Full wwPDB X-ray Structure Validation Report ⓘ

May 31, 2022 – 01:09 pm BST

PDB ID	:	7Q6J
Title	:	Crystal structure of the human GDAP1 CMT2 mutant-H123R
Authors	:	Sutinen, A.; Kursula, P.
Deposited on	:	2021-11-08
Resolution	:	2.20 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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/XrayValidationReportHelp
with specific help available everywhere you see the ⓘ symbol.

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

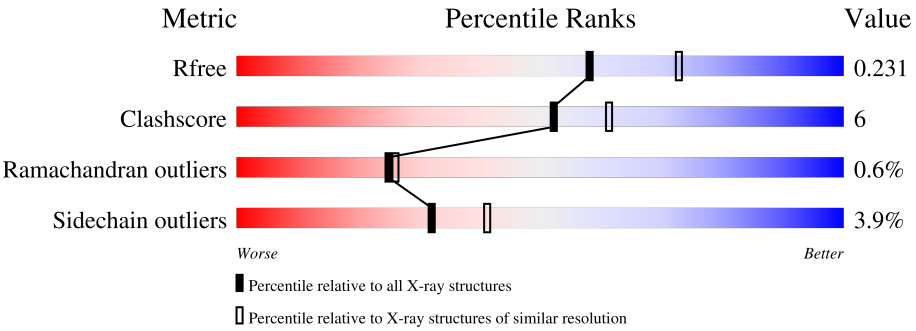
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.28.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.28.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.20 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)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower 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	280	78% 12% • 8%
1	B	280	73% 14% • 11%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 8586 atoms, of which 4247 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, 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.

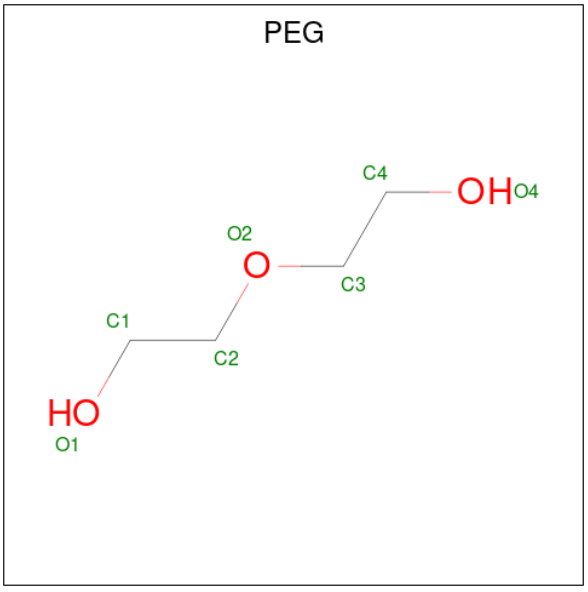
- Molecule 1 is a protein called Ganglioside-induced differentiation-associated protein 1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	257	Total	C	H	N	O	S	0	4	0
			4332	1379	2168	380	396	9			
1	B	248	Total	C	H	N	O	S	0	2	0
			4127	1312	2059	363	384	9			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	123	ARG	HIS	engineered mutation	UNP Q8TB36
B	123	ARG	HIS	engineered mutation	UNP Q8TB36

- Molecule 2 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	H	O	0	0
			17	4	10	3		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C H O 17 4 10 3	0	0

- Molecule 3 is water.

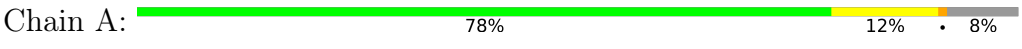
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	72	Total O 72 72	0	0
3	B	21	Total O 21 21	0	0

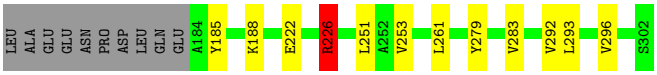
3 Residue-property plots (i)

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: Ganglioside-induced differentiation-associated protein 1

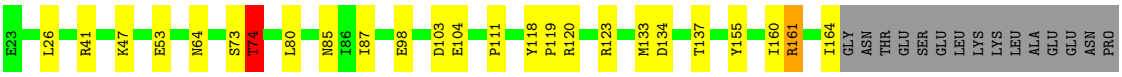
Chain A:





• Molecule 1: Ganglioside-induced differentiation-associated protein 1

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	72.71Å 115.88Å 116.18Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	45.38 – 2.20 45.38 – 2.20	Depositor EDS
% Data completeness (in resolution range)	99.7 (45.38-2.20) 99.7 (45.38-2.20)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.32 (at 2.20Å)	Xtriage
Refinement program	PHENIX 1.18-3861	Depositor
R, R_{free}	0.215 , 0.236 0.210 , 0.231	Depositor DCC
R_{free} test set	1990 reflections (3.95%)	wwPDB-VP
Wilson B-factor (Å^2)	49.4	Xtriage
Anisotropy	0.579	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.004 for -h,l,k	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	8586	wwPDB-VP
Average B, all atoms (Å^2)	80.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.32% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

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

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	0.69	0/2213	0.78	2/2993 (0.1%)
1	B	0.61	2/2113 (0.1%)	0.77	6/2860 (0.2%)
All	All	0.65	2/4326 (0.0%)	0.78	8/5853 (0.1%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	232	GLU	CD-OE2	-5.23	1.19	1.25
1	B	232	GLU	CG-CD	-5.19	1.44	1.51

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	232	GLU	OE1-CD-OE2	-9.03	112.46	123.30
1	B	232	GLU	CG-CD-OE1	8.07	134.45	118.30
1	A	226	ARG	CG-CD-NE	-6.44	98.27	111.80
1	B	227	ASN	O-C-N	-6.20	112.78	122.70
1	B	228	GLU	N-CA-CB	-5.67	100.39	110.60
1	B	232	GLU	CB-CG-CD	-5.39	99.65	114.20
1	B	227	ASN	C-N-CA	-5.27	108.53	121.70
1	A	120	ARG	NE-CZ-NH2	5.02	122.81	120.30

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	2164	2168	2164	21	0
1	B	2068	2059	2057	26	0
2	A	14	20	20	0	0
3	A	72	0	0	0	0
3	B	21	0	0	0	0
All	All	4339	4247	4241	47	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 6.

All (47) 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:26:LEU:HD22	1:A:80:LEU:HD11	1.58	0.83
1:A:222:GLU:OE2	1:A:226:ARG:NH2	2.17	0.78
1:B:197:LEU:HD11	1:B:204:TYR:CG	2.24	0.73
1:B:133:MET:O	1:B:137:THR:HG23	1.90	0.71
1:B:226:ARG:HB3	1:B:237:PRO:HG2	1.76	0.67
1:A:133:MET:O	1:A:137:THR:HG23	1.94	0.66
1:B:227:ASN:HA	1:B:232:GLU:OE2	1.94	0.66
1:B:227:ASN:O	1:B:228:GLU:C	2.33	0.64
1:B:222:GLU:OE2	1:B:226:ARG:NH2	2.30	0.64
1:B:160:ILE:O	1:B:160:ILE:HG22	1.97	0.63
1:B:197:LEU:HD11	1:B:204:TYR:CD2	2.34	0.61
1:B:227:ASN:HA	1:B:232:GLU:CD	2.22	0.60
1:A:138[B]:HIS:CE1	1:A:142:LEU:HD11	2.37	0.59
1:A:24:VAL:HG13	1:A:24:VAL:O	2.02	0.59
1:A:31[A]:TRP:CZ3	1:A:76:GLU:HA	2.39	0.58
1:B:103:ASP:OD1	1:B:104:GLU:N	2.38	0.57
1:A:31[A]:TRP:CH2	1:A:76:GLU:HA	2.41	0.56
1:B:120:ARG:O	1:B:123:ARG:HG2	2.09	0.53
1:B:227:ASN:OD1	1:B:232:GLU:HG3	2.09	0.53
1:B:280:TYR:O	1:B:283:VAL:HG22	2.11	0.51
1:A:137:THR:HG22	1:A:261:LEU:HD11	1.93	0.51
1:A:34:SER:OG	1:A:37:SER:HB3	2.11	0.50
1:A:35:PHE:CE1	1:A:253:VAL:HG12	2.47	0.50
1:A:75:GLY:C	1:A:76:GLU:HG3	2.31	0.49
1:A:118:TYR:HB3	1:A:119:PRO:HD3	1.96	0.47
1:A:251:LEU:HD23	1:A:279:TYR:CZ	2.51	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:75:GLY:O	1:A:76:GLU:HG3	2.16	0.45
1:A:292:VAL:HG12	1:A:293:LEU:CD2	2.47	0.45
1:A:31[A]:TRP:CZ3	1:A:77:VAL:N	2.84	0.45
1:B:64[B]:ASN:CG	1:B:64[B]:ASN:O	2.55	0.45
1:B:251:LEU:HD23	1:B:279:TYR:CZ	2.53	0.44
1:B:85:ASN:HB3	1:B:87:ILE:HD11	2.00	0.43
1:A:47:LYS:NZ	1:A:98:GLU:OE1	2.44	0.43
1:B:73:SER:O	1:B:74:THR:OG1	2.35	0.43
1:B:137:THR:HG22	1:B:261:LEU:HD11	2.00	0.42
1:B:26:LEU:HD22	1:B:80:LEU:HD11	2.00	0.42
1:B:47:LYS:NZ	1:B:98:GLU:OE1	2.41	0.42
1:B:41[A]:ARG:HD2	1:B:53:GLU:OE1	2.20	0.42
1:B:249:VAL:CG2	1:B:250:SER:N	2.83	0.42
1:B:155:TYR:CD1	1:B:155:TYR:C	2.94	0.41
1:A:36[A]:SER:HB2	1:A:90:ALA:HB2	2.01	0.41
1:B:111:PRO:HG2	1:B:118:TYR:HA	2.02	0.41
1:B:118:TYR:HB3	1:B:119:PRO:HD3	2.02	0.41
1:A:65:GLU:OE2	1:A:159:ARG:CD	2.68	0.41
1:A:30:HIS:CE1	1:A:41:ARG:HD3	2.57	0.40
1:B:73:SER:C	1:B:74:THR:OG1	2.59	0.40
1:A:70:ARG:O	1:A:73:SER:HB2	2.22	0.40

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 X-ray entries followed by that with respect to entries of similar resolution.

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	257/280 (92%)	248 (96%)	8 (3%)	1 (0%)	34	37
1	B	246/280 (88%)	235 (96%)	9 (4%)	2 (1%)	19	19
All	All	503/560 (90%)	483 (96%)	17 (3%)	3 (1%)	25	26

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	24	VAL
1	B	74	THR
1	B	161	ARG

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 X-ray entries followed by that with respect to entries of similar resolution.

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	241/258 (93%)	232 (96%)	9 (4%)	34	43
1	B	232/258 (90%)	223 (96%)	9 (4%)	32	41
All	All	473/516 (92%)	455 (96%)	18 (4%)	32	42

All (18) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	72	ASN
1	A	73	SER
1	A	76	GLU
1	A	159	ARG
1	A	185	TYR
1	A	188	LYS
1	A	226	ARG
1	A	283	VAL
1	A	296	VAL
1	B	74	THR
1	B	134	ASP
1	B	161	ARG
1	B	164	ILE
1	B	197	LEU
1	B	198	ASP
1	B	233	GLU
1	B	235	GLN
1	B	296	VAL

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	92	GLN
1	B	235	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 monosaccharides in this entry.

5.6 Ligand geometry

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	$\#\|Z\| > 2$
2	PEG	A	402	-	6,6,6	0.54	0	5,5,5	0.70	0
2	PEG	A	401	-	6,6,6	0.67	0	5,5,5	0.58	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
2	PEG	A	402	-	-	2/4/4/4	-
2	PEG	A	401	-	-	1/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) torsion outliers are listed below:

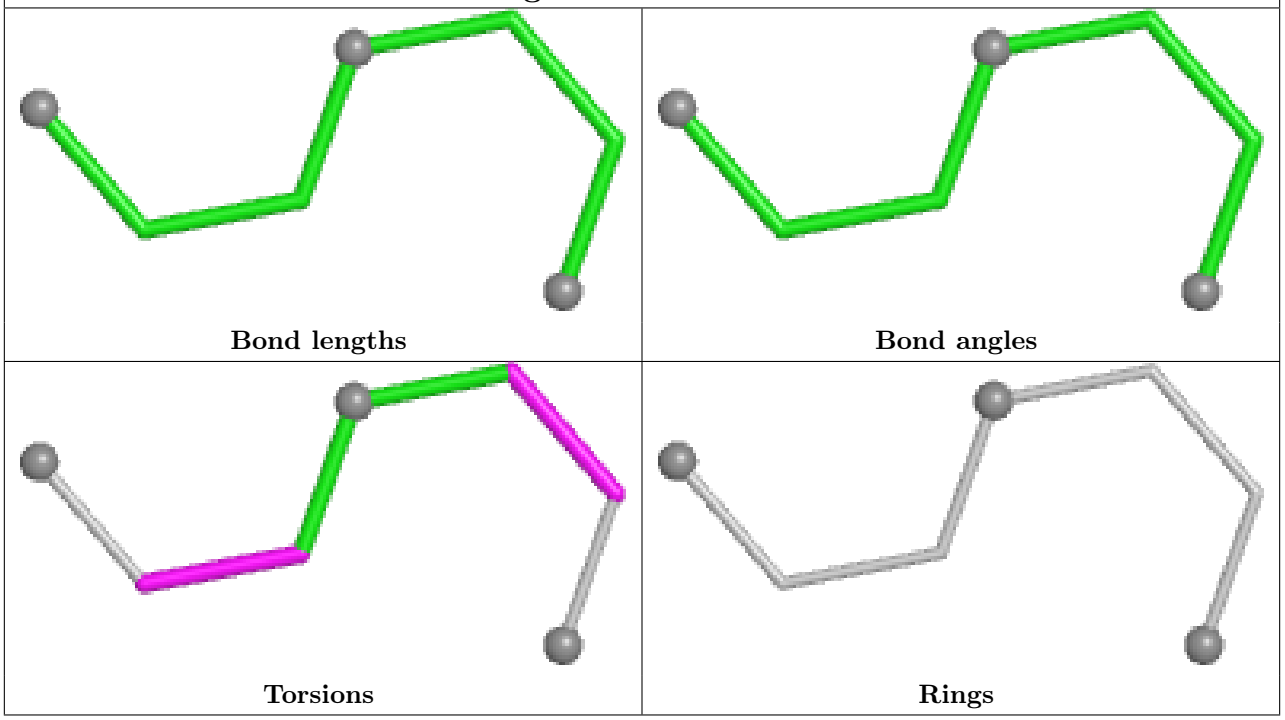
Mol	Chain	Res	Type	Atoms
2	A	402	PEG	O1-C1-C2-O2
2	A	401	PEG	O2-C3-C4-O4
2	A	402	PEG	O2-C3-C4-O4

There are no ring outliers.

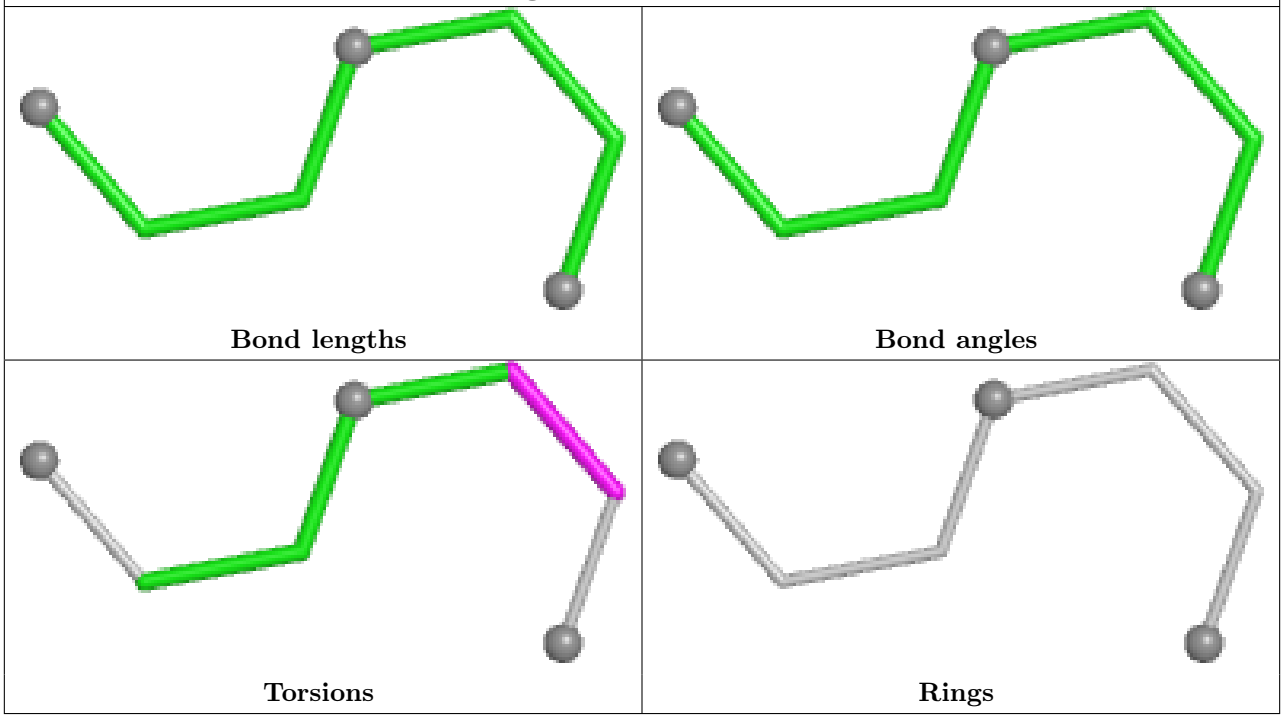
No monomer is involved in short contacts.

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 PEG A 402



Ligand PEG A 401



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.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

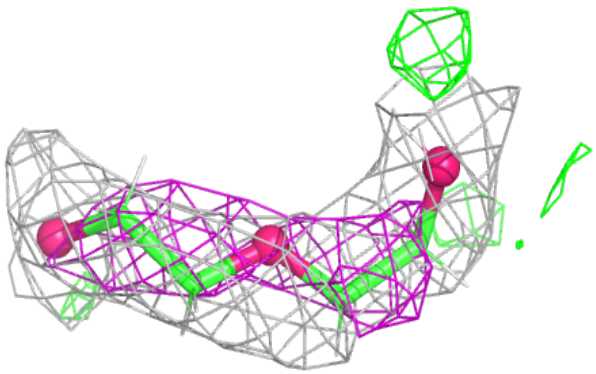
6.4 Ligands

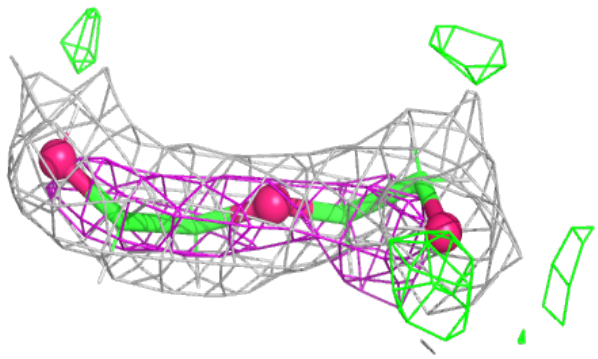
Unable to reproduce the depositors R factor - this section is therefore empty.

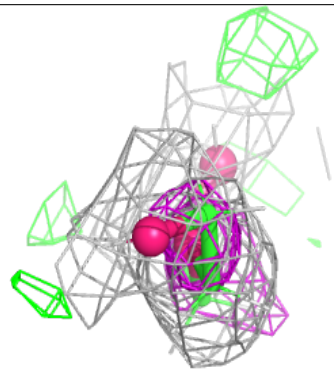
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around PEG A 401:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

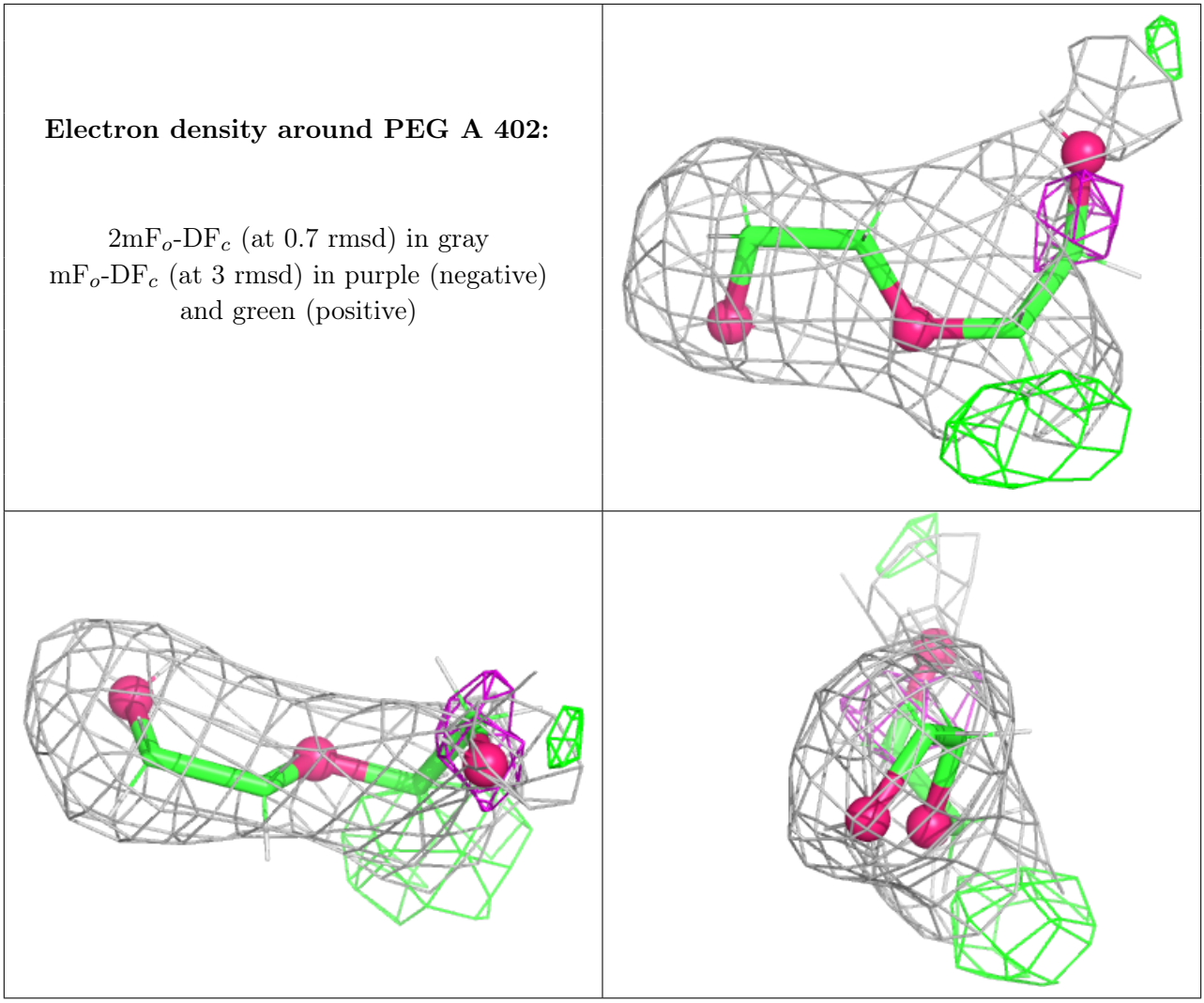






Electron density around PEG A 402:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)



6.5 Other polymers (i)

Unable to reproduce the depositors R factor - this section is therefore empty.