WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Jun 19, 2024 – 10:36 AM EDT

PDB ID	:	4NQD
Title	:	Crystal structure of TCR-MR1 ternary complex and non-covalently bound 5-(2-oxopropylideneamino)-6-D-ribitylaminouracil
Authors	:	Birkinshaw, R.W.; Rossjohn, J.
Deposited on	:	2013-11-25
Resolution	:	2.20 Å(reported)

The following versions of software and data (see references (i)) were used in the production of this report:

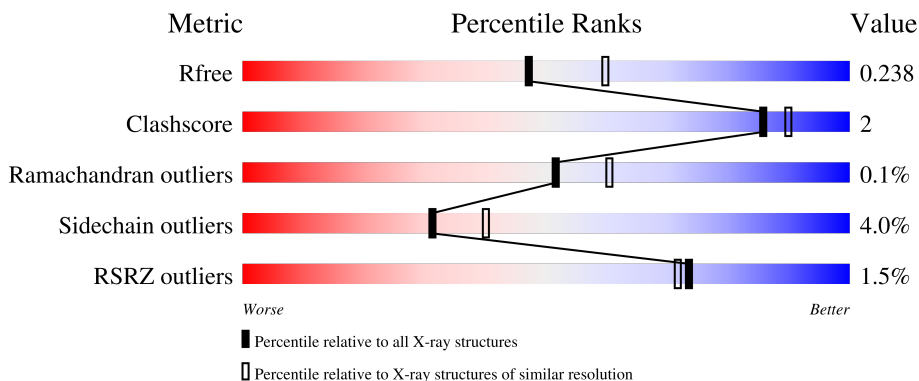|		
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 2.20 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.
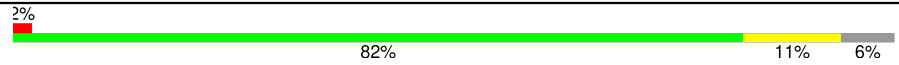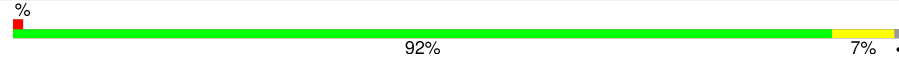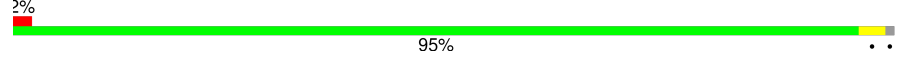
Metric	Percentile Ranks	Value
Rfree		0.238
Clashscore		2
Ramachandran outliers		0.1%
Sidechain outliers		4.0%
RSRZ outliers		1.5%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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 of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. 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 <=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	271	87% 8% •
1	C	271	82% 10% 8%
2	B	99	90% 6% ••
2	F	99	8% 91% 5% ••
3	D	203	89% 9% •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	G	203	2% 82% 11% 6%
4	E	245	% 92% 7%
4	H	245	2% 95%

2　Entry composition ⓘ

There are 7 unique types of molecules in this entry. The entry contains 13111 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 Major histocompatibility complex class I-related gene protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	260	Total 2120	C 1359	N 365	O 386	S 10	0	1	0
1	C	248	Total 2043	C 1310	N 350	O 372	S 11	0	2	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	EXPRESSION TAG	UNP Q95460
A	261	SER	CYS	ENGINEERED MUTATION	UNP Q95460
C	0	MET	-	EXPRESSION TAG	UNP Q95460
C	261	SER	CYS	ENGINEERED MUTATION	UNP Q95460

- Molecule 2 is a protein called Beta-2-microglobulin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	96	Total 785	C 502	N 131	O 150	S 2	0	0	0
2	F	97	Total 769	C 494	N 127	O 146	S 2	0	0	0

- Molecule 3 is a protein called TCR alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	200	Total 1548	C 978	N 248	O 313	S 9	0	0	0
3	G	190	Total 1460	C 929	N 230	O 292	S 9	0	1	0

- Molecule 4 is a protein called TCR beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	243	Total 1907	C 1202	N 333	O 363	S 9	0	2	0
4	H	242	Total 1882	C 1186	N 325	O 362	S 9	0	0	0

- Molecule 5 is a ligand with the chemical component id 2LJ but its atom names do not match the existing wwPDB Chemical Component Dictionary definition for 2LJ. ERROR THIS SHOULD NOT HAPPEN FOLLOWING ANNOTATION.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total 23	C 12	N 4	O 7	0	0
5	C	1	Total 23	C 12	N 4	O 7	0	0

- Molecule 6 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).

GOL

C3

OH O3

C1 C2

HO OH

O1 O2
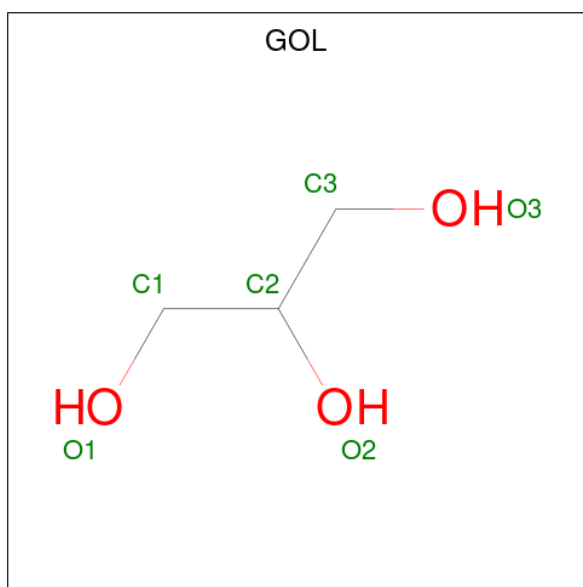
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	D	1	Total 6	C 3	O 3	0	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	113	Total 113	O 113	0	0
7	B	38	Total 38	O 38	0	0

Continued from previous page...

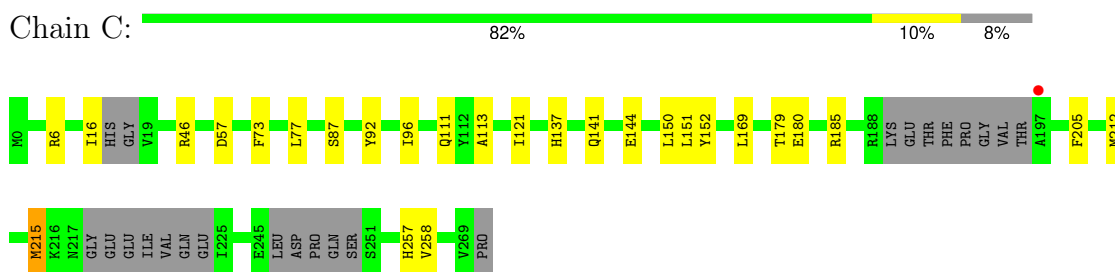| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	C	83	Total 83	O 83	0	0
7	D	84	Total 84	O 84	0	0
7	E	99	Total 99	O 99	0	0
7	F	23	Total 23	O 23	0	0
7	G	57	Total 57	O 57	0	0
7	H	48	Total 48	O 48	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.
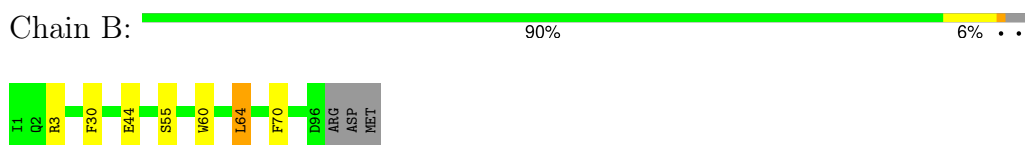
• Molecule 1: Major histocompatibility complex class I-related gene protein
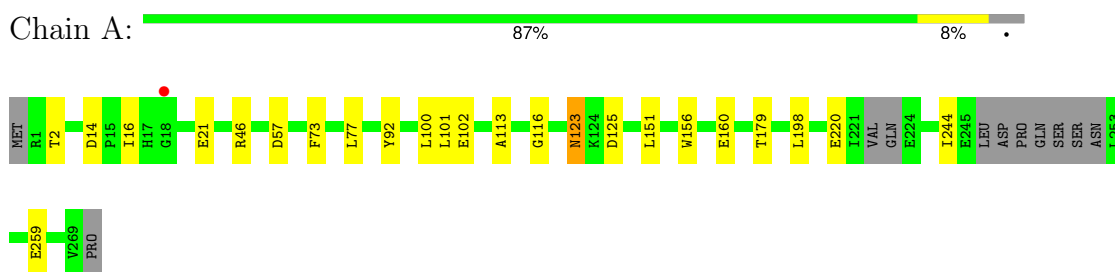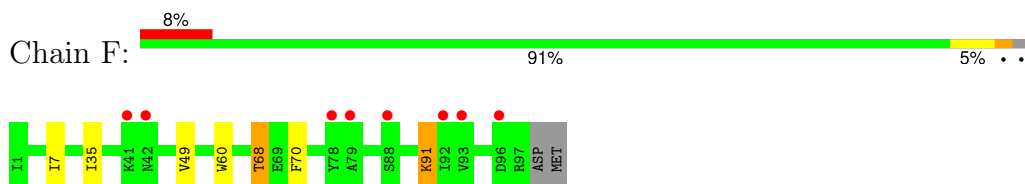
Chain A:

• Molecule 1: Major histocompatibility complex class I-related gene protein

Chain C:

• Molecule 2: Beta-2-microglobulin

Chain B:

• Molecule 2: Beta-2-microglobulin

Chain F:

• Molecule 3: TCR alpha chain

Chain D:

G1 Q25 T26 S27 G28 L51 F60 K68 L78 A89 S93 L97 I107 D148 S149 D150 V151 Y152 R162 A173 K177 S178 D179 S199 P200 GLU SER SER

• Molecule 3: TCR alpha chain

Chain G:

2%
82%   11%   6%

G1 Q25 T26 S27 G28 G31 Y48 L51 L75 L78 Q79 M80 S85 A89 S93 L97 I107 L121 R122 D123 SER LYS SER SER ASP LYS S130 Q145 S146 LYS ASP SER D150 D155 K156 C157 K167 S178 C182 A183 N187

F197 P198 S199 PRO GLU SER SER

• Molecule 4: TCR beta chain

Chain E:

%
92%   7%   •

ASN A2 V12 L21 H29 N65 T86 F105 T113 V128 P131 L144 Q181 P182 A183 L184 N185 R194 H208 F209 R210 C211 Q212 S237 A238 W241 A244 ASP

• Molecule 4: TCR beta chain

Chain H:

2%
95%   • •

ASN ALA G3 E79 T97 F129 A148 H155 L184 A190 R194 A199 Y216 N221 T225 Q226 D227 A244 ASP
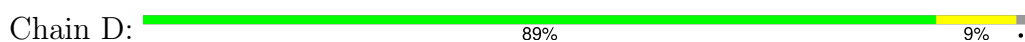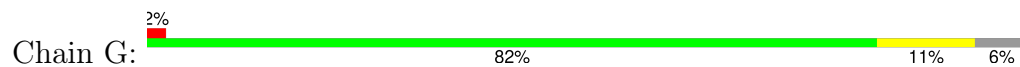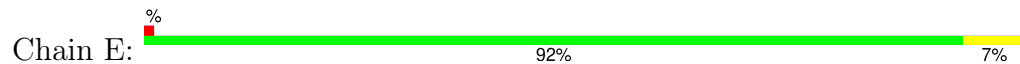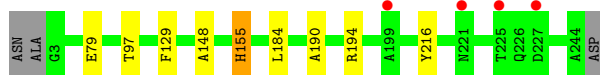
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	215.58Å 68.87Å 142.98Å 90.00° 104.21° 90.00°	Depositor
Resolution (Å)	49.51 – 2.20 49.51 – 2.20	Depositor EDS
% Data completeness (in resolution range)	97.7 (49.51-2.20) 97.6 (49.51-2.20)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.52 (at 2.20Å)	Xtriage
Refinement program	BUSTER 2.10.0	Depositor
R, R_{free}	0.181 , 0.227 0.191 , 0.238	Depositor DCC
R_{free} test set	5109 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å^2)	32.0	Xtriage
Anisotropy	0.465	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 48.0	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	13111	wwPDB-VP
Average B, all atoms (Å^2)	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.71% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 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: GOL, 2LJ

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	0.52	0/2187	0.65	0/2975
1	C	0.53	0/2109	0.66	0/2865
2	B	0.50	0/808	0.64	0/1100
2	F	0.48	0/792	0.64	0/1082
3	D	0.52	0/1583	0.70	0/2148
3	G	0.50	0/1495	0.68	0/2030
4	E	0.51	0/1966	0.68	0/2679
4	H	0.47	0/1933	0.67	0/2635
All	All	0.51	0/12873	0.67	0/17514

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	2120	0	1989	10	0
1	C	2043	0	1919	13	0
2	B	785	0	731	3	0
2	F	769	0	699	4	0
3	D	1548	0	1456	8	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	G	1460	0	1360	8	0
4	E	1907	0	1788	10	0
4	H	1882	0	1754	3	0
5	A	23	0	18	1	0
5	C	23	0	18	2	0
6	D	6	0	8	0	0
7	A	113	0	0	0	0
7	B	38	0	0	0	0
7	C	83	0	0	1	0
7	D	84	0	0	0	0
7	E	99	0	0	0	0
7	F	23	0	0	0	0
7	G	57	0	0	0	0
7	H	48	0	0	0	0
All	All	13111	0	11740	54	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 54 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:E:210:ARG:HE	4:E:212:GLN:HE21	1.16	0.93
3:D:1:GLY:HA2	3:D:27:SER:H	1.36	0.91
3:G:1:GLY:HA2	3:G:27:SER:H	1.53	0.73
4:E:210:ARG:HE	4:E:212:GLN:NE2	1.87	0.70
4:E:210:ARG:HH21	4:E:212:GLN:HE22	1.42	0.65

There are no symmetry-related clashes.

5.3　Torsion angles ⓘ

5.3.1　Protein backbone ⓘ

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	255/271 (94%)	248 (97%)	6 (2%)	1 (0%)	34	37
1	C	240/271 (89%)	234 (98%)	6 (2%)	0	100	100
2	B	94/99 (95%)	94 (100%)	0	0	100	100
2	F	95/99 (96%)	92 (97%)	3 (3%)	0	100	100
3	D	198/203 (98%)	195 (98%)	3 (2%)	0	100	100
3	G	185/203 (91%)	180 (97%)	5 (3%)	0	100	100
4	E	243/245 (99%)	241 (99%)	2 (1%)	0	100	100
4	H	240/245 (98%)	233 (97%)	7 (3%)	0	100	100
All	All	1550/1636 (95%)	1517 (98%)	32 (2%)	1 (0%)	51	60

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	220	GLU

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	221/241 (92%)	211 (96%)	10 (4%)	27	34
1	C	216/241 (90%)	207 (96%)	9 (4%)	30	38
2	B	87/94 (93%)	83 (95%)	4 (5%)	27	34
2	F	82/94 (87%)	78 (95%)	4 (5%)	25	31
3	D	172/180 (96%)	165 (96%)	7 (4%)	30	39
3	G	159/180 (88%)	149 (94%)	10 (6%)	18	20
4	E	204/211 (97%)	199 (98%)	5 (2%)	47	60
4	H	200/211 (95%)	195 (98%)	5 (2%)	47	60
All	All	1341/1452 (92%)	1287 (96%)	54 (4%)	31	40

5 of 54 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	D	162	ARG
2	F	35	ILE
4	H	79	GLU
3	D	199	SER
4	E	181	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	264	HIS
3	D	120	GLN
4	H	155	HIS
4	E	212	GLN
1	C	148	HIS

5.3.3   RNA ⓘ

There are no RNA molecules in this entry.

5.4   Non-standard residues in protein, DNA, RNA chains ⓘ

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

5.5   Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6   Ligand geometry ⓘ

Of 3 ligands modelled in this entry, 2 could not be matched to an existing wwPDB Chemical Component Dictionary definition at this stage - leaving 1 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	GOL	D	301	-	5,5,5	0.09	0	5,5,5	0.17	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	GOL	D	301	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	D	301	GOL	O1-C1-C2-O2
6	D	301	GOL	O1-C1-C2-C3

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers ⓘ

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 ⓘ

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, 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(Å²)	Q<0.9
1	A	260/271 (95%)	-0.36	1 (0%)	92	91	19, 31, 56, 79	0
1	C	248/271 (91%)	-0.36	1 (0%)	92	91	21, 34, 59, 75	0
2	B	96/99 (96%)	-0.35	0	100	100	24, 35, 61, 74	0
2	F	97/99 (97%)	0.36	8 (8%)	11	10	27, 50, 72, 84	0
3	D	200/203 (98%)	-0.45	1 (0%)	91	90	19, 32, 58, 70	0
3	G	190/203 (93%)	0.06	5 (2%)	56	53	21, 39, 71, 94	0
4	E	243/245 (99%)	-0.38	3 (1%)	79	77	22, 32, 49, 75	1 (0%)
4	H	242/245 (98%)	-0.20	4 (1%)	70	68	23, 40, 59, 82	0
All	All	1576/1636 (96%)	-0.26	23 (1%)	73	72	19, 35, 64, 94	1 (0%)

The worst 5 of 23 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	G	178	SER	3.4
2	F	79	ALA	3.3
4	H	225	THR	3.1
4	E	183	ALA	3.1
2	F	78	TYR	2.9

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

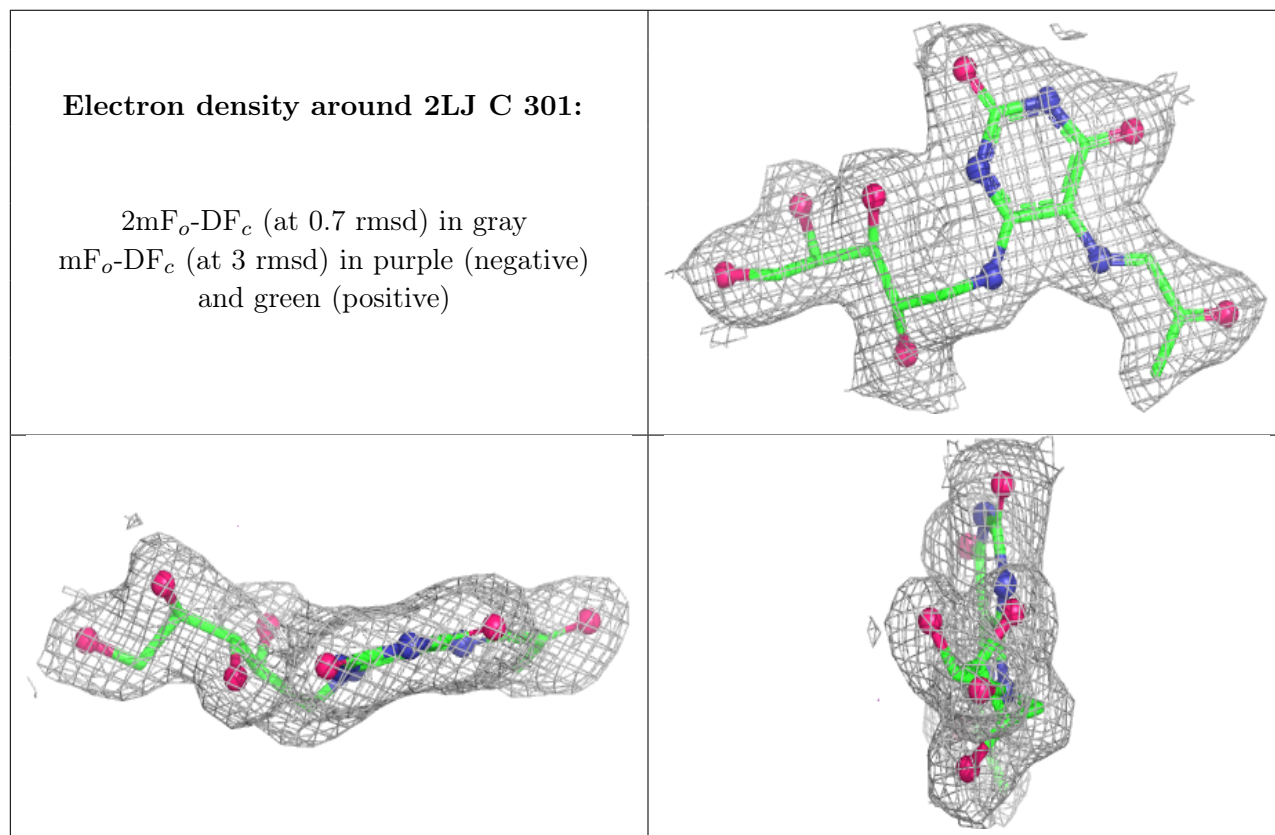There are no non-standard protein/DNA/RNA residues in this entry.

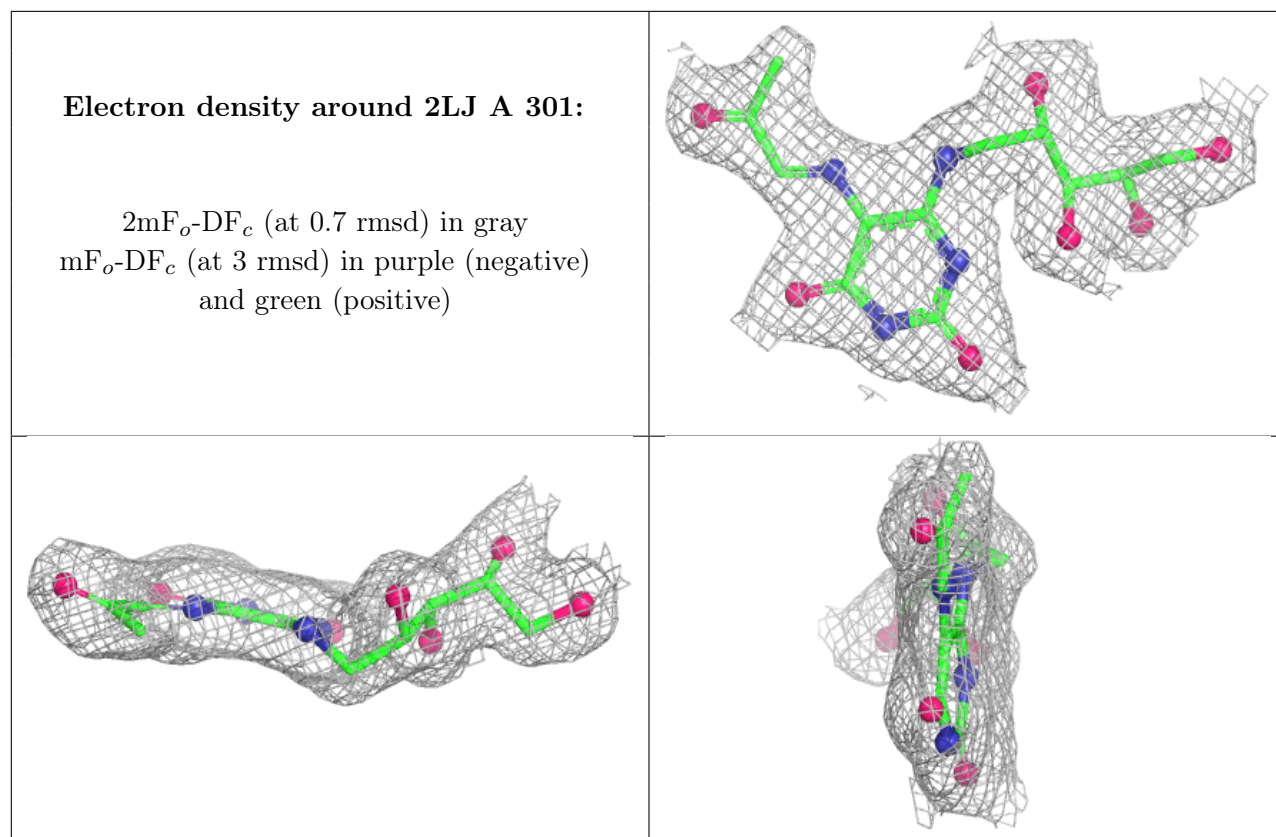## 6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

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, 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	B-factors(Å^2)	Q<0.9
6	GOL	D	301	6/6	0.93	0.14	52,55,56,56	0
5	2LJ	C	301	23/22	0.98	0.10	21,25,28,30	0
5	2LJ	A	301	23/22	0.98	0.14	23,27,30,30	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around 2LJ C 301:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative) and green (positive)

Electron density around 2LJ A 301:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)

6.5 Other polymers ⓘ

There are no such residues in this entry.