



# Full wwPDB X-ray Structure Validation Report ⓘ

Mar 8, 2018 – 06:25 pm GMT

PDB ID : 4JQI  
Title : Structure of active beta-arrestin1 bound to a G protein-coupled receptor phosphopeptide  
Authors : Shukla, A.K.; Manglik, A.; Kruse, A.C.; Xiao, K.; Reis, R.I.; Tseng, W.C.; Staus, D.P.; Hilger, D.; Uysal, S.; Huang, L.H.; Paduch, M.; Shukla, P.T.; Koide, A.; Koide, S.; Weis, W.I.; Kossiakoff, A.A.; Kobilka, B.K.; Lefkowitz, R.J.  
Deposited on : 2013-03-20  
Resolution : 2.60 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](mailto: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 : 1.7.3 (157068), CSD as539be (2018)  
Xtriage (Phenix) : 1.13  
EDS : trunk30967  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Refmac : 5.8.0158  
CCP4 : 7.0 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk30967

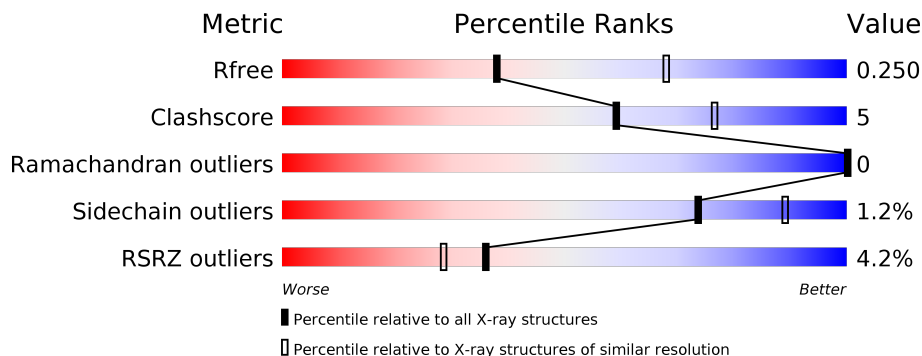
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.60 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))
$R_{free}$	111664	2767 (2.60-2.60)
Clashscore	122126	3110 (2.60-2.60)
Ramachandran outliers	120053	3062 (2.60-2.60)
Sidechain outliers	120020	3062 (2.60-2.60)
RSRZ outliers	108989	2706 (2.60-2.60)

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	401	
2	H	237	
3	L	215	
4	V	29	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
6	EDO	H	303	-	-	-	X

## 2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 5858 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 Beta-arrestin-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	353	2698	1729	459	500	10	0	0	0

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-7	GLY	-	EXPRESSION TAG	UNP P29066
A	-6	SER	-	EXPRESSION TAG	UNP P29066
A	-5	PRO	-	EXPRESSION TAG	UNP P29066
A	-4	GLU	-	EXPRESSION TAG	UNP P29066
A	-3	PHE	-	EXPRESSION TAG	UNP P29066
A	-2	PRO	-	EXPRESSION TAG	UNP P29066
A	-1	GLY	-	EXPRESSION TAG	UNP P29066
A	0	ARG	-	EXPRESSION TAG	UNP P29066
A	1	LEU	-	EXPRESSION TAG	UNP P29066

- Molecule 2 is a protein called Fab30 heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	H	204	1496	950	249	292	5	0	0	0

- Molecule 3 is a protein called Fab30 light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	L	185	1359	856	225	274	4	0	0	0

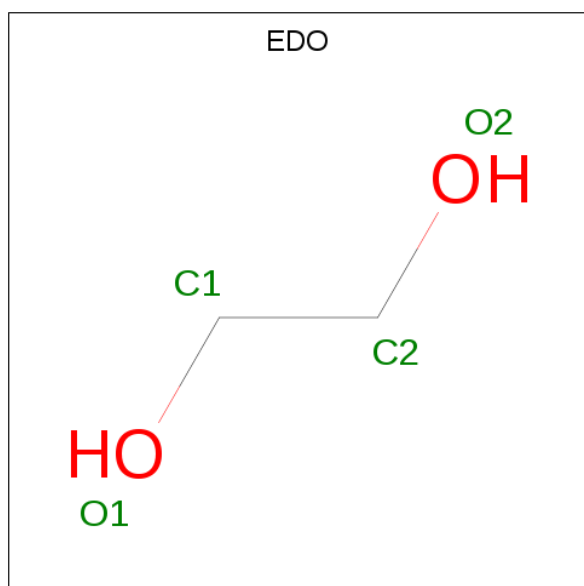
- Molecule 4 is a protein called Vasopressin V2 receptor phosphopeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	N	O	P				S
4	V	21	163	78	21	55	8	1	0	0	0

- Molecule 5 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

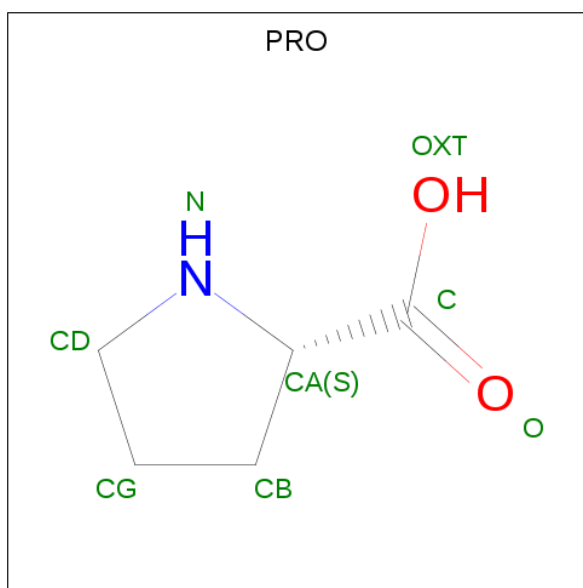
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	H	1	Total	Cl	0	0
			1	1		
5	A	1	Total	Cl	0	0
			1	1		

- Molecule 6 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	1	Total	C O	0	0
			4	2 2		
6	H	1	Total	C O	0	0
			4	2 2		
6	H	1	Total	C O	0	0
			4	2 2		
6	H	1	Total	C O	0	0
			4	2 2		
6	L	1	Total	C O	0	0
			4	2 2		
6	L	1	Total	C O	0	0
			4	2 2		

- Molecule 7 is PROLINE (three-letter code: PRO) (formula: C<sub>5</sub>H<sub>9</sub>NO<sub>2</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
7	L	1	8	5	1	2	0	0

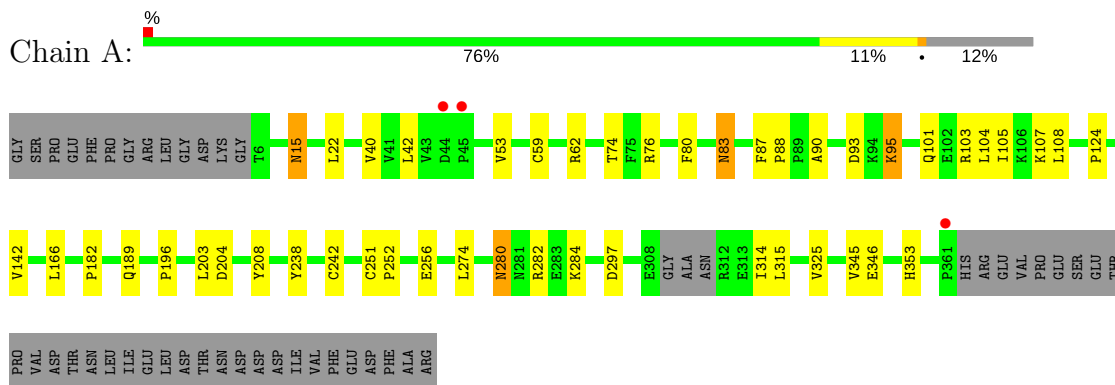
- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	69	Total	O	0	0
			69	69		
8	H	27	Total	O	0	0
			27	27		
8	L	8	Total	O	0	0
			8	8		
8	V	4	Total	O	0	0
			4	4		

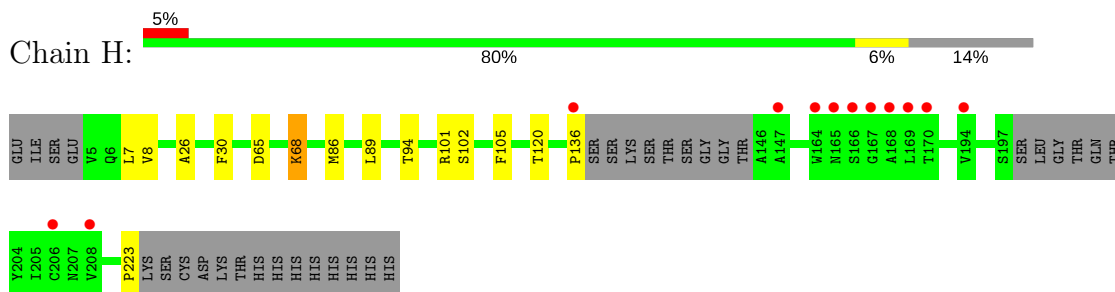
### 3 Residue-property plots i

These plots are drawn for all protein, RNA and DNA 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 and electron density. 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. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). 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.

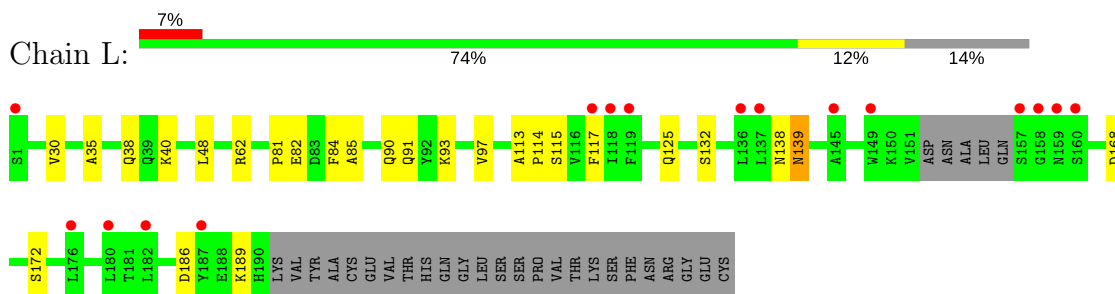
- Molecule 1: Beta-arrestin-1



- Molecule 2: Fab30 heavy chain



- Molecule 3: Fab30 light chain



- Molecule 4: Vasopressin V2 receptor phosphopeptide



ALA	P353	S363	K367	THR
ARG	GLN	S364	D368	SER
GLY	ASP	L365		SER
R346	E356	A366		



## 4 Data and refinement statistics

Property	Value	Source
Space group	I 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	116.84Å 125.13Å 144.20Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.28 – 2.60 39.28 – 2.58	Depositor EDS
% Data completeness (in resolution range)	98.0 (39.28-2.60) 91.6 (39.28-2.58)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.09	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.90 (at 2.58Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: 1.8.1_1168)	Depositor
R, $R_{free}$	0.201 , 0.247 0.201 , 0.250	Depositor DCC
$R_{free}$ test set	1644 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	53.2	Xtrriage
Anisotropy	0.488	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 59.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	5858	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	86.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.01% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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: TPO, CL, EDO, 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	0.25	0/2757	0.47	0/3756
2	H	0.23	0/1537	0.40	0/2105
3	L	0.23	0/1388	0.46	1/1893 (0.1%)
4	V	0.16	0/76	0.40	0/95
All	All	0.24	0/5758	0.45	1/7849 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	L	139	ASN	N-CA-C	5.42	125.65	111.00

There are no chirality outliers.

There are no planarity outliers.

### 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	2698	0	2689	36	0
2	H	1496	0	1364	8	0
3	L	1359	0	1268	13	0
4	V	163	0	103	3	0
5	A	1	0	0	0	0
5	H	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	A	4	0	6	2	0
6	H	12	0	18	0	0
6	L	8	0	12	0	0
7	L	8	0	7	0	0
8	A	69	0	0	3	0
8	H	27	0	0	0	0
8	L	8	0	0	0	0
8	V	4	0	0	0	0
All	All	5858	0	5467	57	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 5.

All (57) 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:30:VAL:HG11	3:L:91:GLN:HB2	1.56	0.85
1:A:93:ASP:HB3	1:A:95:LYS:HD3	1.70	0.74
1:A:182:PRO:HG2	1:A:203:LEU:HB2	1.72	0.70
1:A:107:LYS:NZ	4:V:364:SEP:O	2.27	0.66
2:H:65:ASP:OD1	2:H:68:LYS:HE2	1.97	0.64
1:A:256:GLU:HB2	8:A:551:HOH:O	1.99	0.62
3:L:38:GLN:HB2	3:L:48:LEU:HD11	1.82	0.61
1:A:353:HIS:ND1	6:A:402:EDO:O2	2.31	0.58
2:H:86:MET:HB3	2:H:89:LEU:HD21	1.86	0.58
1:A:252:PRO:O	1:A:284:LYS:NZ	2.37	0.58
1:A:22:LEU:HD12	1:A:166:LEU:HD23	1.87	0.56
1:A:107:LYS:NZ	4:V:363:SEP:OG	2.39	0.56
1:A:83:ASN:H	1:A:83:ASN:HD22	1.55	0.54
3:L:138:ASN:OD1	3:L:139:ASN:ND2	2.41	0.54
1:A:15:ASN:H	1:A:15:ASN:HD22	1.57	0.53
3:L:62:ARG:NH2	3:L:82:GLU:OE2	2.41	0.52
2:H:8:VAL:HG13	2:H:26:ALA:HB3	1.90	0.52
1:A:62:ARG:NH1	1:A:74:THR:OG1	2.43	0.52
1:A:280:ASN:H	1:A:280:ASN:HD22	1.58	0.52
1:A:238:TYR:OH	1:A:346:GLU:OE1	2.28	0.51
1:A:90:ALA:HB3	1:A:93:ASP:HB2	1.93	0.51
1:A:101:GLN:O	1:A:105:ILE:HG12	2.11	0.50
3:L:168:ASP:O	3:L:172:SER:N	2.39	0.50
1:A:124:PRO:HB2	1:A:142:VAL:HG11	1.93	0.50
1:A:274:LEU:CD2	8:A:551:HOH:O	2.60	0.49

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:103:ARG:NH2	4:V:366:ALA:O	2.41	0.48
3:L:91:GLN:O	3:L:97:VAL:HG23	2.15	0.47
1:A:15:ASN:N	1:A:15:ASN:HD22	2.11	0.47
3:L:81:PRO:HA	3:L:84:PHE:HE2	1.79	0.46
3:L:35:ALA:HB3	3:L:90:GLN:HE21	1.80	0.46
1:A:53:VAL:HB	1:A:87:PHE:HB3	1.98	0.45
1:A:204:ASP:OD1	1:A:208:TYR:OH	2.19	0.45
1:A:59:CYS:HB3	1:A:80:PHE:HB3	1.99	0.45
3:L:113:ALA:HA	3:L:114:PRO:HD3	1.88	0.44
1:A:124:PRO:HG3	1:A:315:LEU:O	2.18	0.43
1:A:314:ILE:HG22	1:A:315:LEU:O	2.18	0.43
1:A:42:LEU:HB2	1:A:108:LEU:HD23	2.01	0.43
1:A:282:ARG:NH2	1:A:297:ASP:OD1	2.50	0.43
1:A:189:GLN:HG2	1:A:196:PRO:HB3	2.01	0.43
3:L:40:LYS:HG2	3:L:85:ALA:HB2	2.00	0.42
1:A:325:VAL:HB	1:A:345:VAL:HG12	2.01	0.42
3:L:115:SER:OG	3:L:138:ASN:HB3	2.19	0.42
1:A:107:LYS:HG2	1:A:108:LEU:HD12	2.02	0.42
2:H:101:ARG:HG2	2:H:102:SER:N	2.35	0.42
1:A:242:CYS:O	1:A:315:LEU:HA	2.20	0.41
2:H:7:LEU:HD21	2:H:30:PHE:HZ	1.85	0.41
2:H:94:THR:HG23	2:H:120:THR:HA	2.02	0.41
1:A:62:ARG:NH2	1:A:76:ARG:HD2	2.36	0.41
1:A:274:LEU:HD22	8:A:547:HOH:O	2.21	0.41
3:L:125:GLN:HE22	3:L:132:SER:HB2	1.85	0.41
3:L:186:ASP:HA	3:L:189:LYS:HD3	2.02	0.41
2:H:136:PRO:HD2	2:H:223:PRO:HG3	2.02	0.41
1:A:280:ASN:N	1:A:280:ASN:HD22	2.15	0.41
1:A:251:CYS:SG	1:A:284:LYS:HD3	2.62	0.40
1:A:40:VAL:HG22	1:A:104:LEU:HD13	2.03	0.40
1:A:87:PHE:HA	1:A:88:PRO:HA	1.94	0.40
6:A:402:EDO:H22	2:H:105:PHE:CD2	2.56	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	349/401 (87%)	333 (95%)	16 (5%)	0	100	100
2	H	198/237 (84%)	189 (96%)	9 (4%)	0	100	100
3	L	181/215 (84%)	176 (97%)	5 (3%)	0	100	100
4	V	9/29 (31%)	7 (78%)	2 (22%)	0	100	100
All	All	737/882 (84%)	705 (96%)	32 (4%)	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 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	292/356 (82%)	288 (99%)	4 (1%)	69	87
2	H	153/200 (76%)	152 (99%)	1 (1%)	85	95
3	L	144/190 (76%)	142 (99%)	2 (1%)	69	87
4	V	7/16 (44%)	7 (100%)	0	100	100
All	All	596/762 (78%)	589 (99%)	7 (1%)	74	89

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	15	ASN
1	A	83	ASN
1	A	95	LYS
1	A	280	ASN
2	H	68	LYS
3	L	93	LYS
3	L	117	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (10) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	15	ASN
1	A	83	ASN
1	A	162	ASN
1	A	222	ASN
1	A	280	ASN
2	H	31	ASN
3	L	28	GLN
3	L	90	GLN
3	L	125	GLN
3	L	161	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](#)

8 non-standard protein/DNA/RNA residues 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
4	TPO	V	347	4	9,10,11	1.22	1 (11%)	11,14,16	1.84	2 (18%)
4	SEP	V	350	4	9,9,10	1.58	2 (22%)	9,12,14	1.70	2 (22%)
4	SEP	V	357	4	9,9,10	1.52	1 (11%)	9,12,14	1.74	2 (22%)
4	TPO	V	359	4	9,10,11	1.24	0	11,14,16	1.17	1 (9%)
4	TPO	V	360	4	9,10,11	1.20	0	11,14,16	1.68	1 (9%)
4	SEP	V	362	4	9,9,10	1.55	1 (11%)	9,12,14	2.12	2 (22%)
4	SEP	V	363	4	9,9,10	1.54	2 (22%)	9,12,14	1.58	2 (22%)
4	SEP	V	364	4	9,9,10	1.54	2 (22%)	9,12,14	1.25	1 (11%)

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
4	TPO	V	347	4	-	0/8/11/13	0/0/0/0
4	SEP	V	350	4	-	0/5/8/10	0/0/0/0
4	SEP	V	357	4	-	0/5/8/10	0/0/0/0
4	TPO	V	359	4	-	0/8/11/13	0/0/0/0
4	TPO	V	360	4	-	0/8/11/13	0/0/0/0
4	SEP	V	362	4	-	0/5/8/10	0/0/0/0
4	SEP	V	363	4	-	0/5/8/10	0/0/0/0
4	SEP	V	364	4	-	0/5/8/10	0/0/0/0

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	V	363	SEP	CA-C	2.00	1.52	1.50
4	V	364	SEP	CA-C	2.04	1.52	1.50
4	V	347	TPO	CA-C	2.06	1.53	1.50
4	V	350	SEP	CA-C	2.18	1.53	1.50
4	V	363	SEP	P-O1P	3.20	1.61	1.50
4	V	350	SEP	P-O1P	3.21	1.61	1.50
4	V	364	SEP	P-O1P	3.23	1.61	1.50
4	V	357	SEP	P-O1P	3.24	1.61	1.50
4	V	362	SEP	P-O1P	3.26	1.61	1.50

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	V	347	TPO	P-OG1-CB	-5.33	107.12	123.21
4	V	360	TPO	P-OG1-CB	-4.91	108.38	123.21
4	V	362	SEP	P-OG-CB	-3.77	107.90	118.30
4	V	357	SEP	P-OG-CB	-3.21	109.46	118.30
4	V	350	SEP	P-OG-CB	-2.86	110.41	118.30
4	V	363	SEP	P-OG-CB	-2.85	110.44	118.30
4	V	364	SEP	P-OG-CB	-2.46	111.52	118.30
4	V	359	TPO	P-OG1-CB	-2.37	116.06	123.21
4	V	347	TPO	O-C-CA	-2.13	120.12	125.09
4	V	363	SEP	OG-CB-CA	2.91	111.04	108.17
4	V	357	SEP	OG-CB-CA	3.34	111.46	108.17
4	V	350	SEP	OG-CB-CA	3.44	111.55	108.17
4	V	362	SEP	OG-CB-CA	4.45	112.55	108.17

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	V	363	SEP	1	0
4	V	364	SEP	1	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 9 ligands modelled in this entry, 2 are monoatomic - leaving 7 for Mogul analysis.

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	EDO	A	402	-	3,3,3	0.44	0	2,2,2	0.36	0
6	EDO	H	302	-	3,3,3	0.46	0	2,2,2	0.38	0
6	EDO	H	303	-	3,3,3	0.47	0	2,2,2	0.37	0
6	EDO	H	304	-	3,3,3	0.44	0	2,2,2	0.28	0
6	EDO	L	301	-	3,3,3	0.47	0	2,2,2	0.38	0
6	EDO	L	302	-	3,3,3	0.47	0	2,2,2	0.35	0
7	PRO	L	303	-	5,8,8	0.50	0	6,10,10	1.05	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	EDO	A	402	-	-	0/1/1/1	0/0/0/0
6	EDO	H	302	-	-	0/1/1/1	0/0/0/0
6	EDO	H	303	-	-	0/1/1/1	0/0/0/0

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	EDO	H	304	-	-	0/1/1/1	0/0/0/0
6	EDO	L	301	-	-	0/1/1/1	0/0/0/0
6	EDO	L	302	-	-	0/1/1/1	0/0/0/0
7	PRO	L	303	-	-	0/0/11/11	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	402	EDO	2	0

## 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

In the following table, the column labelled ‘#RSRZ > 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q < 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	353/401 (88%)	-0.19	3 (0%) 86 84	38, 67, 126, 162	0
2	H	204/237 (86%)	0.23	12 (5%) 22 16	45, 94, 159, 182	0
3	L	185/215 (86%)	0.37	16 (8%) 10 7	58, 102, 156, 178	1 (0%)
4	V	13/29 (44%)	1.10	1 (7%) 13 9	66, 94, 133, 142	0
All	All	755/882 (85%)	0.08	32 (4%) 36 28	38, 80, 148, 182	1 (0%)

All (32) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	L	149	TRP	5.4
2	H	194	VAL	4.9
3	L	159	ASN	4.0
3	L	136	LEU	3.7
2	H	206	CYS	3.6
1	A	361	PRO	3.6
1	A	44	ASP	3.3
2	H	136	PRO	3.2
2	H	170	THR	3.1
3	L	182	LEU	3.1
3	L	117	PHE	3.0
3	L	187	TYR	2.9
2	H	166	SER	2.9
2	H	208	VAL	2.9
2	H	168	ALA	2.9
2	H	164	TRP	2.8
3	L	145	ALA	2.7
4	V	366	ALA	2.7
3	L	1	SER	2.6
2	H	169	LEU	2.6
1	A	45	PRO	2.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
3	L	119	PHE	2.5
3	L	160	SER	2.5
3	L	157	SER	2.4
2	H	167	GLY	2.4
3	L	180	LEU	2.4
3	L	176	LEU	2.4
2	H	147	ALA	2.3
2	H	165	ASN	2.3
3	L	118	ILE	2.2
3	L	137	LEU	2.1
3	L	158	GLY	2.1

## 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
4	TPO	V	347	11/12	0.87	0.16	88,119,144,170	0
4	SEP	V	357	10/11	0.88	0.13	69,91,133,146	0
4	TPO	V	359	11/12	0.92	0.11	77,83,182,184	0
4	SEP	V	364	10/11	0.93	0.16	69,79,95,105	0
4	SEP	V	362	10/11	0.95	0.13	64,71,77,78	0
4	SEP	V	350	10/11	0.95	0.14	63,75,104,128	0
4	SEP	V	363	10/11	0.96	0.14	64,73,84,88	0
4	TPO	V	360	11/12	0.97	0.12	55,62,75,75	0

## 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
6	EDO	L	302	4/4	0.71	0.39	88,90,98,101	0
6	EDO	H	303	4/4	0.72	0.42	70,96,102,105	0
6	EDO	H	302	4/4	0.79	0.26	72,85,88,101	0
6	EDO	H	304	4/4	0.86	0.38	84,107,118,119	0
5	CL	A	401	1/1	0.87	0.17	84,84,84,84	1
6	EDO	L	301	4/4	0.87	0.20	88,89,90,93	0
5	CL	H	301	1/1	0.90	0.12	86,86,86,86	0
7	PRO	L	303	8/8	0.91	0.38	61,127,150,161	0
6	EDO	A	402	4/4	0.96	0.39	68,90,92,101	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.