



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 06:39 PM GMT

PDB ID : 4MBS
Title : Crystal Structure of the CCR5 Chemokine Receptor
Authors : Tan, Q.; Zhu, Y.; Han, G.W.; Li, J.; Fenalti, G.; Liu, H.; Cherezov, V.;
Stevens, R.C.; GPCR Network (GPCR); Zhao, Q.; Wu, B.
Deposited on : 2013-08-19
Resolution : 2.71 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

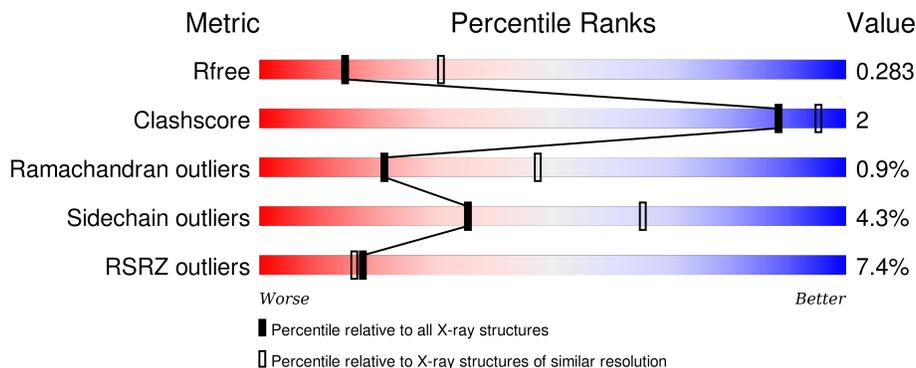
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.71 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	2439 (2.74-2.70)
Clashscore	102246	2771 (2.74-2.70)
Ramachandran outliers	100387	2726 (2.74-2.70)
Sidechain outliers	100360	2727 (2.74-2.70)
RSRZ outliers	91569	2443 (2.74-2.70)

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	414	
1	B	414	

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
4	OLC	A	1104	-	-	-	X
4	OLC	A	1105	-	-	-	X
4	OLC	A	1106	-	-	-	X
4	OLC	A	1107	-	-	-	X
4	OLC	A	1108	-	-	-	X
4	OLC	B	1104	-	-	-	X
4	OLC	B	1105	-	-	-	X
4	OLC	B	1106	-	-	-	X

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 5764 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 Chimera protein of C-C chemokine receptor type 5 and Rubredoxin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	346	2742	1828	426	469	19	0	0	0
1	B	346	2744	1828	427	470	19	0	0	0

There are 32 discrepancies between the modelled and reference sequences:

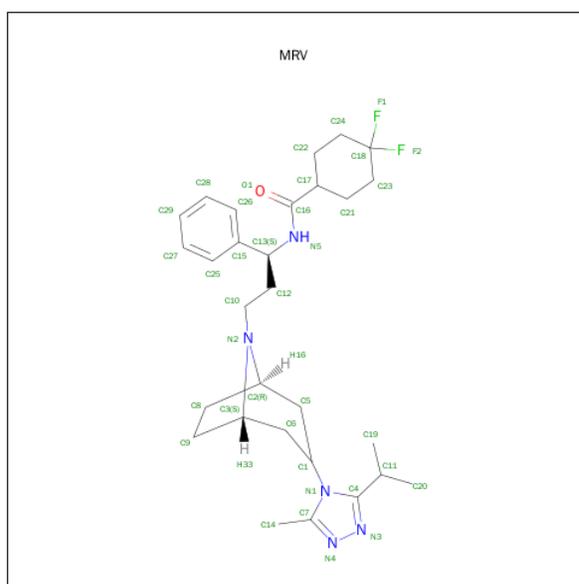
Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	EXPRESSION TAG	UNP P51681
A	0	ALA	-	EXPRESSION TAG	UNP P51681
A	1	PRO	-	EXPRESSION TAG	UNP P51681
A	58	TYR	CYS	ENGINEERED MUTATION	UNP P51681
A	163	ASN	GLY	ENGINEERED MUTATION	UNP P51681
A	233	ASP	ALA	ENGINEERED MUTATION	UNP P51681
A	303	GLU	LYS	ENGINEERED MUTATION	UNP P51681
A	353	GLY	-	EXPRESSION TAG	UNP P51681
A	354	ARG	-	EXPRESSION TAG	UNP P51681
A	355	PRO	-	EXPRESSION TAG	UNP P51681
A	356	LEU	-	EXPRESSION TAG	UNP P51681
A	357	GLU	-	EXPRESSION TAG	UNP P51681
A	358	VAL	-	EXPRESSION TAG	UNP P51681
A	359	LEU	-	EXPRESSION TAG	UNP P51681
A	360	PHE	-	EXPRESSION TAG	UNP P51681
A	361	GLN	-	EXPRESSION TAG	UNP P51681
B	-1	GLY	-	EXPRESSION TAG	UNP P51681
B	0	ALA	-	EXPRESSION TAG	UNP P51681
B	1	PRO	-	EXPRESSION TAG	UNP P51681
B	58	TYR	CYS	ENGINEERED MUTATION	UNP P51681
B	163	ASN	GLY	ENGINEERED MUTATION	UNP P51681
B	233	ASP	ALA	ENGINEERED MUTATION	UNP P51681
B	303	GLU	LYS	ENGINEERED MUTATION	UNP P51681
B	353	GLY	-	EXPRESSION TAG	UNP P51681

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	354	ARG	-	EXPRESSION TAG	UNP P51681
B	355	PRO	-	EXPRESSION TAG	UNP P51681
B	356	LEU	-	EXPRESSION TAG	UNP P51681
B	357	GLU	-	EXPRESSION TAG	UNP P51681
B	358	VAL	-	EXPRESSION TAG	UNP P51681
B	359	LEU	-	EXPRESSION TAG	UNP P51681
B	360	PHE	-	EXPRESSION TAG	UNP P51681
B	361	GLN	-	EXPRESSION TAG	UNP P51681

- Molecule 2 is 4,4-DIFLUORO-N-[(1S)-3-{(3-EXO)-3-[3-METHYL-5-(PROPAN-2-YL)-4H-1,2,4-TRIAZOL-4-YL]-8-AZABICYCLO[3.2.1]OCT-8-YL}-1-PHENYLPROPYL]CYCLOHEXANECARBOXAMIDE (three-letter code: MRV) (formula: C₂₉H₄₁F₂N₅O).

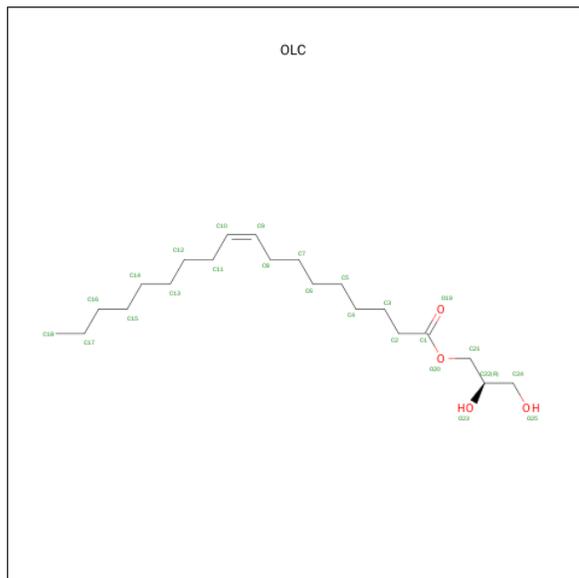


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
2	A	1	Total	C	F	N	O	0	0
			37	29	2	5	1		
2	B	1	Total	C	F	N	O	0	0
			37	29	2	5	1		

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Zn	0	0
			1	1		
3	A	1	Total	Zn	0	0
			1	1		

- Molecule 4 is (2R)-2,3-DIHYDROXYPROPYL (9Z)-OCTADEC-9-ENOATE (three-letter code: OLC) (formula: C₂₁H₄₀O₄).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			16	12	4		
4	A	1	Total	C	O	0	0
			13	9	4		
4	A	1	Total	C	O	0	0
			15	11	4		
4	A	1	Total	C	O	0	0
			25	21	4		
4	A	1	Total	C	O	0	0
			19	15	4		
4	A	1	Total	C	O	0	0
			12	8	4		
4	B	1	Total	C	O	0	0
			17	13	4		
4	B	1	Total	C	O	0	0
			12	8	4		
4	B	1	Total	C	O	0	0
			13	9	4		
4	B	1	Total	C	O	0	0
			12	8	4		

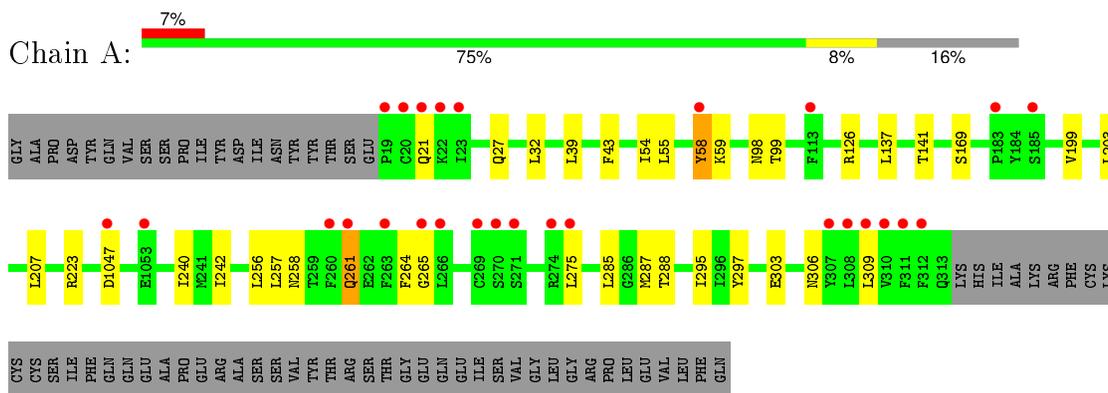
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	22	Total 22	O 22	0	0
5	B	26	Total 26	O 26	0	0

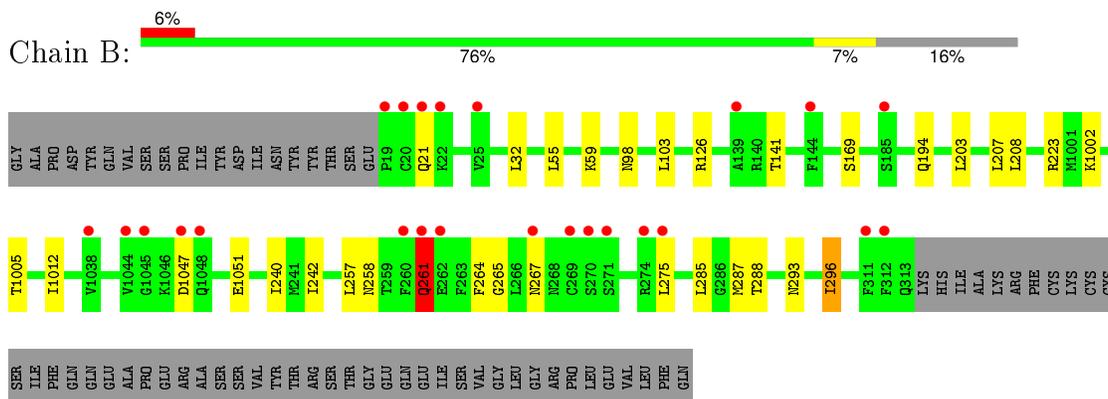
3 Residue-property plots [i](#)

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: Chimera protein of C-C chemokine receptor type 5 and Rubredoxin



- Molecule 1: Chimera protein of C-C chemokine receptor type 5 and Rubredoxin



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	72.92Å 103.52Å 137.47Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	36.33 – 2.71 35.97 – 2.71	Depositor EDS
% Data completeness (in resolution range)	95.5 (36.33-2.71) 95.6 (35.97-2.71)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.50 (at 2.72Å)	Xtrriage
Refinement program	BUSTER 2.10.0	Depositor
R, R_{free}	0.216 , 0.263 0.224 , 0.283	Depositor DCC
R_{free} test set	1416 reflections (5.40%)	DCC
Wilson B-factor (Å ²)	61.2	Xtrriage
Anisotropy	0.392	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 68.5	EDS
Estimated twinning fraction	No twinning to report.	Xtrriage
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Outliers	2 of 27630 reflections (0.007%)	Xtrriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	5764	wwPDB-VP
Average B, all atoms (Å ²)	72.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 38.92 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 3.4299e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, OLC, MRV

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	0.50	0/2820	0.60	0/3850
1	B	0.49	0/2821	0.61	0/3848
All	All	0.50	0/5641	0.60	0/7698

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	2742	0	2679	15	0
1	B	2744	0	2686	13	0
2	A	37	0	41	0	0
2	B	37	0	41	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	100	0	133	1	0
4	B	54	0	63	0	0
5	A	22	0	0	0	0
5	B	26	0	0	0	0
All	All	5764	0	5643	28	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 28 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:B:240:ILE:HG12	1:B:296:ILE:HD11	1.61	0.81
1:B:257:LEU:HA	1:B:261:GLN:HB2	1.71	0.73
1:A:257:LEU:HA	1:A:261:GLN:HB2	1.73	0.70
1:B:261:GLN:HA	1:B:264:PHE:HB2	1.72	0.70
1:A:261:GLN:HA	1:A:264:PHE:HB2	1.72	0.69

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	344/414 (83%)	332 (96%)	9 (3%)	3 (1%)	21	47
1	B	344/414 (83%)	332 (96%)	9 (3%)	3 (1%)	21	47
All	All	688/828 (83%)	664 (96%)	18 (3%)	6 (1%)	21	47

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	261	GLN
1	B	261	GLN
1	A	21	GLN
1	B	21	GLN
1	B	265	GLY

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	289/371 (78%)	277 (96%)	12 (4%)	36	66
1	B	289/371 (78%)	276 (96%)	13 (4%)	34	63
All	All	578/742 (78%)	553 (96%)	25 (4%)	35	65

5 of 25 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	303	GLU
1	B	103	LEU
1	B	287	MET
1	B	32	LEU
1	B	126	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	102	GLN
1	A	163	ASN
1	A	194	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](#)

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 14 ligands modelled in this entry, 2 are monoatomic - leaving 12 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
2	MRV	A	1101	-	34,41,41	2.19	4 (11%)	43,60,60	1.77	5 (11%)
4	OLC	A	1103	-	15,15,24	1.20	1 (6%)	16,16,25	1.05	1 (6%)
4	OLC	A	1104	-	12,12,24	1.31	1 (8%)	13,13,25	1.21	1 (7%)
4	OLC	A	1105	-	14,14,24	1.21	1 (7%)	15,15,25	0.96	1 (6%)
4	OLC	A	1106	-	24,24,24	0.94	1 (4%)	25,25,25	0.86	1 (4%)
4	OLC	A	1107	-	18,18,24	1.11	1 (5%)	19,19,25	1.08	2 (10%)
4	OLC	A	1108	-	11,11,24	1.38	1 (9%)	12,12,25	1.09	2 (16%)
2	MRV	B	1101	-	34,41,41	2.32	6 (17%)	43,60,60	1.73	6 (13%)
4	OLC	B	1103	-	16,16,24	1.16	1 (6%)	17,17,25	1.01	1 (5%)
4	OLC	B	1104	-	11,11,24	1.32	1 (9%)	12,12,25	1.07	2 (16%)
4	OLC	B	1105	-	12,12,24	1.36	1 (8%)	13,13,25	1.06	2 (15%)
4	OLC	B	1106	-	11,11,24	1.39	1 (9%)	12,12,25	1.09	1 (8%)

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
2	MRV	A	1101	-	-	0/21/58/58	0/3/5/5
4	OLC	A	1103	-	-	0/15/15/24	0/0/0/0
4	OLC	A	1104	-	-	0/12/12/24	0/0/0/0
4	OLC	A	1105	-	-	0/14/14/24	0/0/0/0
4	OLC	A	1106	-	-	0/24/24/24	0/0/0/0
4	OLC	A	1107	-	-	0/18/18/24	0/0/0/0
4	OLC	A	1108	-	-	0/11/11/24	0/0/0/0
2	MRV	B	1101	-	-	0/21/58/58	0/3/5/5
4	OLC	B	1103	-	-	0/16/16/24	0/0/0/0
4	OLC	B	1104	-	-	0/11/11/24	0/0/0/0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	OLC	B	1105	-	-	0/12/12/24	0/0/0/0
4	OLC	B	1106	-	-	0/11/11/24	0/0/0/0

The worst 5 of 20 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1101	MRV	C15-C13	-8.27	1.40	1.52
2	B	1101	MRV	C15-C13	-7.94	1.40	1.52
2	A	1101	MRV	N3-N4	-7.64	1.22	1.37
2	B	1101	MRV	N3-N4	-6.91	1.23	1.37
2	B	1101	MRV	F2-C18	-4.72	1.32	1.38

The worst 5 of 25 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	1101	MRV	C6-C3-C9	-4.09	108.80	112.61
2	A	1101	MRV	C6-C3-C9	-3.13	109.69	112.61
4	A	1107	OLC	O20-C1-O19	-2.39	117.32	123.49
4	B	1104	OLC	O20-C1-O19	-2.00	118.32	123.49
4	A	1108	OLC	C21-O20-C1	2.06	122.61	116.85

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
4	A	1106	OLC	1	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	346/414 (83%)	0.43	27 (7%) 16 14	47, 66, 131, 167	0
1	B	346/414 (83%)	0.30	24 (6%) 20 19	44, 67, 117, 154	0
All	All	692/828 (83%)	0.37	51 (7%) 17 16	44, 66, 130, 167	0

The worst 5 of 51 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	19	PRO	9.5
1	A	19	PRO	7.3
1	A	21	GLN	7.2
1	A	20	CYS	7.0
1	A	261	GLN	6.1

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
4	OLC	A	1108	12/25	0.80	0.41	13.90	83,89,91,92	0
4	OLC	A	1104	13/25	0.68	0.46	7.90	87,101,106,108	0
4	OLC	B	1106	12/25	0.73	0.37	6.96	90,92,94,94	0
4	OLC	A	1107	19/25	0.73	0.36	5.89	75,91,96,97	0
4	OLC	A	1106	25/25	0.53	0.37	4.71	80,103,117,118	0
4	OLC	A	1105	15/25	0.72	0.34	3.85	92,99,107,107	0
4	OLC	B	1104	12/25	0.84	0.32	3.38	98,105,112,113	0
4	OLC	B	1105	13/25	0.72	0.42	1.61	107,113,117,118	0
4	OLC	A	1103	16/25	0.85	0.25	1.36	97,99,100,100	0
4	OLC	B	1103	17/25	0.78	0.31	1.24	98,101,103,103	0
2	MRV	B	1101	37/37	0.95	0.16	-0.58	47,54,60,61	0
2	MRV	A	1101	37/37	0.95	0.16	-0.63	44,52,59,61	0
3	ZN	B	1102	1/1	0.99	0.06	-1.96	72,72,72,72	0
3	ZN	A	1102	1/1	0.99	0.06	-2.56	61,61,61,61	0

6.5 Other polymers [i](#)

There are no such residues in this entry.