



wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 28, 2022 – 03:37 pm GMT

PDB ID : 6QH6
Title : AP2 clathrin adaptor core with two cargo peptides in open+ conformation
Authors : Wrobel, A.G.; Owen, D.J.; McCoy, A.J.; Evans, P.R.
Deposited on : 2019-01-15
Resolution : 5.00 Å(reported)

This is a wwPDB X-ray Structure 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/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.8.4, CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.31.3
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) : 2.31.3

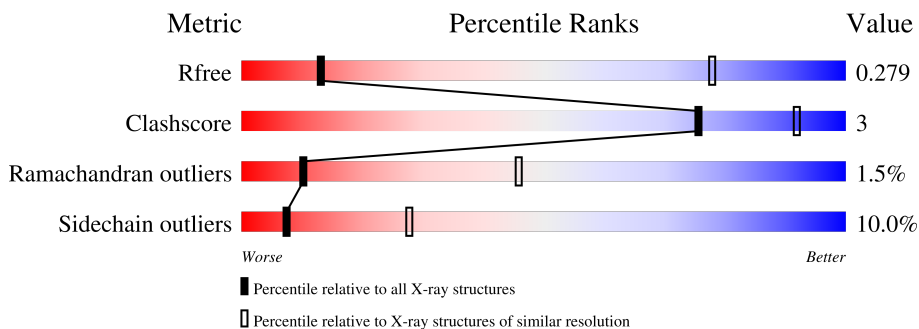
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 5.00 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)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1140 (6.20-3.80)
Clashscore	141614	1000 (6.16-3.82)
Ramachandran outliers	138981	1146 (6.20-3.80)
Sidechain outliers	138945	1122 (6.20-3.80)

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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	621	83% 15% ..
2	B	592	68% 15% . 15%
3	M	446	49% 7% . 43%
4	N	446	26% 6% 68%
5	P	6	100%
6	Q	10	60% 20% 10% 10%
7	S	142	87% 12% .

2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 13342 atoms, of which 0 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 AP-2 complex subunit alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	614	4844	3085	834	904	21	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	8	ASP	GLU	conflict	UNP Q66HM2

- Molecule 2 is a protein called AP-2 complex subunit beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	502	3974	2529	661	763	21	0	0	0

- Molecule 3 is a protein called AP-2 complex subunit mu.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	M	255	2033	1310	350	359	14	0	0	0

There are 11 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
M	237	MET	-	insertion	UNP Q3ZC13
M	238	GLU	-	insertion	UNP Q3ZC13
M	239	GLN	-	insertion	UNP Q3ZC13
M	240	LYS	-	insertion	UNP Q3ZC13
M	241	LEU	-	insertion	UNP Q3ZC13
M	242	ILE	-	insertion	UNP Q3ZC13
M	243	SER	-	insertion	UNP Q3ZC13
M	244	GLU	-	insertion	UNP Q3ZC13
M	245	GLU	-	insertion	UNP Q3ZC13

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
M	246	ASP	-	insertion	UNP Q3ZC13
M	247	LEU	-	insertion	UNP Q3ZC13

- Molecule 4 is a protein called AP-2 complex subunit mu.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	N	143	1157	742	203	207	5	0	0	0

There are 11 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
N	237	MET	-	insertion	UNP P84092
N	238	GLU	-	insertion	UNP P84092
N	239	GLN	-	insertion	UNP P84092
N	240	LYS	-	insertion	UNP P84092
N	241	LEU	-	insertion	UNP P84092
N	242	ILE	-	insertion	UNP P84092
N	243	SER	-	insertion	UNP P84092
N	244	GLU	-	insertion	UNP P84092
N	245	GLU	-	insertion	UNP P84092
N	246	ASP	-	insertion	UNP P84092
N	247	LEU	-	insertion	UNP P84092

- Molecule 5 is a protein called TGN38 CARGO PEPTIDE.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
5	P	6	57	34	11	12	0	0	0

- Molecule 6 is a protein called CD4 CARGO PEPTIDE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	N	O	P				S
6	Q	9	77	46	14	15	1	1	10	0	0

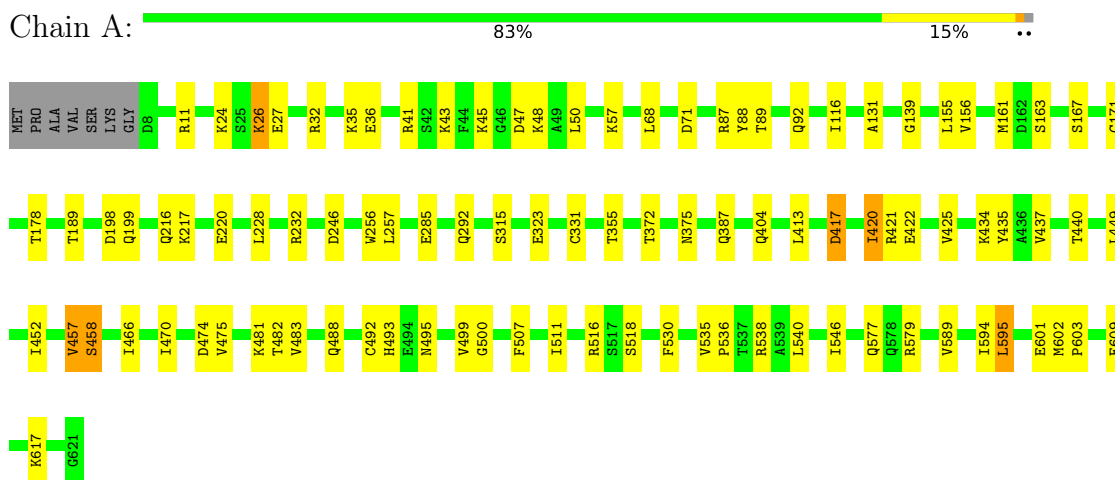
- Molecule 7 is a protein called AP-2 complex subunit sigma.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
7	S	142	1200	778	200	215	7	0	0	0

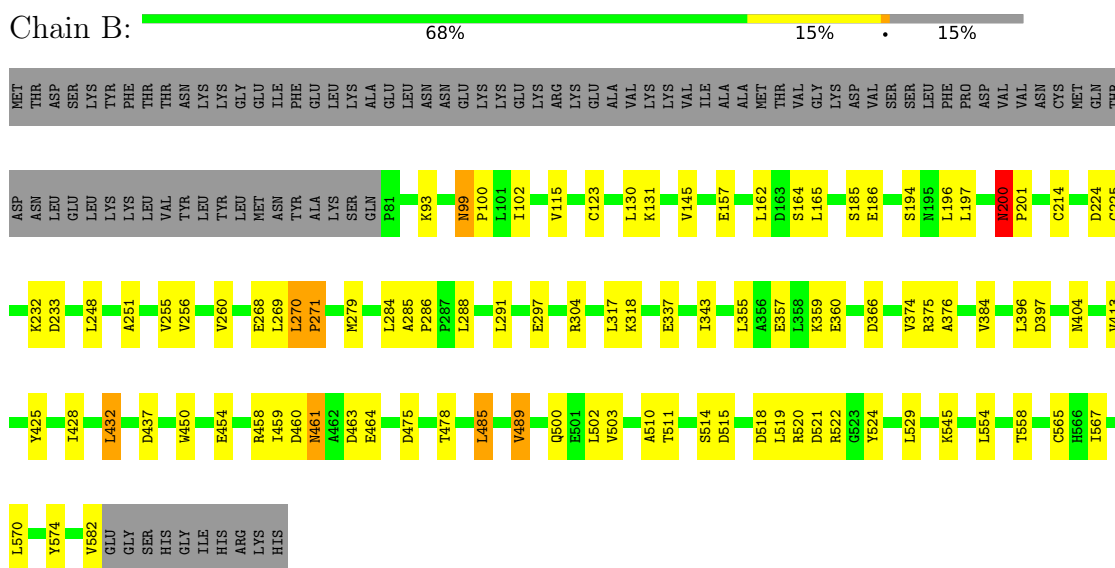
3 Residue-property plots

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: AP-2 complex subunit alpha

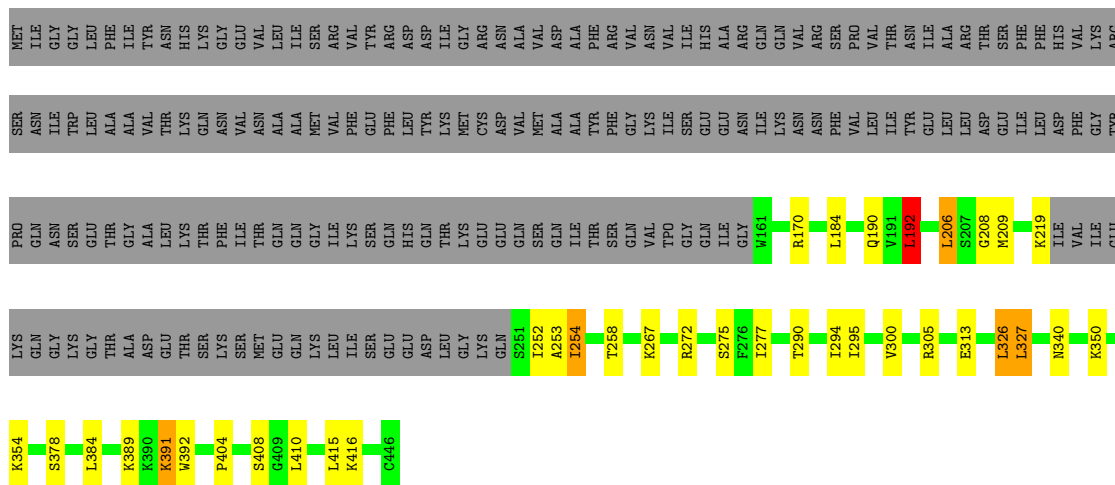


- Molecule 2: AP-2 complex subunit beta

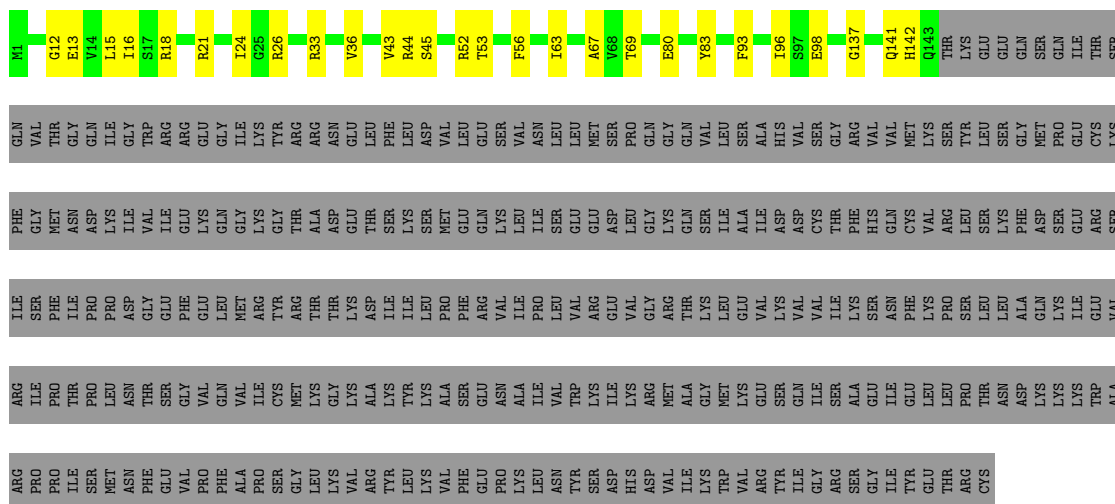


- Molecule 3: AP-2 complex subunit mu





- Molecule 4: AP-2 complex subunit mu



- Molecule 5: TGN38 CARGO PEPTIDE

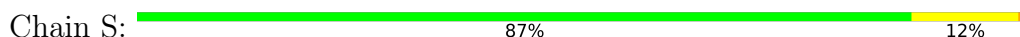


There are no outlier residues recorded for this chain.

- Molecule 6: CD4 CARGO PEPTIDE



- Molecule 7: AP-2 complex subunit sigma





4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α , β , γ	180.50Å 180.50Å 210.06Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	91.00 – 5.00 90.78 – 5.00	Depositor EDS
% Data completeness (in resolution range)	99.4 (91.00-5.00) 99.4 (90.78-5.00)	Depositor EDS
R_{merge}	0.20	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.10 (at 5.12Å)	Xtriage
Refinement program	REFMAC 5.8.0189	Depositor
R, R_{free}	0.259 , 0.285 0.244 , 0.279	Depositor DCC
R_{free} test set	742 reflections (4.79%)	wwPDB-VP
Wilson B-factor (Å ²)	256.6	Xtriage
Anisotropy	0.015	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	(Not available) , (Not available)	EDS
L-test for twinning ²	$\langle L \rangle = 0.34$, $\langle L^2 \rangle = 0.17$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.57	EDS
Total number of atoms	13342	wwPDB-VP
Average B, all atoms (Å ²)	325.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.86% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

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

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	0.51	0/4929	0.75	1/6680 (0.0%)
2	B	0.52	0/4038	0.76	1/5488 (0.0%)
3	M	0.56	0/2076	0.76	2/2797 (0.1%)
4	N	0.57	0/1180	0.70	0/1592
5	P	0.65	0/57	0.70	0/74
6	Q	0.50	0/65	0.86	0/82
7	S	0.52	0/1224	0.73	0/1650
All	All	0.53	0/13569	0.75	4/18363 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
2	B	0	1
3	M	0	1
6	Q	0	1
7	S	0	1
All	All	0	4

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	M	192	LEU	CA-CB-CG	6.13	129.39	115.30
3	M	206	LEU	CA-CB-CG	5.63	128.25	115.30
1	A	595	LEU	CA-CB-CG	5.51	127.98	115.30
2	B	162	LEU	CA-CB-CG	5.28	127.44	115.30

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	271	PRO	Peptide
3	M	326	LEU	Peptide
6	Q	2	MET	Mainchain
7	S	138	LEU	Peptide

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	A	4844	0	4956	33	0
2	B	3974	0	4061	35	0
3	M	2033	0	2097	6	0
4	N	1157	0	1158	10	0
5	P	57	0	53	0	0
6	Q	77	0	83	0	0
7	S	1200	0	1195	10	0
All	All	13342	0	13603	83	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 3.

The worst 5 of 83 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:387:GLN:HE22	1:A:420:ILE:HG21	1.53	0.73
2:B:297:GLU:HB2	4:N:83:TYR:OH	1.88	0.72
2:B:200:ASN:HB3	2:B:201:PRO:HD2	1.75	0.69
2:B:343:ILE:HD11	2:B:376:ALA:HB1	1.75	0.69
2:B:200:ASN:HB3	2:B:201:PRO:CD	2.24	0.66

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	612/621 (99%)	567 (93%)	36 (6%)	9 (2%)	10	46
2	B	500/592 (84%)	448 (90%)	46 (9%)	6 (1%)	13	50
3	M	251/446 (56%)	214 (85%)	30 (12%)	7 (3%)	5	32
4	N	141/446 (32%)	129 (92%)	11 (8%)	1 (1%)	22	62
5	P	4/6 (67%)	4 (100%)	0	0	100	100
6	Q	6/10 (60%)	4 (67%)	1 (17%)	1 (17%)	0	3
7	S	140/142 (99%)	130 (93%)	9 (6%)	1 (1%)	22	62
All	All	1654/2263 (73%)	1496 (90%)	133 (8%)	25 (2%)	10	46

5 of 25 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	458	SER
1	A	601	GLU
2	B	200	ASN
3	M	252	ILE
3	M	254	ILE

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	538/543 (99%)	488 (91%)	50 (9%)	9	30
2	B	452/533 (85%)	403 (89%)	49 (11%)	6	25

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	M	227/397 (57%)	201 (88%)	26 (12%)	5	23
4	N	124/398 (31%)	111 (90%)	13 (10%)	7	26
5	P	6/6 (100%)	6 (100%)	0	100	100
6	Q	8/9 (89%)	6 (75%)	2 (25%)	0	4
7	S	131/131 (100%)	123 (94%)	8 (6%)	18	46
All	All	1486/2017 (74%)	1338 (90%)	148 (10%)	7	28

5 of 148 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	M	340	ASN
7	S	54	ASN
3	M	384	LEU
4	N	44	ARG
1	A	609	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	461	ASN
7	S	48	ASN
3	M	329	GLN
7	S	86	ASN
4	N	122	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](#)

1 non-standard protein/DNA/RNA residue 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	$\# Z > 2$
6	SEP	Q	3	6	8,9,10	0.72	0	8,12,14	1.20	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
6	SEP	Q	3	6	-	2/5/8/10	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	Q	3	SEP	CA-CB-OG-P
6	Q	3	SEP	CB-OG-P-O3P

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues

The following chains have linkage breaks:

Mol	Chain	Number of breaks
6	Q	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	Q	3:SEP	C	4:GLN	N	1.92

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.

6.5 Other polymers

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