



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 06:08 PM GMT

PDB ID : 4KNG
Title : Crystal structure of human LGR5-RSPO1-RNF43
Authors : Chen, P.H.; He, X.
Deposited on : 2013-05-09
Resolution : 2.50 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 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
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) : trunk26865

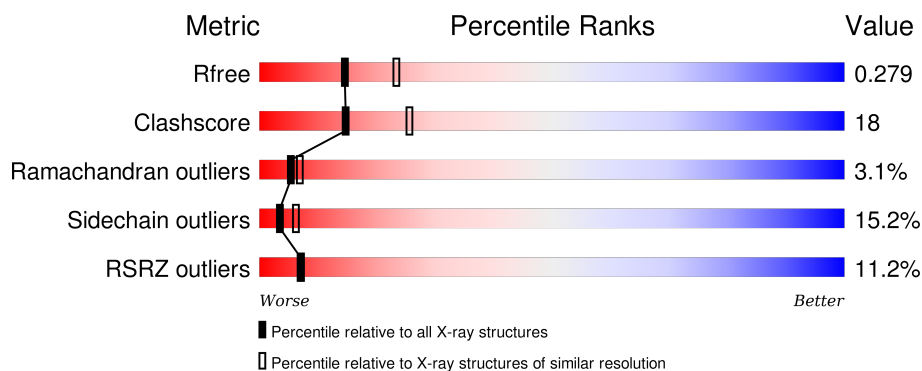
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.50 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.



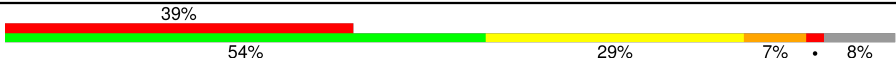
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	3553 (2.50-2.50)
Clashscore	102246	4242 (2.50-2.50)
Ramachandran outliers	100387	4156 (2.50-2.50)
Sidechain outliers	100360	4158 (2.50-2.50)
RSRZ outliers	91569	3562 (2.50-2.50)

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	531	<div> <div>8%</div> <div> <div></div> <div>54%</div> <div>26%</div> <div>6% •</div> <div>13%</div> </div> </div>
1	B	531	<div> <div>5%</div> <div> <div></div> <div>55%</div> <div>24%</div> <div>7% •</div> <div>13%</div> </div> </div>
2	M	115	<div> <div>10%</div> <div> <div></div> <div>64%</div> <div>17%</div> <div>•</div> <div>15%</div> </div> </div>
2	P	115	<div> <div>10%</div> <div> <div></div> <div>53%</div> <div>23%</div> <div>6%</div> <div>18%</div> </div> </div>
3	E	160	<div> <div>6%</div> <div> <div></div> <div>54%</div> <div>31%</div> <div>8% •</div> <div>7%</div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	160	

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 11252 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 Leucine-rich repeat-containing G-protein coupled receptor 5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	461	Total	C	N	O	S	0	1	0
			3622	2307	626	673	16			
1	B	463	Total	C	N	O	S	0	0	0
			3629	2312	627	674	16			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	30	GLY	-	EXPRESSION TAG	UNP O75473
A	31	PRO	-	EXPRESSION TAG	UNP O75473
A	558	ALA	-	EXPRESSION TAG	UNP O75473
A	559	ALA	-	EXPRESSION TAG	UNP O75473
A	560	ALA	-	EXPRESSION TAG	UNP O75473
B	30	GLY	-	EXPRESSION TAG	UNP O75473
B	31	PRO	-	EXPRESSION TAG	UNP O75473
B	558	ALA	-	EXPRESSION TAG	UNP O75473
B	559	ALA	-	EXPRESSION TAG	UNP O75473
B	560	ALA	-	EXPRESSION TAG	UNP O75473

- Molecule 2 is a protein called R-spondin-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	M	98	Total	C	N	O	S	0	0	0
			750	465	131	136	18			
2	P	94	Total	C	N	O	S	0	0	0
			720	448	127	128	17			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
M	33	GLY	-	EXPRESSION TAG	UNP Q2MKA7
M	34	PRO	-	EXPRESSION TAG	UNP Q2MKA7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
M	145	ALA	-	EXPRESSION TAG	UNP Q2MKA7
M	146	ALA	-	EXPRESSION TAG	UNP Q2MKA7
M	147	ALA	-	EXPRESSION TAG	UNP Q2MKA7
P	33	GLY	-	EXPRESSION TAG	UNP Q2MKA7
P	34	PRO	-	EXPRESSION TAG	UNP Q2MKA7
P	145	ALA	-	EXPRESSION TAG	UNP Q2MKA7
P	146	ALA	-	EXPRESSION TAG	UNP Q2MKA7
P	147	ALA	-	EXPRESSION TAG	UNP Q2MKA7

- Molecule 3 is a protein called E3 ubiquitin-protein ligase RNF43.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	149	Total	C	N	O	S	0	0	0
			1152	734	200	212	6			
3	F	148	Total	C	N	O	S	0	0	0
			1145	729	199	211	6			

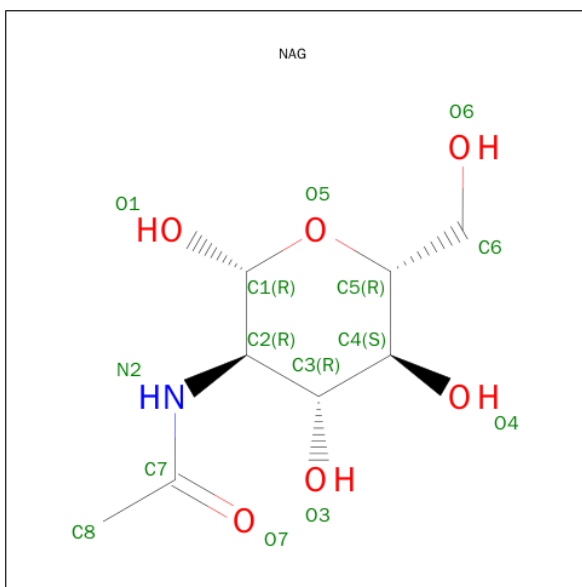
There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	42	GLY	-	EXPRESSION TAG	UNP Q68DV7
E	43	PRO	-	EXPRESSION TAG	UNP Q68DV7
E	199	ALA	-	EXPRESSION TAG	UNP Q68DV7
E	200	ALA	-	EXPRESSION TAG	UNP Q68DV7
E	201	ALA	-	EXPRESSION TAG	UNP Q68DV7
F	42	GLY	-	EXPRESSION TAG	UNP Q68DV7
F	43	PRO	-	EXPRESSION TAG	UNP Q68DV7
F	199	ALA	-	EXPRESSION TAG	UNP Q68DV7
F	200	ALA	-	EXPRESSION TAG	UNP Q68DV7
F	201	ALA	-	EXPRESSION TAG	UNP Q68DV7

- Molecule 4 is NICKEL (II) ION (three-letter code: NI) (formula: Ni).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total	Ni	0	0
			1	1		

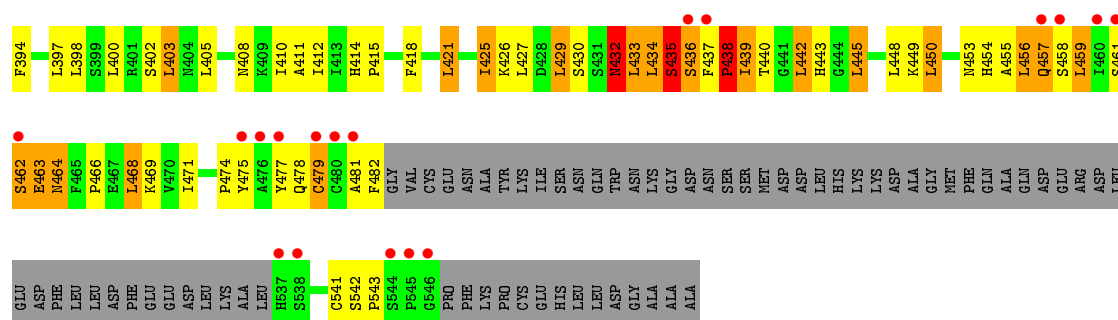
- Molecule 5 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula: C₈H₁₅NO₆).



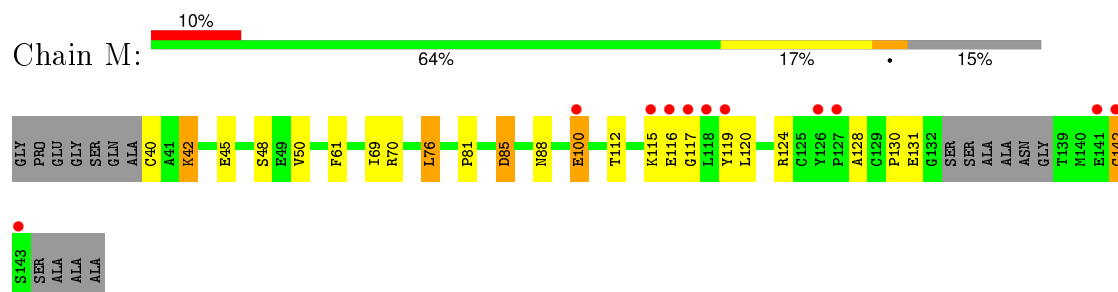
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	N	O	0	0
			14	8	1	5		
5	B	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 6 is water.

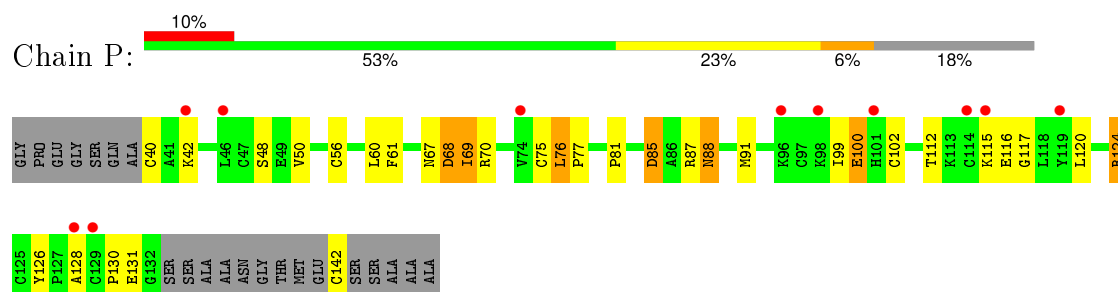
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	70	Total	O	0	0
			70	70		
6	B	85	Total	O	0	0
			85	85		
6	M	21	Total	O	0	0
			21	21		
6	P	7	Total	O	0	0
			7	7		
6	E	18	Total	O	0	0
			18	18		
6	F	4	Total	O	0	0
			4	4		



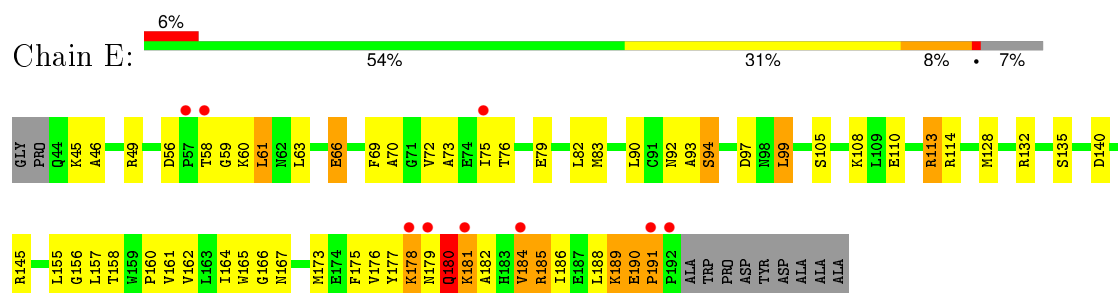
• Molecule 2: R-spondin-1



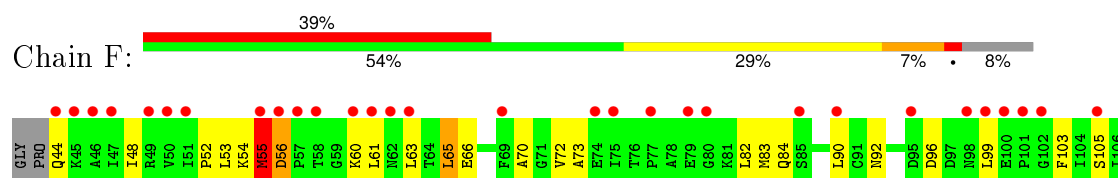
• Molecule 2: R-spondin-1

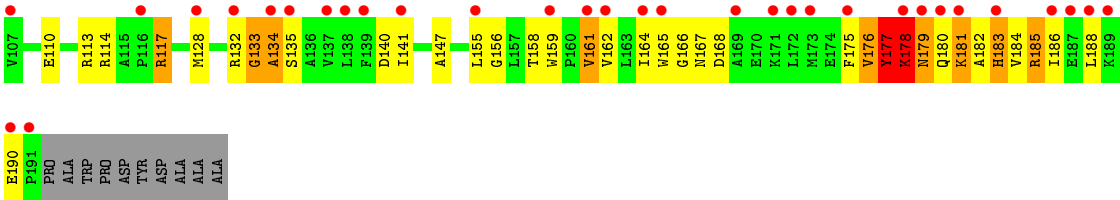


• Molecule 3: E3 ubiquitin-protein ligase RNF43



• Molecule 3: E3 ubiquitin-protein ligase RNF43





4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	104.58Å 120.97Å 181.01Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.51 – 2.50 39.51 – 2.50	Depositor EDS
% Data completeness (in resolution range)	97.1 (39.51-2.50) 97.1 (39.51-2.50)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.63 (at 2.51Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, R_{free}	0.232 , 0.276 0.232 , 0.279	Depositor DCC
R_{free} test set	3888 reflections (5.26%)	DCC
Wilson B-factor (Å ²)	43.4	Xtriage
Anisotropy	0.321	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 54.1	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	6 of 77914 reflections (0.008%)	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	11252	wwPDB-VP
Average B, all atoms (Å ²)	76.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 16.19% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: NI, NAG

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	0.61	0/3704	0.87	9/5043 (0.2%)
1	B	0.72	1/3712 (0.0%)	0.93	8/5055 (0.2%)
2	M	0.73	0/765	0.90	0/1025
2	P	0.64	0/735	0.87	0/985
3	E	0.54	0/1175	0.80	0/1593
3	F	0.48	0/1167	0.88	2/1581 (0.1%)
All	All	0.64	1/11258 (0.0%)	0.89	19/15282 (0.1%)

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
1	A	0	1
1	B	0	8
2	P	0	1
3	E	0	3
3	F	0	3
All	All	0	16

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	168	TRP	CG-CD1	5.66	1.44	1.36

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	86	LEU	CA-CB-CG	-7.22	98.69	115.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	442	LEU	CA-CB-CG	6.30	129.79	115.30
1	B	96	ARG	NE-CZ-NH1	6.22	123.41	120.30
1	A	214	VAL	CB-CA-C	-6.15	99.71	111.40
1	A	170	ASP	N-CA-CB	-6.15	99.54	110.60

There are no chirality outliers.

5 of 16 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	409	LYS	Peptide
1	B	432	ASN	Peptide
1	B	57	LEU	Peptide
1	B	59	GLU	Peptide
1	B	61	PRO	Peptide

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	3622	0	3626	129	0
1	B	3629	0	3634	146	0
2	M	750	0	719	17	0
2	P	720	0	692	26	0
3	E	1152	0	1182	43	0
3	F	1145	0	1175	47	0
4	B	1	0	0	0	0
5	A	14	0	13	0	0
5	B	14	0	13	0	0
6	A	70	0	0	8	0
6	B	85	0	0	7	0
6	E	18	0	0	4	0
6	F	4	0	0	1	0
6	M	21	0	0	2	0
6	P	7	0	0	8	0
All	All	11252	0	11054	394	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 18.

The worst 5 of 394 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:F:177:TYR:O	3:F:179:ASN:HA	1.44	1.17
1:B:450:LEU:HD12	1:B:453:ASN:HD22	0.98	1.15
1:B:459:LEU:HD21	1:B:478:GLN:CB	1.76	1.14
1:B:459:LEU:CD2	1:B:478:GLN:HB2	1.78	1.14
1:B:459:LEU:HD21	1:B:478:GLN:HB2	1.22	1.08

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	458/531 (86%)	389 (85%)	58 (13%)	11 (2%)	7	11
1	B	459/531 (86%)	392 (85%)	56 (12%)	11 (2%)	7	11
2	M	94/115 (82%)	82 (87%)	10 (11%)	2 (2%)	9	14
2	P	91/115 (79%)	80 (88%)	8 (9%)	3 (3%)	5	6
3	E	147/160 (92%)	120 (82%)	19 (13%)	8 (5%)	2	2
3	F	146/160 (91%)	119 (82%)	19 (13%)	8 (6%)	2	2
All	All	1395/1612 (86%)	1182 (85%)	170 (12%)	43 (3%)	5	7

5 of 43 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	410	ILE
1	A	430	SER
1	B	84	ASN
1	B	430	SER
1	B	438	PRO

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	418/476 (88%)	356 (85%)	62 (15%)	4	7
1	B	419/476 (88%)	354 (84%)	65 (16%)	3	6
2	M	87/95 (92%)	77 (88%)	10 (12%)	7	13
2	P	83/95 (87%)	72 (87%)	11 (13%)	5	9
3	E	124/130 (95%)	106 (86%)	18 (14%)	4	7
3	F	123/130 (95%)	98 (80%)	25 (20%)	1	2
All	All	1254/1402 (89%)	1063 (85%)	191 (15%)	3	6

5 of 191 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	256	LEU
1	B	421	LEU
3	F	99	LEU
1	B	278	VAL
1	B	330	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 39 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	141	GLN
1	B	218	HIS
2	P	88	ASN
1	B	166	HIS
1	B	199	HIS

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](#)

Of 3 ligands modelled in this entry, 1 is monoatomic - leaving 2 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	$\# Z > 2$
5	NAG	A	601	1	14,14,15	0.76	1 (7%)	15,19,21	1.94	3 (20%)
5	NAG	B	602	1	14,14,15	0.84	0	15,19,21	1.97	3 (20%)

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
5	NAG	A	601	1	-	0/6/23/26	0/1/1/1
5	NAG	B	602	1	-	0/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	601	NAG	C1-C2	2.29	1.55	1.52

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	601	NAG	C4-C3-C2	-5.16	103.20	111.23

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	B	602	NAG	C1-O5-C5	-5.02	105.87	112.25
5	A	601	NAG	O3-C3-C2	2.34	113.74	109.11
5	B	602	NAG	O5-C5-C6	2.76	113.32	107.35
5	B	602	NAG	C6-C5-C4	2.88	120.11	113.02

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	461/531 (86%)	0.39	40 (8%) 13 13	25, 66, 145, 192	0
1	B	463/531 (87%)	0.19	25 (5%) 29 33	23, 52, 120, 159	0
2	M	98/115 (85%)	0.27	11 (11%) 7 7	32, 59, 110, 123	0
2	P	94/115 (81%)	0.52	11 (11%) 6 6	31, 83, 113, 125	0
3	E	149/160 (93%)	0.34	9 (6%) 25 28	40, 79, 132, 151	0
3	F	148/160 (92%)	2.10	62 (41%) 0 0	78, 133, 169, 200	0
All	All	1413/1612 (87%)	0.50	158 (11%) 7 7	23, 68, 148, 200	0

The worst 5 of 158 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	480	CYS	12.8
3	F	191	PRO	11.6
1	A	481	ALA	9.3
1	A	539	VAL	9.1
1	B	480	CYS	8.3

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(Å ²)	Q<0.9
4	NI	B	601	1/1	0.99	0.12	-	44,44,44,44	0
5	NAG	A	601	14/15	0.90	0.31	-	66,82,88,95	0
5	NAG	B	602	14/15	0.93	0.28	-	54,60,66,67	0

6.5 Other polymers [i](#)

There are no such residues in this entry.