



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 17, 2016 – 03:57 PM EDT

PDB ID : 5G53
Title : Structure of the adenosine A2A receptor bound to an engineered G protein
Authors : Carpenter, B.; Nehme, R.; Warne, T.; Leslie, A.G.W.; Tate, C.G.
Deposited on : 2016-05-19
Resolution : 3.40 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 : 1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20027939
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
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-20027939

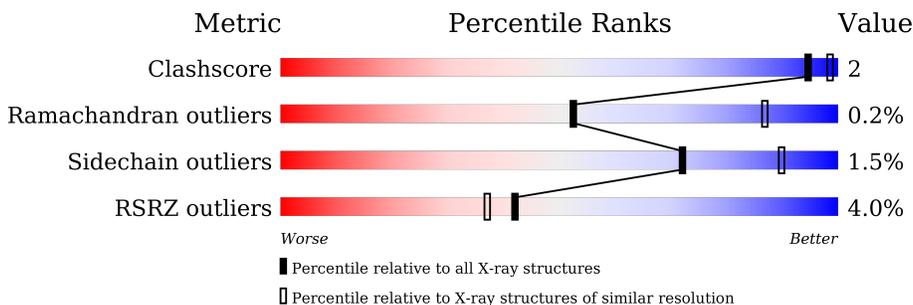
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.40 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))
Clashscore	102246	1611 (3.50-3.30)
Ramachandran outliers	100387	1571 (3.50-3.30)
Sidechain outliers	100360	1571 (3.50-3.30)
RSRZ outliers	91569	1485 (3.50-3.30)

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	314	 6% 85% 5% 10%
1	B	314	 6% 87% 10%
2	C	229	 3% 79% 6% 14%
2	D	229	 5% 78% 19%

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 7359 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 ADENOSINE RECEPTOR A2A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	283	Total	C	N	O	S	0	0	0
			2186	1458	353	355	20			
1	B	284	Total	C	N	O	S	0	0	0
			2142	1428	343	352	19			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	154	ALA	ASN	ENGINEERED MUTATION	UNP P29274
A	309	GLU	-	EXPRESSION TAG	UNP P29274
A	310	ASN	-	EXPRESSION TAG	UNP P29274
A	311	LEU	-	EXPRESSION TAG	UNP P29274
A	312	TYR	-	EXPRESSION TAG	UNP P29274
A	313	PHE	-	EXPRESSION TAG	UNP P29274
A	314	GLN	-	EXPRESSION TAG	UNP P29274
B	154	ALA	ASN	ENGINEERED MUTATION	UNP P29274
B	309	GLU	-	EXPRESSION TAG	UNP P29274
B	310	ASN	-	EXPRESSION TAG	UNP P29274
B	311	LEU	-	EXPRESSION TAG	UNP P29274
B	312	TYR	-	EXPRESSION TAG	UNP P29274
B	313	PHE	-	EXPRESSION TAG	UNP P29274
B	314	GLN	-	EXPRESSION TAG	UNP P29274

- Molecule 2 is a protein called ENGINEERED DOMAIN OF HUMAN G ALPHA S LONG ISOFORM.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	C	196	Total	C	N	O	S	0	0	0
			1515	971	261	278	5			
2	D	185	Total	C	N	O	S	0	0	0
			1404	899	236	264	5			

There are 60 discrepancies between the modelled and reference sequences:

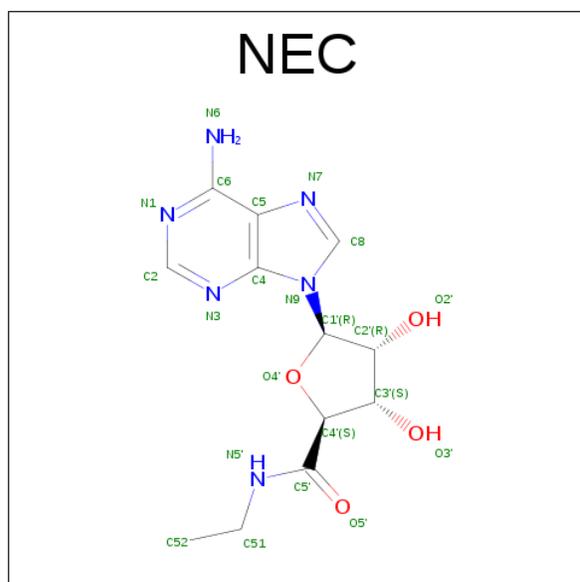
Chain	Residue	Modelled	Actual	Comment	Reference
C	25	GLY	-	EXPRESSION TAG	UNP P63092
C	49	ASP	GLY	ENGINEERED MUTATION	UNP P63092
C	50	ASN	GLU	ENGINEERED MUTATION	UNP P63092
C	61	ARG	-	LINKER	UNP P63092
C	193	ILE	-	LINKER	UNP P63092
C	194	TYR	-	LINKER	UNP P63092
C	195	HIS	-	LINKER	UNP P63092
C	196	GLY	-	LINKER	UNP P63092
C	197	GLY	-	LINKER	UNP P63092
C	198	SER	-	LINKER	UNP P63092
C	199	GLY	-	LINKER	UNP P63092
C	200	GLY	-	LINKER	UNP P63092
C	201	SER	-	LINKER	UNP P63092
C	202	GLY	-	LINKER	UNP P63092
C	203	GLY	-	LINKER	UNP P63092
C	249	ASP	ALA	ENGINEERED MUTATION	UNP Q5JWF2
C	252	ASP	SER	ENGINEERED MUTATION	UNP Q5JWF2
C	.	-	ASN	DELETION	UNP Q5JWF2
C	.	-	MET	DELETION	UNP Q5JWF2
C	.	-	VAL	DELETION	UNP Q5JWF2
C	.	-	ILE	DELETION	UNP Q5JWF2
C	.	-	ARG	DELETION	UNP Q5JWF2
C	.	-	GLU	DELETION	UNP Q5JWF2
C	.	-	ASP	DELETION	UNP Q5JWF2
C	.	-	ASN	DELETION	UNP Q5JWF2
C	.	-	GLN	DELETION	UNP Q5JWF2
C	.	-	THR	DELETION	UNP Q5JWF2
C	272	ASP	LEU	ENGINEERED MUTATION	UNP Q5JWF2
C	372	ALA	ILE	ENGINEERED MUTATION	UNP Q5JWF2
C	375	ILE	VAL	ENGINEERED MUTATION	UNP Q5JWF2
D	25	GLY	-	EXPRESSION TAG	UNP P63092
D	49	ASP	GLY	ENGINEERED MUTATION	UNP P63092
D	50	ASN	GLU	ENGINEERED MUTATION	UNP P63092
D	61	ARG	-	LINKER	UNP P63092
D	193	ILE	-	LINKER	UNP P63092
D	194	TYR	-	LINKER	UNP P63092
D	195	HIS	-	LINKER	UNP P63092
D	196	GLY	-	LINKER	UNP P63092
D	197	GLY	-	LINKER	UNP P63092
D	198	SER	-	LINKER	UNP P63092
D	199	GLY	-	LINKER	UNP P63092
D	200	GLY	-	LINKER	UNP P63092
D	201	SER	-	LINKER	UNP P63092

Continued on next page...

Continued from previous page...

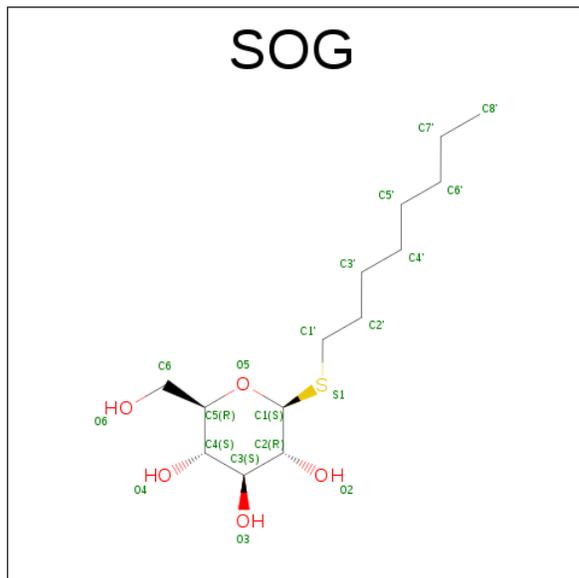
Chain	Residue	Modelled	Actual	Comment	Reference
D	202	GLY	-	LINKER	UNP P63092
D	203	GLY	-	LINKER	UNP P63092
D	249	ASP	ALA	ENGINEERED MUTATION	UNP Q5JWF2
D	252	ASP	SER	ENGINEERED MUTATION	UNP Q5JWF2
D	.	-	ASN	DELETION	UNP Q5JWF2
D	.	-	MET	DELETION	UNP Q5JWF2
D	.	-	VAL	DELETION	UNP Q5JWF2
D	.	-	ILE	DELETION	UNP Q5JWF2
D	.	-	ARG	DELETION	UNP Q5JWF2
D	.	-	GLU	DELETION	UNP Q5JWF2
D	.	-	ASP	DELETION	UNP Q5JWF2
D	.	-	ASN	DELETION	UNP Q5JWF2
D	.	-	GLN	DELETION	UNP Q5JWF2
D	.	-	THR	DELETION	UNP Q5JWF2
D	272	ASP	LEU	ENGINEERED MUTATION	UNP Q5JWF2
D	372	ALA	ILE	ENGINEERED MUTATION	UNP Q5JWF2
D	375	ILE	VAL	ENGINEERED MUTATION	UNP Q5JWF2

- Molecule 3 is N-ETHYL-5'-CARBOXAMIDO ADENOSINE (three-letter code: NEC) (formula: C₁₂H₁₆N₆O₄).



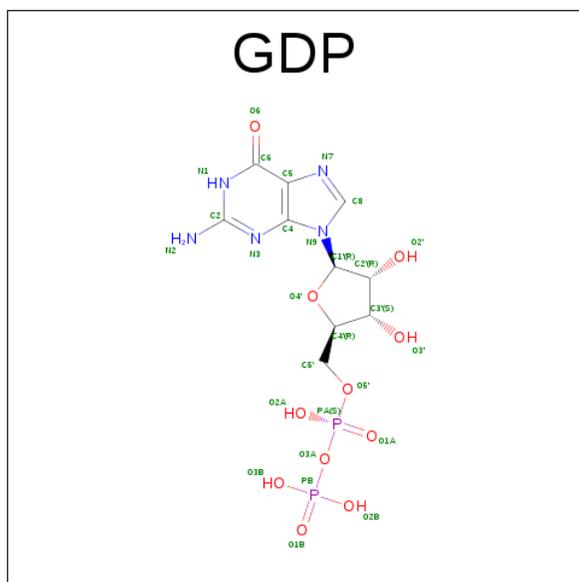
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	
			Total	C	N			O
3	A	1	Total	C	N	O	0	0
			22	12	6	4		
3	B	1	Total	C	N	O	0	0
			22	12	6	4		

- Molecule 4 is SUGAR (2-HYDROXYMETHYL-6-OCTYLSULFANYL-TETRAHYDRO-PYRAN-3,4,5-TRIOL) (three-letter code: SOG) (formula: $C_{14}H_{28}O_5S$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	O	S		
4	A	1	Total	C	O	S	0	0
			20	14	5	1		
4	A	1	Total	C	O	S	0	0
			20	14	5	1		

- Molecule 5 is GUANOSINE-5'-DIPHOSPHATE (three-letter code: GDP) (formula: $C_{10}H_{15}N_5O_{11}P_2$).

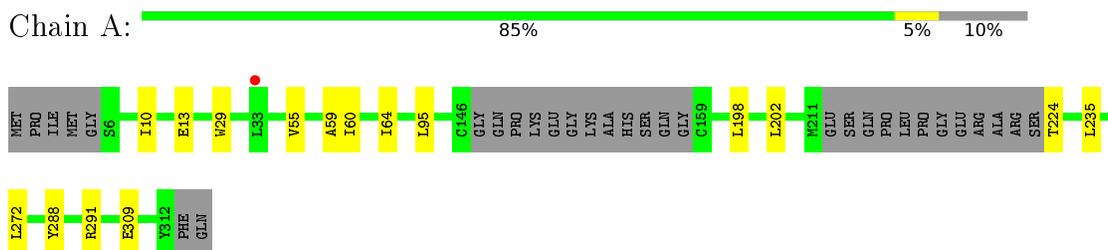


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
5	C	1	28	10	5	11	2	0	0

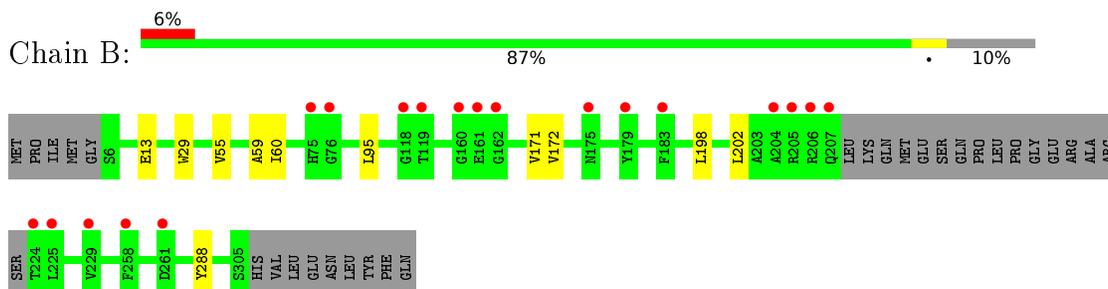
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

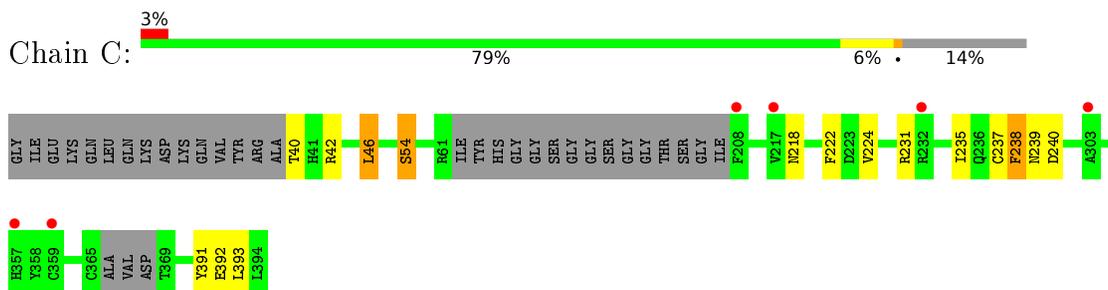
- Molecule 1: ADENOSINE RECEPTOR A2A



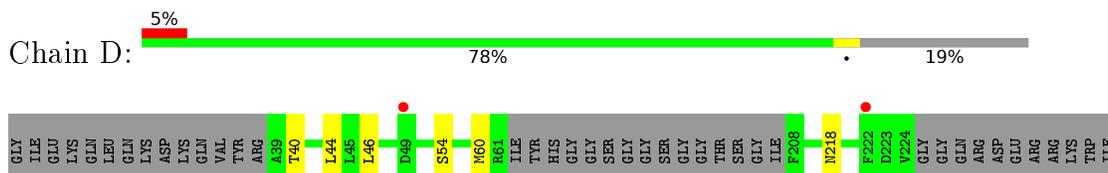
- Molecule 1: ADENOSINE RECEPTOR A2A



- Molecule 2: ENGINEERED DOMAIN OF HUMAN G ALPHA S LONG ISOFORM



- Molecule 2: ENGINEERED DOMAIN OF HUMAN G ALPHA S LONG ISOFORM





4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	90.63Å 111.81Å 161.30Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	91.89 – 3.40 40.33 – 3.40	Depositor EDS
% Data completeness (in resolution range)	89.9 (91.89-3.40) 90.0 (40.33-3.40)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.47 (at 3.40Å)	Xtrriage
Refinement program	REFMAC 5.8.0144	Depositor
R, R_{free}	0.284 , 0.315 0.287 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	DCC
Wilson B-factor (Å ²)	79.1	Xtrriage
Anisotropy	0.079	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.25 , 28.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.40$, $\langle L^2 \rangle = 0.22$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.82	EDS
Total number of atoms	7359	wwPDB-VP
Average B, all atoms (Å ²)	79.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.76% 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: GDP, NEC, SOG

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	0.41	0/2240	0.56	0/3062
1	B	0.40	0/2197	0.53	0/3009
2	C	0.42	0/1550	0.61	1/2107 (0.0%)
2	D	0.42	0/1433	0.60	0/1952
All	All	0.41	0/7420	0.57	1/10130 (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	C	1	0

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
2	C	239	ASN	N-CA-C	5.89	126.90	111.00

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	C	239	ASN	CA

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	2186	0	2237	8	0
1	B	2142	0	2161	5	0
2	C	1515	0	1382	12	0
2	D	1404	0	1277	2	0
3	A	22	0	16	0	0
3	B	22	0	16	0	0
4	A	40	0	56	0	0
5	C	28	0	12	1	0
All	All	7359	0	7157	24	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 2.

The worst 5 of 24 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:C:237:CYS:O	2:C:238:PHE:HB2	1.87	0.74
1:A:235:LEU:HD13	2:C:393:LEU:HD21	1.70	0.73
2:C:54:SER:N	5:C:400:GDP:O2B	2.30	0.65
2:C:222:PHE:CD1	2:C:238:PHE:HZ	2.19	0.60
1:A:235:LEU:CD1	2:C:393:LEU:HD21	2.33	0.58

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	277/314 (88%)	270 (98%)	7 (2%)	0	100	100
1	B	280/314 (89%)	273 (98%)	7 (2%)	0	100	100
2	C	190/229 (83%)	182 (96%)	6 (3%)	2 (1%)	17	61
2	D	179/229 (78%)	172 (96%)	7 (4%)	0	100	100
All	All	926/1086 (85%)	897 (97%)	27 (3%)	2 (0%)	52	87

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	238	PHE
2	C	224	VAL

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	233/264 (88%)	229 (98%)	4 (2%)	68	88
1	B	223/264 (84%)	222 (100%)	1 (0%)	93	97
2	C	149/201 (74%)	146 (98%)	3 (2%)	63	87
2	D	138/201 (69%)	135 (98%)	3 (2%)	60	86
All	All	743/930 (80%)	732 (98%)	11 (2%)	72	90

5 of 11 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	29	TRP
2	C	46	LEU
2	D	46	LEU
1	A	309	GLU
2	C	392	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	157	GLN
2	D	41	HIS
1	B	280	ASN
1	A	284	ASN
1	B	284	ASN

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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

5 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
3	NEC	A	400	-	21,24,24	1.01	1 (4%)	17,35,35	2.37	4 (23%)
4	SOG	A	501	-	20,20,20	0.95	1 (5%)	23,25,25	1.49	3 (13%)
4	SOG	A	502	-	20,20,20	1.00	2 (10%)	23,25,25	0.76	1 (4%)
3	NEC	B	400	-	21,24,24	0.98	1 (4%)	17,35,35	2.36	3 (17%)
5	GDP	C	400	-	24,30,30	1.33	2 (8%)	26,47,47	1.81	5 (19%)

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
3	NEC	A	400	-	-	0/7/27/27	0/3/3/3
4	SOG	A	501	-	-	0/11/31/31	0/1/1/1
4	SOG	A	502	-	-	0/11/31/31	0/1/1/1
3	NEC	B	400	-	-	0/7/27/27	0/3/3/3
5	GDP	C	400	-	-	0/12/32/32	0/3/3/3

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	502	SOG	C1'-S1	-3.57	1.76	1.81
4	A	501	SOG	C1'-S1	-3.31	1.76	1.81
4	A	502	SOG	C1-S1	-2.20	1.77	1.80
3	B	400	NEC	C5-C4	2.93	1.47	1.40
3	A	400	NEC	C5-C4	2.98	1.47	1.40

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	400	NEC	N3-C2-N1	-7.91	122.66	128.87
3	A	400	NEC	N3-C2-N1	-7.88	122.68	128.87
5	C	400	GDP	C5-C6-N1	-4.29	117.92	123.52
5	C	400	GDP	N3-C2-N1	-3.50	122.80	127.56
3	B	400	NEC	C1'-N9-C4	-2.98	123.48	126.81

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	C	400	GDP	1	0

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues

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	283/314 (90%)	-0.26	1 (0%) 93 91	40, 67, 94, 134	0
1	B	284/314 (90%)	0.07	19 (6%) 21 20	51, 78, 113, 126	0
2	C	196/229 (85%)	0.01	6 (3%) 52 48	51, 87, 119, 131	0
2	D	185/229 (80%)	0.24	12 (6%) 22 21	54, 86, 117, 133	0
All	All	948/1086 (87%)	-0.01	38 (4%) 42 37	40, 77, 113, 134	0

The worst 5 of 38 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	306	SER	5.0
1	B	119	THR	4.1
2	D	311	TYR	3.8
1	B	161	GLU	3.7
1	B	258	PHE	3.6

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
3	NEC	A	400	22/22	0.94	0.27	1.90	56,62,67,68	0
3	NEC	B	400	22/22	0.96	0.24	0.08	68,73,77,78	0
5	GDP	C	400	28/28	0.85	0.16	-0.18	62,69,73,73	0
4	SOG	A	501	20/20	0.94	0.17	-0.18	59,65,71,71	0
4	SOG	A	502	20/20	0.75	0.26	-	97,140,148,148	0

6.5 Other polymers [i](#)

There are no such residues in this entry.