



wwPDB X-ray Structure Validation Summary Report ⓘ

Jul 6, 2017 – 05:04 PM EDT

PDB ID : 3SN6
Title : Crystal structure of the beta2 adrenergic receptor-Gs protein complex
Authors : Rasmussen, S.G.F.; DeVree, B.T.; Zou, Y.; Kruse, A.C.; Chung, K.Y.; Kobilka, T.S.; Thian, F.S.; Chae, P.S.; Pardon, E.; Calinski, D.; Mathiesen, J.M.; Shah, S.T.A.; Lyons, J.A.; Caffrey, M.; Gellman, S.H.; Steyaert, J.; Skinotitis, G.; Weis, W.I.; Sunahara, R.K.; Kobilka, B.K.
Deposited on : unknown
Resolution : 3.20 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

<http://wwpdb.org/validation/2016/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	:	1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20029824
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20029824

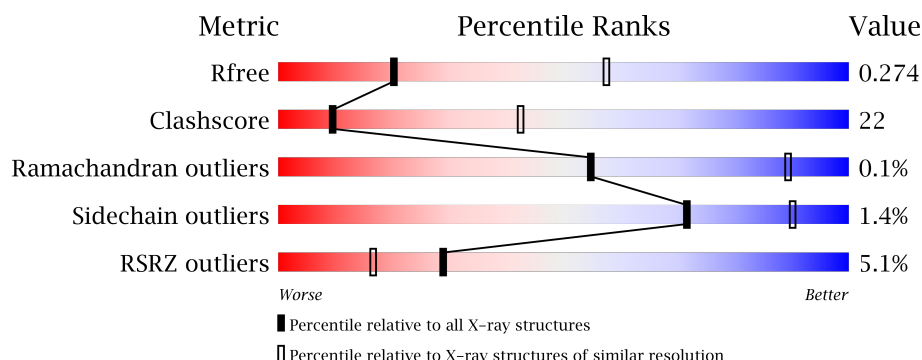
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.20 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}	100719	1015 (3.22-3.18)
Clashscore	112137	1009 (3.20-3.20)
Ramachandran outliers	110173	1118 (3.22-3.18)
Sidechain outliers	110143	1117 (3.22-3.18)
RSRZ outliers	101464	1020 (3.22-3.18)

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 ≥ 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	380	<div> <div>6%</div> <div>58%</div> <div>33%</div> <div>8%</div> </div>
2	B	351	<div> <div>%</div> <div>57%</div> <div>39%</div> <div>%</div> </div>
3	G	68	<div> <div>%</div> <div>68%</div> <div>18%</div> <div>15%</div> </div>
4	R	514	<div> <div>8%</div> <div>49%</div> <div>36%</div> <div>14%</div> </div>
5	N	138	<div> <div>51%</div> <div>41%</div> <div>7%</div> </div>

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 criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
6	P0G	R	1601	-	-	-	X

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 10274 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 Guanine nucleotide-binding protein G(s) subunit alpha isoforms short.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	349	Total	C	N	O	S	0	0	0
			2814	1786	494	522	12			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	86	SER	GLY	ENGINEERED MUTATION	UNP P04896

- Molecule 2 is a protein called Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	340	Total	C	N	O	S	0	0	0
			2592	1600	463	508	21			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-10	MET	-	expression tag	UNP P54311
B	-9	HIS	-	expression tag	UNP P54311
B	-8	HIS	-	expression tag	UNP P54311
B	-7	HIS	-	expression tag	UNP P54311
B	-6	HIS	-	expression tag	UNP P54311
B	-5	HIS	-	expression tag	UNP P54311
B	-4	HIS	-	expression tag	UNP P54311
B	-3	GLY	-	expression tag	UNP P54311
B	-2	SER	-	expression tag	UNP P54311
B	-1	LEU	-	expression tag	UNP P54311
B	0	LEU	-	expression tag	UNP P54311
B	1	GLN	-	expression tag	UNP P54311

- Molecule 3 is a protein called Guanine nucleotide-binding protein G(I)/G(S)/G(O) subunit gamma-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	G	58	Total	C	N	O	S	0	0	0
			438	274	76	85	3			

- Molecule 4 is a protein called Endolysin,Beta-2 adrenergic receptor.

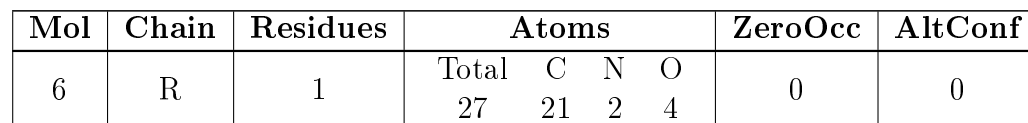
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	R	443	Total	C	N	O	S	0	0	0
			3433	2234	572	605	22			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
R	987	ASP	-	expression tag	UNP P00720
R	988	TYR	-	expression tag	UNP P00720
R	989	LYS	-	expression tag	UNP P00720
R	990	ASP	-	expression tag	UNP P00720
R	991	ASP	-	expression tag	UNP P00720
R	992	ASP	-	expression tag	UNP P00720
R	993	ASP	-	expression tag	UNP P00720
R	994	ALA	-	expression tag	UNP P00720
R	995	GLU	-	expression tag	UNP P00720
R	996	ASN	-	expression tag	UNP P00720
R	997	LEU	-	expression tag	UNP P00720
R	998	TYR	-	expression tag	UNP P00720
R	999	PHE	-	expression tag	UNP P00720
R	1000	GLN	-	expression tag	UNP P00720
R	1001	GLY	-	expression tag	UNP P00720
R	1012	GLY	ARG	conflict	UNP P00720
R	1054	THR	CYS	engineered mutation	UNP P00720
R	1097	ALA	CYS	engineered mutation	UNP P00720
R	1137	ARG	ILE	conflict	UNP P00720
R	?	ALA	-	linker	UNP P00720
R	?	ALA	-	linker	UNP P00720
R	96	THR	MET	engineered mutation	UNP P07550
R	98	THR	MET	engineered mutation	UNP P07550
R	187	GLU	ASN	engineered mutation	UNP P07550

- Molecule 5 is a protein called Camelid antibody VHH fragment.

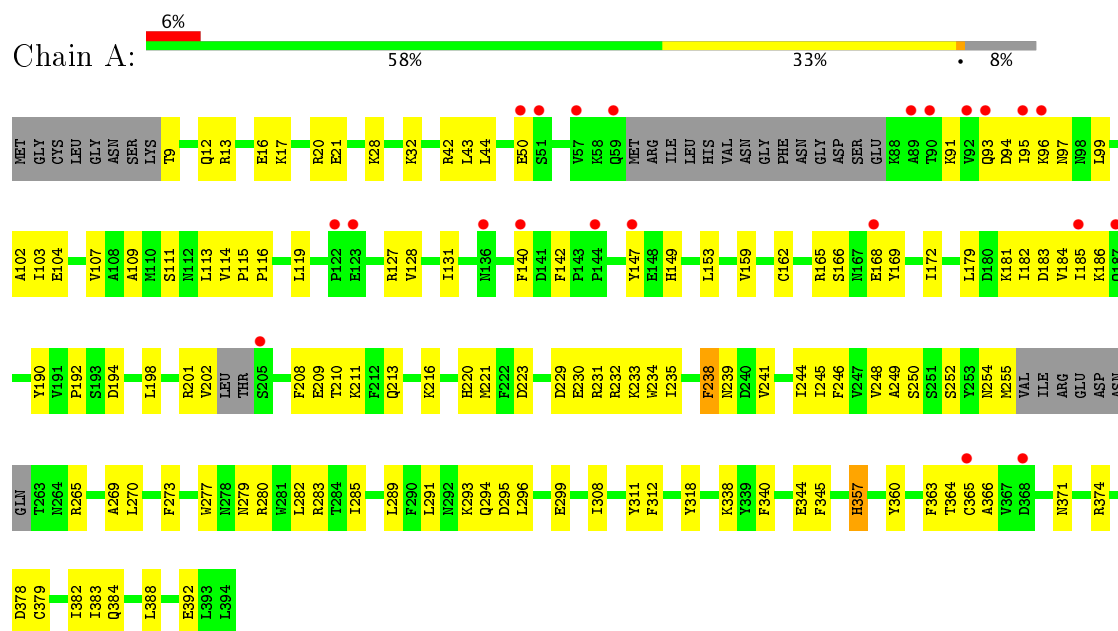
- Molecule 6 is 8-[(1R)-2-[[1,1-dimethyl-2-(2-methylphenyl)ethyl]amino]-1-hydroxyethyl]-5-hydroxy-2H-1,4-benzoxazin-3(4H)-one (three-letter code: P0G) (formula: $C_{21}H_{26}N_2O_4$).



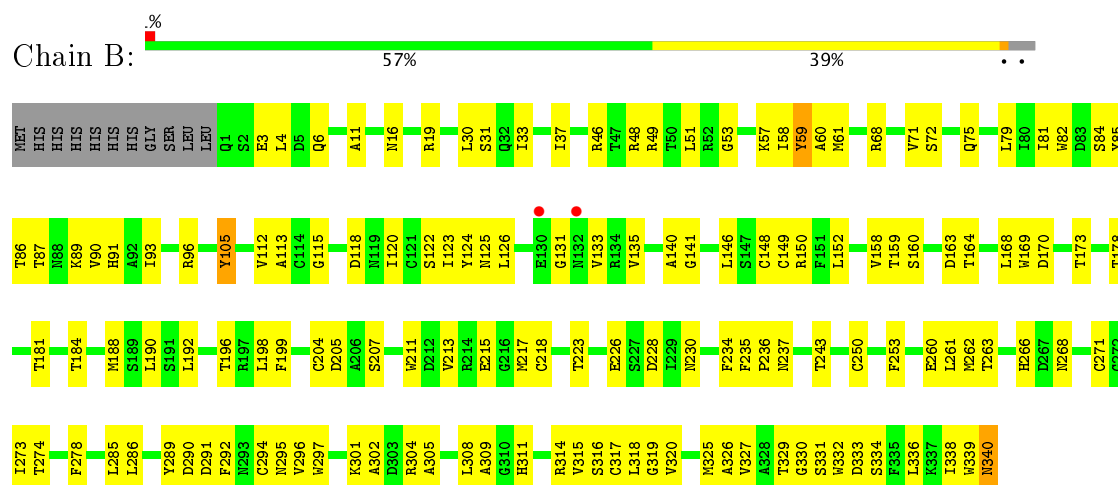
3 Residue-property plots

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.

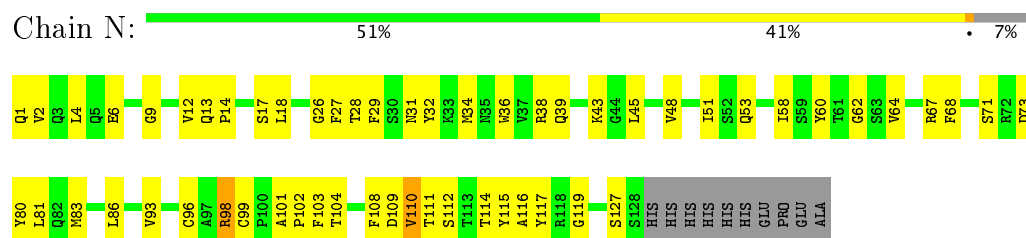
- Molecule 1: Guanine nucleotide-binding protein G(s) subunit alpha isoforms short



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



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



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	119.34Å 64.56Å 131.24Å 90.00° 91.67° 90.00°	Depositor
Resolution (Å)	40.67 – 3.20 40.68 – 3.20	Depositor EDS
% Data completeness (in resolution range)	92.8 (40.67-3.20) 92.9 (40.68-3.20)	Depositor EDS
R_{merge}	0.16	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.80 (at 3.18Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7_650)	Depositor
R, R_{free}	0.226 , 0.277 0.225 , 0.274	Depositor DCC
R_{free} test set	1556 reflections (5.01%)	DCC
Wilson B-factor (Å ²)	68.1	Xtriage
Anisotropy	0.051	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.25 , 61.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.028 for h,-k,-l	Xtriage
F_o, F_c correlation	0.87	EDS
Total number of atoms	10274	wwPDB-VP
Average B, all atoms (Å ²)	92.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.93% 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: P0G

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	0.23	0/2870	0.40	0/3879
2	B	0.23	0/2639	0.46	0/3580
3	G	0.22	0/444	0.38	0/601
4	R	0.23	0/3502	0.40	0/4762
5	N	0.28	0/990	0.53	1/1341 (0.1%)
All	All	0.23	0/10445	0.43	1/14163 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	N	99	CYS	CA-CB-SG	-5.45	104.19	114.00

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	2814	0	2719	110	0
2	B	2592	0	2485	134	0
3	G	438	0	443	11	0
4	R	3433	0	3429	143	0
5	N	970	0	930	59	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	R	27	0	25	7	0
All	All	10274	0	10031	440	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 22.

The worst 5 of 440 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)
4:R:132:TYR:HB2	4:R:218:VAL:HG13	1.43	0.99
4:R:137:SER:HB3	4:R:140:LYS:HD3	1.43	0.97
1:A:103:ILE:HG22	1:A:179:LEU:HD21	1.51	0.93
1:A:114:VAL:HB	1:A:115:PRO:HD3	1.48	0.92
4:R:293:ASN:HD21	6:R:1601:P0G:HAMA	1.39	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 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	341/380 (90%)	328 (96%)	13 (4%)	0	100	100
2	B	338/351 (96%)	323 (96%)	15 (4%)	0	100	100
3	G	56/68 (82%)	54 (96%)	2 (4%)	0	100	100
4	R	435/514 (85%)	407 (94%)	28 (6%)	0	100	100
5	N	126/138 (91%)	117 (93%)	8 (6%)	1 (1%)	22	65
All	All	1296/1451 (89%)	1229 (95%)	66 (5%)	1 (0%)	55	89

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
5	N	110	VAL

5.3.2 Protein sidechains ⓘ

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	296/342 (86%)	294 (99%)	2 (1%)	87	96
2	B	278/293 (95%)	273 (98%)	5 (2%)	64	87
3	G	46/56 (82%)	46 (100%)	0	100	100
4	R	358/441 (81%)	352 (98%)	6 (2%)	66	88
5	N	104/115 (90%)	102 (98%)	2 (2%)	62	86
All	All	1082/1247 (87%)	1067 (99%)	15 (1%)	71	90

5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	340	ASN
4	R	1104	PHE
4	R	336	PHE
2	B	234	PHE
4	R	332	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	293	ASN
3	G	5	ASN
5	N	1	GLN
2	B	268	ASN
5	N	13	GLN

5.3.3 RNA ⓘ

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 carbohydrates 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	$\# Z > 2$
6	P0G	R	1601	-	28,29,29	1.30	3 (10%)	36,42,42	1.45	4 (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
6	P0G	R	1601	-	-	0/15/24/24	0/3/3/3

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	R	1601	P0G	CAW-CAZ	-4.88	1.48	1.52
6	R	1601	P0G	CBA-NAP	-3.53	1.45	1.49
6	R	1601	P0G	CAX-NAQ	-2.14	1.35	1.39

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	R	1601	P0G	CAX-NAQ-CAS	-3.26	120.34	124.53

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	R	1601	P0G	OAR-CAY-CAX	-2.86	117.44	121.21
6	R	1601	P0G	OAR-CAY-CAW	3.64	121.75	116.59
6	R	1601	P0G	CAN-NAP-CBA	5.10	122.01	116.34

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	R	1601	P0G	7	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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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, 95th 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 ²)	Q<0.9
1	A	349/380 (91%)	0.32	22 (6%) 21 12	32, 93, 156, 191	0
2	B	340/351 (96%)	-0.15	2 (0%) 89 83	29, 59, 101, 162	0
3	G	58/68 (85%)	-0.20	1 (1%) 70 57	36, 80, 132, 146	0
4	R	443/514 (86%)	0.49	42 (9%) 9 5	56, 122, 175, 214	0
5	N	128/138 (92%)	-0.22	0 100 100	31, 59, 93, 115	0
All	All	1318/1451 (90%)	0.18	67 (5%) 29 16	29, 89, 158, 214	0

The worst 5 of 67 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	R	192	ASP	5.4
1	A	147	TYR	4.6
1	A	136	ASN	4.2
4	R	294	ILE	4.0
4	R	144	LEU	4.0

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

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

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. LLDF column lists the quality of electron

density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th 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	LLDF	B-factors(\AA^2)	Q<0.9
6	P0G	R	1601	27/27	0.60	0.57	1.88	152,166,192,195	0

6.5 Other polymers [i](#)

There are no such residues in this entry.