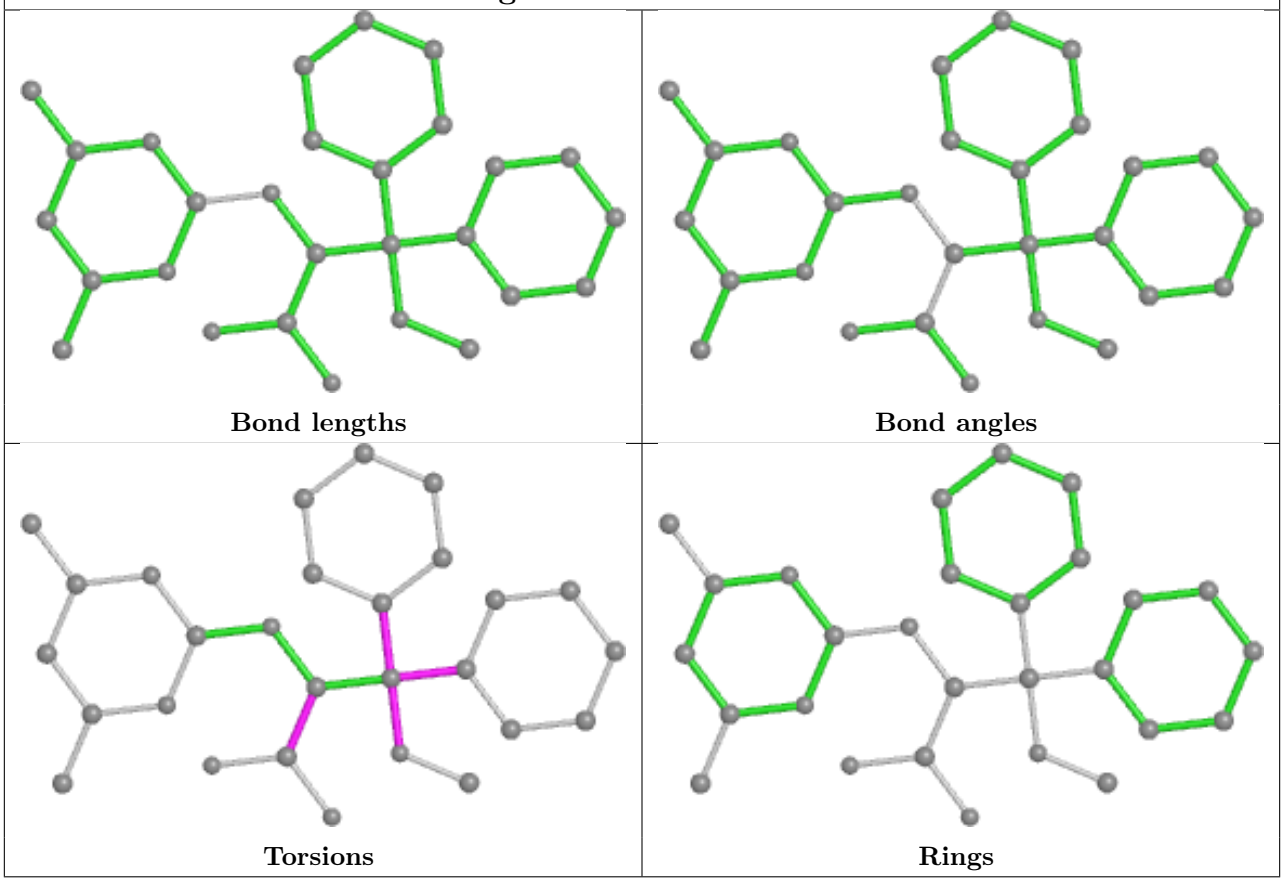
There are no ring outliers.

No monomer is involved in short contacts.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and

any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand A1D5J R 1201



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.