wwPDB EM Validation Summary Report

Oct 28, 2024 – 01:29 PM JST

PDB ID	:	8YUV
EMDB ID	:	EMD-39584
Title	:	Cryo-EM structure of the immepip-bound H3R-Gi complex
Authors	:	Shen, Q.; Tang, X.; Wen, X.; Cheng, S.; Xiao, P.; Zang, S.; Shen, D.; Jiang, L.; Zheng, Y.; Zhang, H.; Xu, H.; Mao, C.; Zhang, M.; Hu, W.; Sun, J.; Chen, Z.; Zhang, Y.
Deposited on	:	2024-03-27
Resolution	:	3.00 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 (1)) were used in the production of this report:

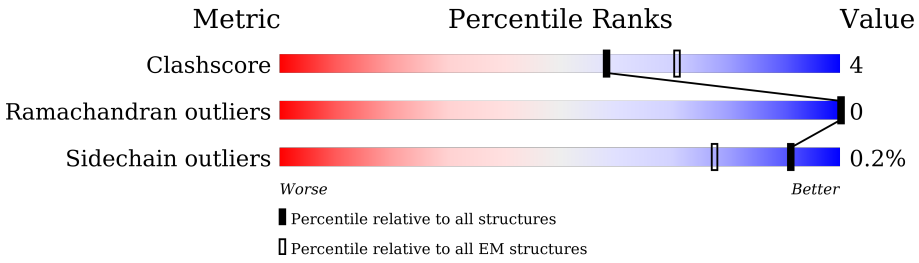
EMDB validation analysis	:	FAILED
Mogul	:	1.8.5 (274361), CSD as541be (2020)
MolProbity	:	4.02b-467
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ	:	FAILED
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, 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 3.00 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	A	354	61% · 36%
2	B	358	90% · 5%
3	G	71	79% · 18%
4	S	269	81% 6% 13%
5	R	461	53% 10% 37%

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 9102 atoms, of which 61 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) subunit alpha-1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	225	1811	1150	301	346	14	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	47	ASN	SER	conflict	UNP P63096
A	203	ALA	GLY	conflict	UNP P63096
A	245	ALA	GLU	conflict	UNP P63096
A	326	SER	ALA	conflict	UNP P63096

- 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
			Total	C	N	O	S		
2	B	339	2607	1607	468	511	21	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

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
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	GLY	-	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	58	Total	C	N	O	S	0	0
			444	277	79	85	3		

- Molecule 4 is a protein called scFv16.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	S	234	Total	C	N	O	S	0	0
			1795	1137	297	351	10		

- Molecule 5 is a protein called Histamine H3 receptor.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	R	289	Total	C	N	O	S	0	0
			2344	1559	389	384	12		

There are 16 discrepancies between the modelled and reference sequences:

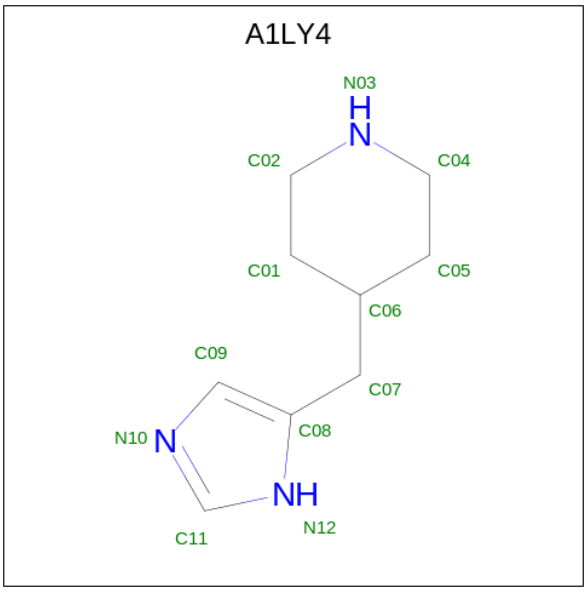
Chain	Residue	Modelled	Actual	Comment	Reference
R	-15	ASP	-	expression tag	UNP Q9Y5N1
R	-14	TYR	-	expression tag	UNP Q9Y5N1
R	-13	LYS	-	expression tag	UNP Q9Y5N1
R	-12	ASP	-	expression tag	UNP Q9Y5N1
R	-11	ASP	-	expression tag	UNP Q9Y5N1
R	-10	ASP	-	expression tag	UNP Q9Y5N1
R	-9	ASP	-	expression tag	UNP Q9Y5N1
R	-8	LYS	-	expression tag	UNP Q9Y5N1
R	-7	LEU	-	expression tag	UNP Q9Y5N1
R	-6	GLU	-	expression tag	UNP Q9Y5N1
R	-5	VAL	-	expression tag	UNP Q9Y5N1
R	-4	LEU	-	expression tag	UNP Q9Y5N1

Continued on next page...

Continued from previous page...

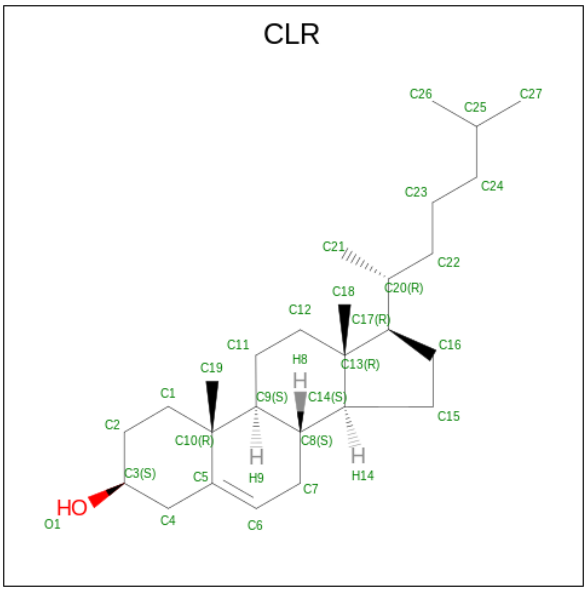
Chain	Residue	Modelled	Actual	Comment	Reference
R	-3	PHE	-	expression tag	UNP Q9Y5N1
R	-2	GLN	-	expression tag	UNP Q9Y5N1
R	-1	GLY	-	expression tag	UNP Q9Y5N1
R	0	PRO	-	expression tag	UNP Q9Y5N1

- Molecule 6 is 4-(1H-imidazol-5-ylmethyl)piperidine (three-letter code: A1LY4) (formula: $C_9H_{15}N_3$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				AltConf
			Total	C	H	N	
6	R	1	27	9	15	3	0

- Molecule 7 is CHOLESTEROL (three-letter code: CLR) (formula: $C_{27}H_{46}O$).



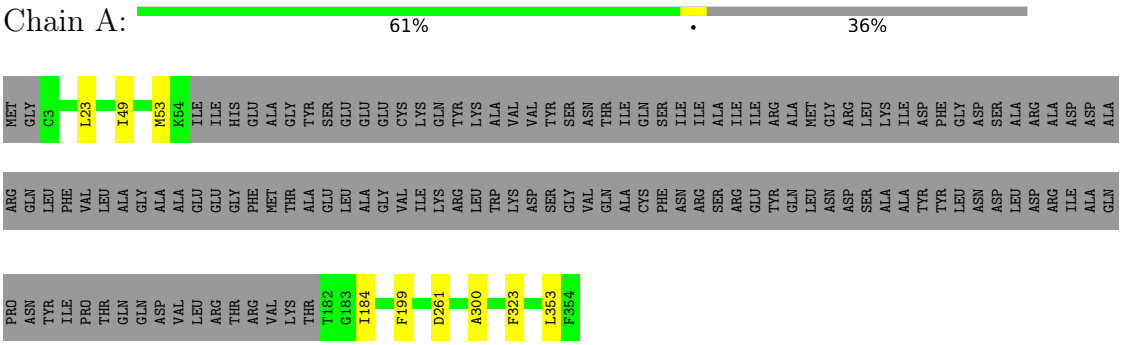
Mol	Chain	Residues	Atoms				AltConf
7	R	1	Total	C	H	O	0
			74	27	46	1	

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.

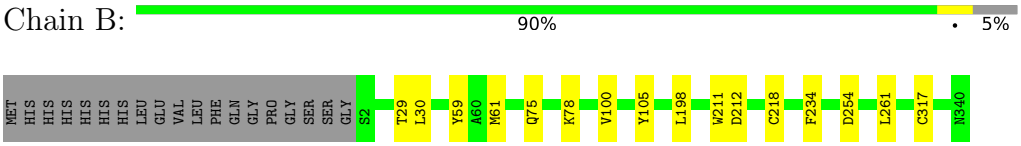
• Molecule 1: Guanine nucleotide-binding protein G(i) subunit alpha-1

Chain A:



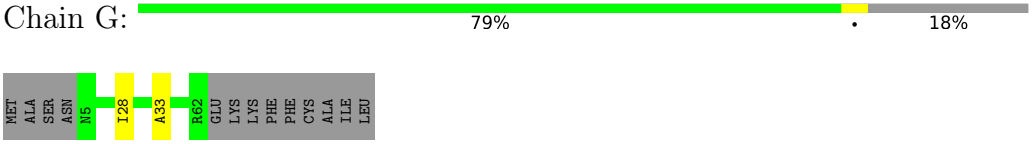
• Molecule 2: Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-1

Chain B:



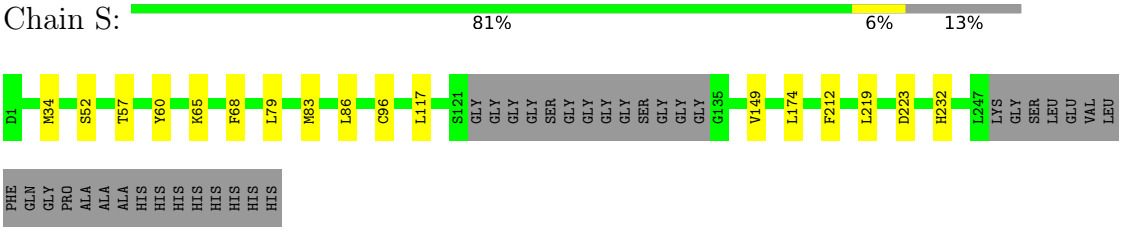
• Molecule 3: Guanine nucleotide-binding protein G(I)/G(S)/G(O) subunit gamma-2

Chain G:



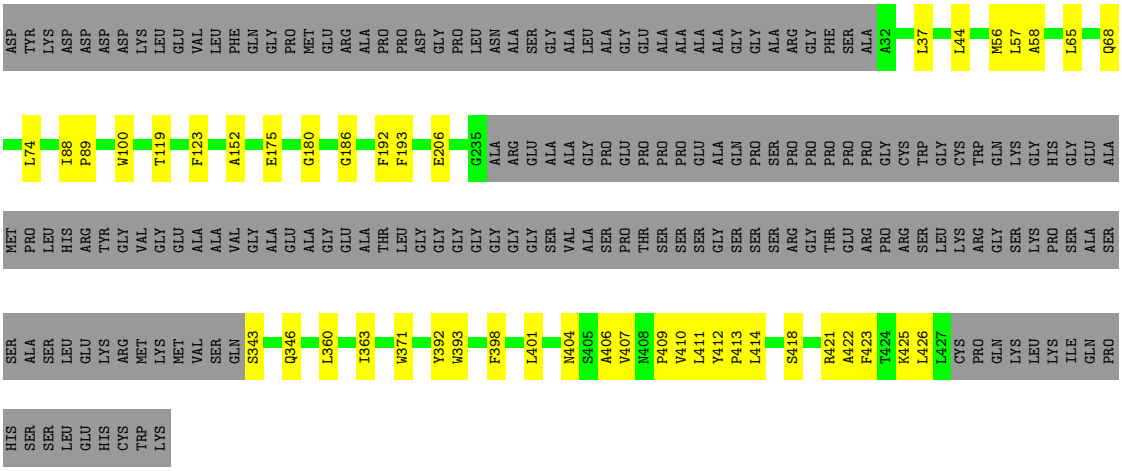
• Molecule 4: scFv16

Chain S:



• Molecule 5: Histamine H3 receptor

Chain R:


ASP TYR LYS ASP ASP ASP ASP LYS LEU GLU VAL LEU PHE GLN GLY PRO MET GLU ARG ALA PRO PRO ASP GLY PRO LEU ASN ALA SER GLY ALA LEU ALA GLY GLU ALA ALA ALA GLY GLY ALA ARG GLY PHE SER ALA A32 L37 L44 M56 L57 A58 L65 Q68
L74 I88 P89 W100 T119 F123 A152 E175 G180 G186 F192 F193 E206 G235 ALA ARG GLU ALA ALA GLY PRO GLU PRO PRO PRO GLU ALA GLN PRO SER PRO PRO PRO PRO GLY CYS TRP GLY CYS TRP GLN LYS GLY HIS GLY GLU ALA
MET PRO LEU HIS ARG TYR GLY VAL GLY GLU ALA ALA VAL GLY ALA GLU ALA GLY GLU ALA THR LEU GLY GLY GLY GLY GLY GLY GLY SER VAL ALA SER PRO THR SER SER SER GLY SER SER SER ARG GLY THR GLU ARG PRO ARG SER LEU LYS ARG GLY SER LYS PRO SER ALA SER
SER ALA SER LEU GLU LYS ARG MET LYS MET VAL SER GLN S343 Q346 L360 I363 W371 Y392 W393 F398 L401 N404 S405 A406 V407 N408 P409 V410 L411 Y412 P413 L414 S418 R421 A422 F423 T424 K425 L426 L427 CYS PRO GLN LYS LEU LYS ILE GLN PRO
HIS SER SER LEU GLU HIS CYS TRP LYS

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	330449	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 ($e^-/Å^2$)	62	Depositor
Minimum defocus (nm)	500	Depositor
Maximum defocus (nm)	2500	Depositor
Magnification	Not provided	
Image detector	GATAN K2 SUMMIT (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: A1LY4, CLR

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	0.90	0/1841	0.66	0/2471
2	B	0.83	0/2654	0.78	0/3597
3	G	0.80	0/450	0.65	0/608
4	S	0.64	0/1839	0.67	0/2493
5	R	0.81	0/2414	0.74	0/3292
All	All	0.80	0/9198	0.72	0/12461

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

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	1811	0	1799	6	0
2	B	2607	0	2510	11	0
3	G	444	0	454	1	0
4	S	1795	0	1727	24	0
5	R	2344	0	2367	40	0
6	R	12	15	0	4	0
7	R	28	46	46	2	0
All	All	9041	61	8903	80	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 4.

The worst 5 of 80 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)
2:B:254:ASP:HB2	2:B:261:LEU:HD11	1.35	1.05
5:R:119:THR:HG23	5:R:206:GLU:HG3	1.36	1.05
2:B:254:ASP:HB2	2:B:261:LEU:CD1	1.99	0.92
4:S:68:PHE:CZ	4:S:83:MET:HE2	2.07	0.89
5:R:412:TYR:HB2	5:R:413:PRO:HD3	1.54	0.88

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	A	221/354 (62%)	215 (97%)	6 (3%)	0	100	100
2	B	337/358 (94%)	324 (96%)	13 (4%)	0	100	100
3	G	56/71 (79%)	56 (100%)	0	0	100	100
4	S	230/269 (86%)	228 (99%)	2 (1%)	0	100	100
5	R	285/461 (62%)	276 (97%)	9 (3%)	0	100	100
All	All	1129/1513 (75%)	1099 (97%)	30 (3%)	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	A	200/305 (66%)	200 (100%)	0	100 100
2	B	282/298 (95%)	280 (99%)	2 (1%)	81 91
3	G	47/58 (81%)	47 (100%)	0	100 100
4	S	198/217 (91%)	198 (100%)	0	100 100
5	R	249/372 (67%)	249 (100%)	0	100 100
All	All	976/1250 (78%)	974 (100%)	2 (0%)	91 97

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	105	TYR
2	B	234	PHE

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

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)

2 ligands are 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	$\#\|Z\| > 2$
6	A1LY4	R	501	-	9,13,13	0.50	0	11,16,16	1.88	2 (18%)
7	CLR	R	502	-	31,31,31	1.21	1 (3%)	48,48,48	1.53	9 (18%)

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
6	A1LY4	R	501	-	-	2/3/12/12	1/2/2/2
7	CLR	R	502	-	-	3/10/68/68	0/4/4/4

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	R	502	CLR	C16-C17	3.20	1.61	1.54

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	R	501	A1LY4	C01-C06-C07	3.57	120.24	111.88
6	R	501	A1LY4	C02-C01-C06	-3.51	106.71	112.14
7	R	502	CLR	C12-C11-C9	3.44	119.07	113.11
7	R	502	CLR	C15-C14-C13	3.28	107.79	103.84
7	R	502	CLR	C1-C2-C3	3.07	114.40	110.47

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	R	501	A1LY4	C01-C06-C07-C08
7	R	502	CLR	C17-C20-C22-C23
7	R	502	CLR	C21-C20-C22-C23
6	R	501	A1LY4	C05-C06-C07-C08
7	R	502	CLR	C23-C24-C25-C27

All (1) ring outliers are listed below:

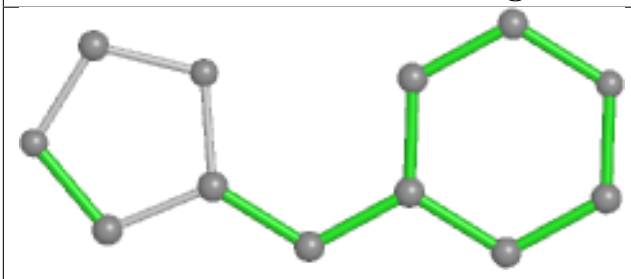
Mol	Chain	Res	Type	Atoms
6	R	501	A1LY4	C01-C02-C04-C05-C06-N03

2 monomers are involved in 6 short contacts:

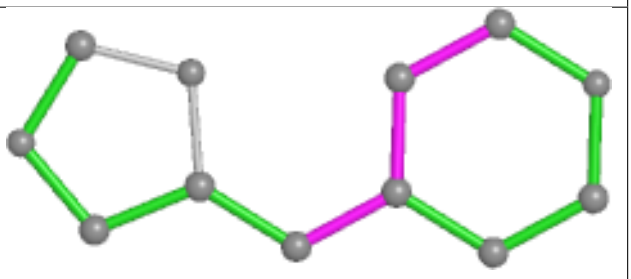
Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	R	501	A1LY4	4	0
7	R	502	CLR	2	0

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.

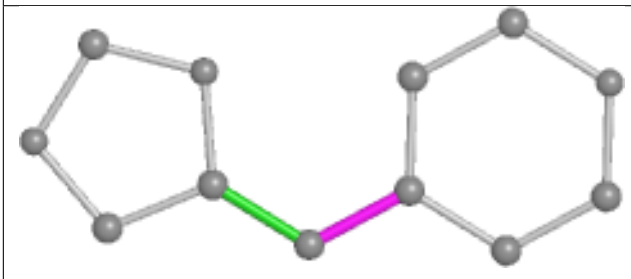
Ligand A1LY4 R 501



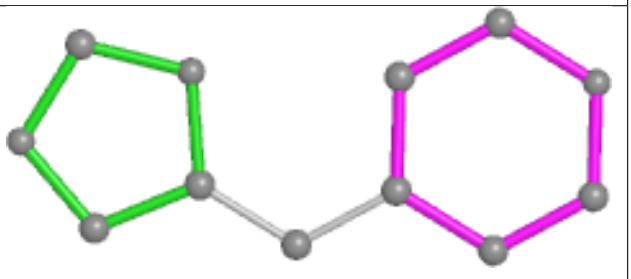
Bond lengths



Bond angles

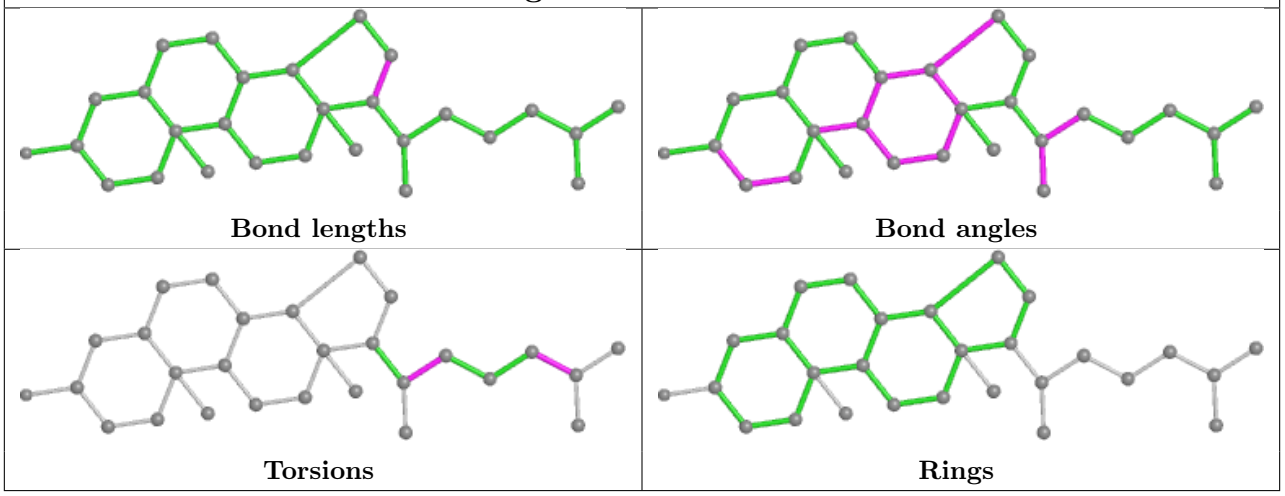


Torsions



Rings

Ligand CLR R 502



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.