wwPDB X-ray Structure Validation Summary Report (i)

Jun 18, 2024 – 10:31 AM EDT

PDB ID	:	4LCC
Title	:	Crystal structure of a human MAIT TCR in complex with a bacterial antigen bound to humanized bovine MR1
Authors	:	Lopez-Sagaseta, J.; Adams, E.J.
Deposited on	:	2013-06-21
Resolution	:	3.26 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

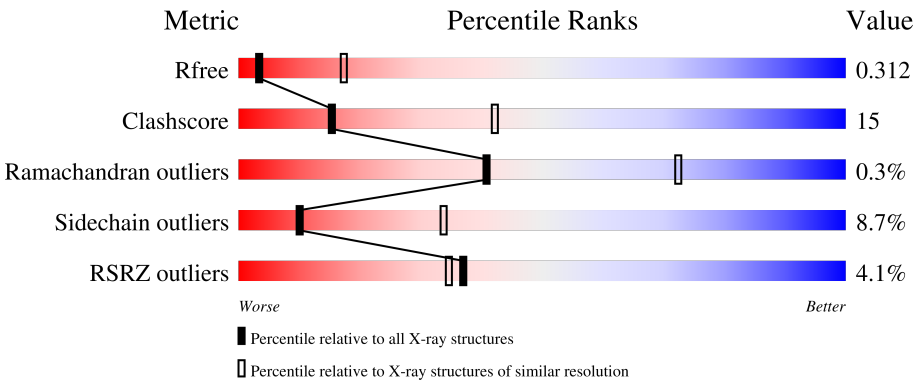
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.26 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}	130704	1191 (3.30-3.22)
Clashscore	141614	1251 (3.30-3.22)
Ramachandran outliers	138981	1229 (3.30-3.22)
Sidechain outliers	138945	1228 (3.30-3.22)
RSRZ outliers	127900	1154 (3.30-3.22)

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	C	392	4%, 60%, 22%, •, 13%
2	A	208	4%, 55%, 19%, •, 25%
3	B	253	2%, 59%, 30%, •, 6%

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	1XL	C	401	-	-	-	X

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 5827 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 Beta-2-microglobulin, MHC class I-related protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	C	341	2763	1778	470	501	14	0	0	0

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	185	MET	ALA	ENGINEERED MUTATION	UNP C1ITJ8
C	260	GLN	ARG	ENGINEERED MUTATION	UNP C1ITJ8
C	264	LEU	GLN	ENGINEERED MUTATION	UNP C1ITJ8
C	391	GLY	-	EXPRESSION TAG	UNP C1ITJ8
C	392	GLY	-	EXPRESSION TAG	UNP C1ITJ8
C	99	GLY	-	LINKER	UNP P01888
C	100	GLY	-	LINKER	UNP P01888
C	101	GLY	-	LINKER	UNP P01888
C	102	GLY	-	LINKER	UNP P01888
C	103	SER	-	LINKER	UNP P01888
C	104	GLY	-	LINKER	UNP P01888
C	105	GLY	-	LINKER	UNP P01888
C	106	SER	-	LINKER	UNP P01888
C	107	GLY	-	LINKER	UNP P01888
C	108	SER	-	LINKER	UNP P01888
C	109	GLY	-	LINKER	UNP P01888
C	110	GLY	-	LINKER	UNP P01888
C	111	GLY	-	LINKER	UNP P01888
C	112	GLY	-	LINKER	UNP P01888
C	113	SER	-	LINKER	UNP P01888

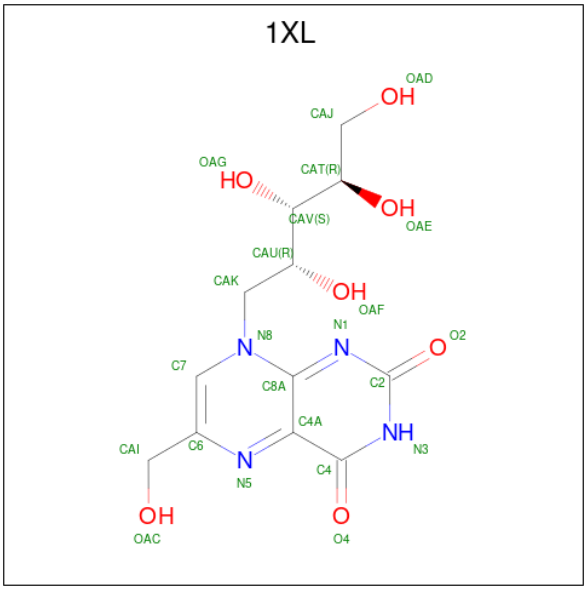
- Molecule 2 is a protein called Human MAIT TCR alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	A	157	1196	759	192	240	5	0	0	0

- Molecule 3 is a protein called Human MAIT TCR beta chain.

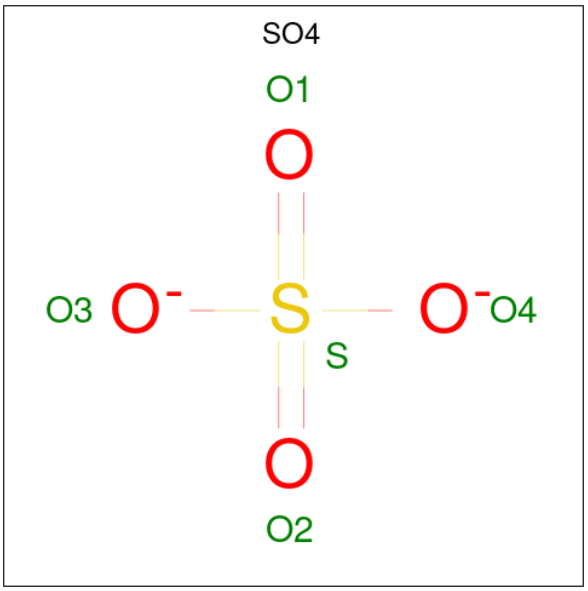
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	B	238	1824	1154	307	354	9	0	0	0

- Molecule 4 is 1-deoxy-1-[6-(hydroxymethyl)-2,4-dioxo-3,4-dihydropteridin-8(2H)-yl]-D-arabinitol (three-letter code: 1XL) (formula: $C_{12}H_{16}N_4O_7$).



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

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



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

- Molecule 6 is water.

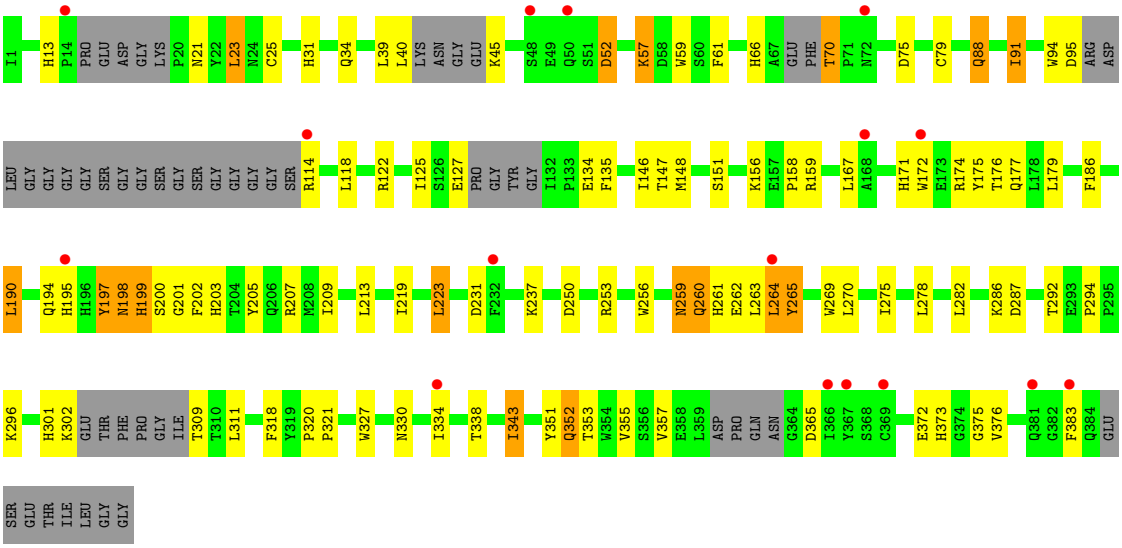
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total O 1 1	0	0

3 Residue-property plots (i)

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.

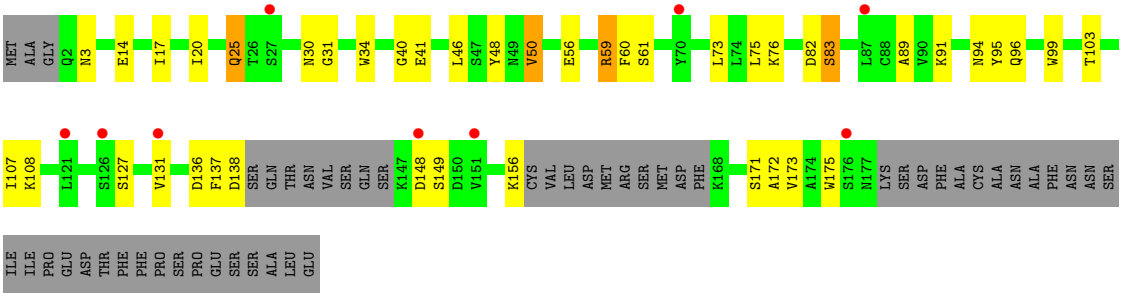
• Molecule 1: Beta-2-microglobulin, MHC class I-related protein

Chain C: 4% 60% 22% • 13%



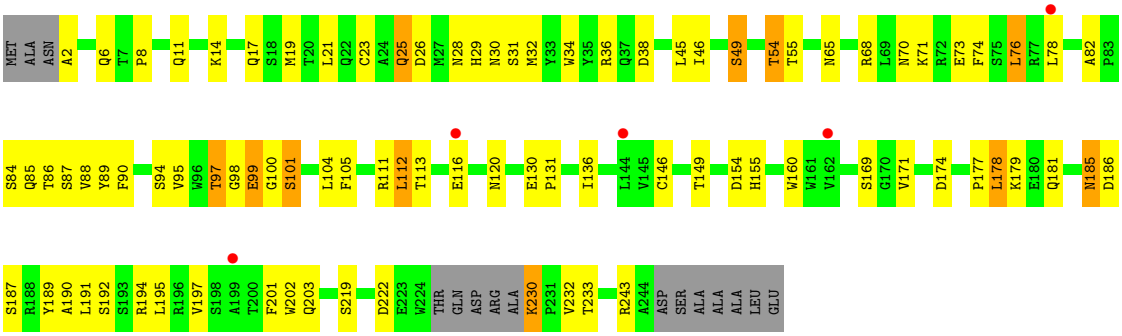
• Molecule 2: Human MAIT TCR alpha chain

Chain A: 4% 55% 19% • 25%



• Molecule 3: Human MAIT TCR beta chain

Chain B: 2% 59% 30% • 6%


MET ALA ASN A2 Q6 T7 P8 Q11 K14 Q17 S18 M19 T20 L21 Q22 C23 A24 Q25 D26 M27 N28 H29 N30 S31 M32 Y33 W34 Y35 R36 Q37 D38 L45 I46 S49 T54 T55 N65 R68 L69 N70 K71 R72 E73 F74 S75 L76 R77 L78 A82 P83
S84 Q85 T86 S87 V88 Y89 F90 S94 V95 W96 T97 G98 E99 G100 S101 L104 F105 R111 L112 T113 E116 N120 E130 P131 I136 L144 V145 C146 T149 D154 H155 W160 W161 V162 S169 G170 V171 D174 P177 L178 K179 E180 Q181 N185 D186
S187 R188 Y189 A190 L191 S192 S193 R194 L195 R196 V197 S198 A199 T200 F201 W202 Q203 S219 D222 E223 W224 THR GLN ASP ARG ALA K230 P231 V232 T233 R243 A244 ASP SER ALA ALA ALA LEU GLU

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	85.90Å 88.58Å 155.50Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.32 – 3.26 48.32 – 3.26	Depositor EDS
% Data completeness (in resolution range)	97.0 (48.32-3.26) 97.0 (48.32-3.26)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.89 (at 3.25Å)	Xtriage
Refinement program	PHENIX 1.8.2_1309	Depositor
R, R_{free}	0.257 , 0.312 0.260 , 0.312	Depositor DCC
R_{free} test set	938 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å^2)	40.0	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 65.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.008 for k,h,-l	Xtriage
F_o,F_c correlation	0.80	EDS
Total number of atoms	5827	wwPDB-VP
Average B, all atoms (Å^2)	88.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.73% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: 1XL, SO4

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	C	0.30	0/2845	0.50	1/3865 (0.0%)
2	A	0.36	0/1221	0.53	0/1660
3	B	0.33	0/1874	0.52	0/2559
All	All	0.32	0/5940	0.52	1/8084 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	198	ASN	N-CA-C	5.75	126.53	111.00

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	C	2763	0	2551	78	0
2	A	1196	0	1098	32	0
3	B	1824	0	1676	64	0
4	C	23	0	16	4	0
5	A	5	0	0	1	0
5	B	10	0	0	0	0
5	C	5	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	A	1	0	0	0	0
All	All	5827	0	5341	166	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 15.

The worst 5 of 166 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:C:127:GLU:HG2	1:C:202:PHE:HE1	1.19	1.04
1:C:127:GLU:HG2	1:C:202:PHE:CE1	1.95	1.02
3:B:120:ASN:ND2	3:B:186:ASP:OD2	1.94	1.01
1:C:34:GLN:NE2	1:C:127:GLU:OE2	1.99	0.95
2:A:59:ARG:NH1	5:A:301:SO4:O1	2.04	0.89

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	C	325/392 (83%)	305 (94%)	18 (6%)	2 (1%)	25 59
2	A	151/208 (73%)	144 (95%)	7 (5%)	0	100 100
3	B	234/253 (92%)	230 (98%)	4 (2%)	0	100 100
All	All	710/853 (83%)	679 (96%)	29 (4%)	2 (0%)	41 72

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	198	ASN
1	C	260	GLN

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	C	288/343 (84%)	263 (91%)	25 (9%)	10	34
2	A	126/183 (69%)	118 (94%)	8 (6%)	18	47
3	B	192/215 (89%)	172 (90%)	20 (10%)	7	25
All	All	606/741 (82%)	553 (91%)	53 (9%)	10	34

5 of 53 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	A	83	SER
3	B	49	SER
3	B	178	LEU
2	A	94	ASN
2	A	136	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	177	GLN
3	B	120	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 monosaccharides 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	$\#\|Z\| > 2$
5	SO4	C	402	-	4,4,4	0.23	0	6,6,6	0.14	0
4	1XL	C	401	-	24,24,24	3.61	4 (16%)	28,34,34	3.46	12 (42%)
5	SO4	B	301	-	4,4,4	0.21	0	6,6,6	0.15	0
5	SO4	B	302	-	4,4,4	0.22	0	6,6,6	0.11	0
5	SO4	A	301	-	4,4,4	0.22	0	6,6,6	0.25	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
4	1XL	C	401	-	-	5/16/16/16	0/2/2/2

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	C	401	1XL	CAI-C6	-16.31	1.37	1.50
4	C	401	1XL	C8A-N1	3.55	1.40	1.33
4	C	401	1XL	C7-C6	2.69	1.40	1.36
4	C	401	1XL	C8A-N8	-2.31	1.34	1.38

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	C	401	1XL	CAK-N8-C7	14.32	130.55	120.04
4	C	401	1XL	C6-C7-N8	-4.83	116.11	122.09
4	C	401	1XL	C4A-C8A-N8	3.99	122.19	116.48
4	C	401	1XL	C4-N3-C2	-3.20	119.97	125.64
4	C	401	1XL	C8A-C4A-N5	-2.83	119.11	124.80

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	C	401	1XL	CAU-CAK-N8-C7
4	C	401	1XL	CAU-CAK-N8-C8A
4	C	401	1XL	C7-C6-CAI-OAC
4	C	401	1XL	N5-C6-CAI-OAC
4	C	401	1XL	CAJ-CAT-CAV-OAG

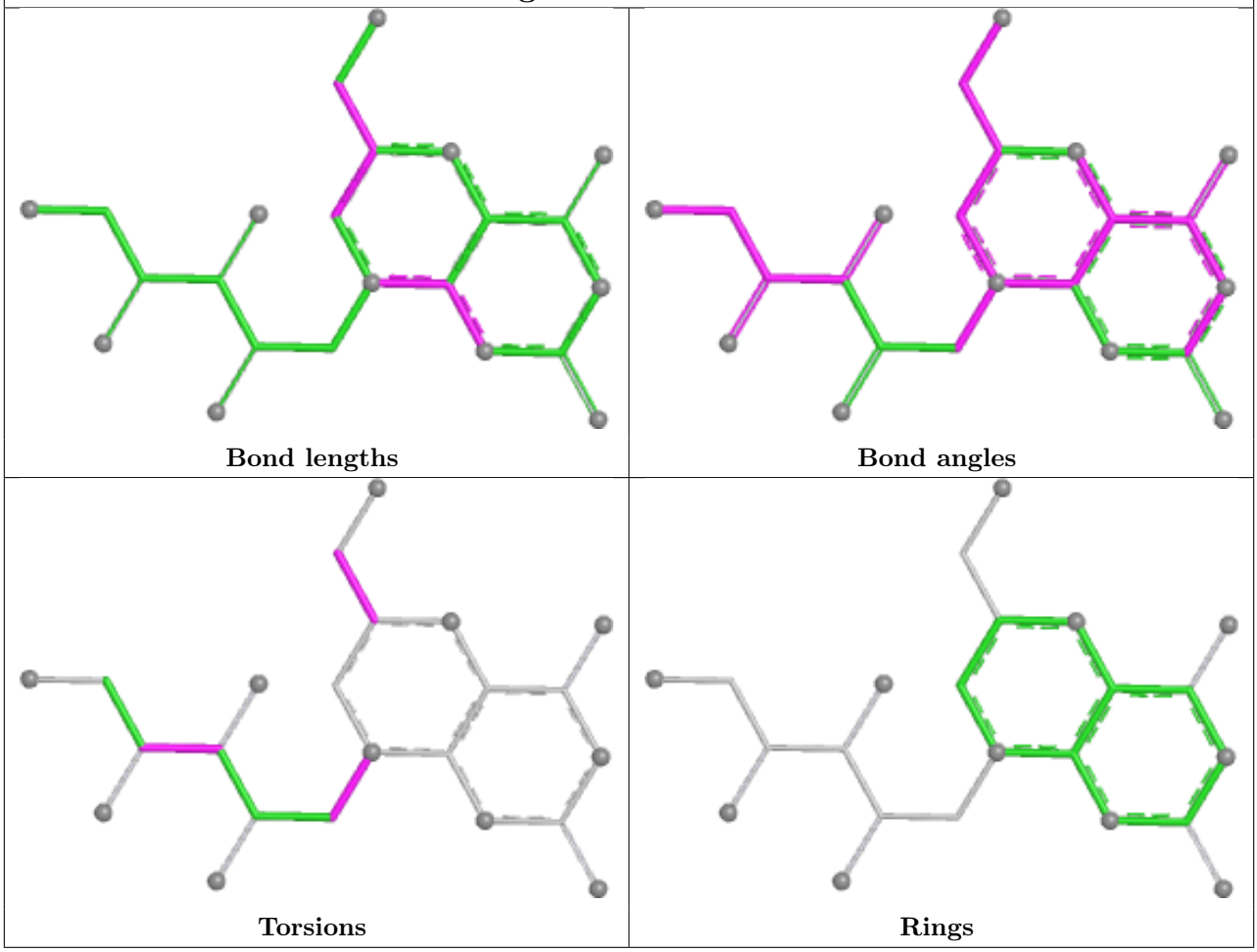
There are no ring outliers.

2 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	C	401	1XL	4	0
5	A	301	SO4	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand 1XL C 401



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	C	341/392 (86%)	0.55	16 (4%)	31	28	55, 90, 130, 167	0
2	A	157/208 (75%)	0.55	9 (5%)	23	22	55, 81, 113, 154	0
3	B	238/253 (94%)	0.49	5 (2%)	63	61	44, 82, 121, 139	0
All	All	736/853 (86%)	0.53	30 (4%)	37	34	44, 86, 126, 167	0

The worst 5 of 30 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	383	PHE	3.8
1	C	334	ILE	2.9
2	A	121	LEU	2.9
2	A	27	SER	2.8
2	A	151	VAL	2.7

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 monosaccharides in this entry.

6.4 Ligands (i)

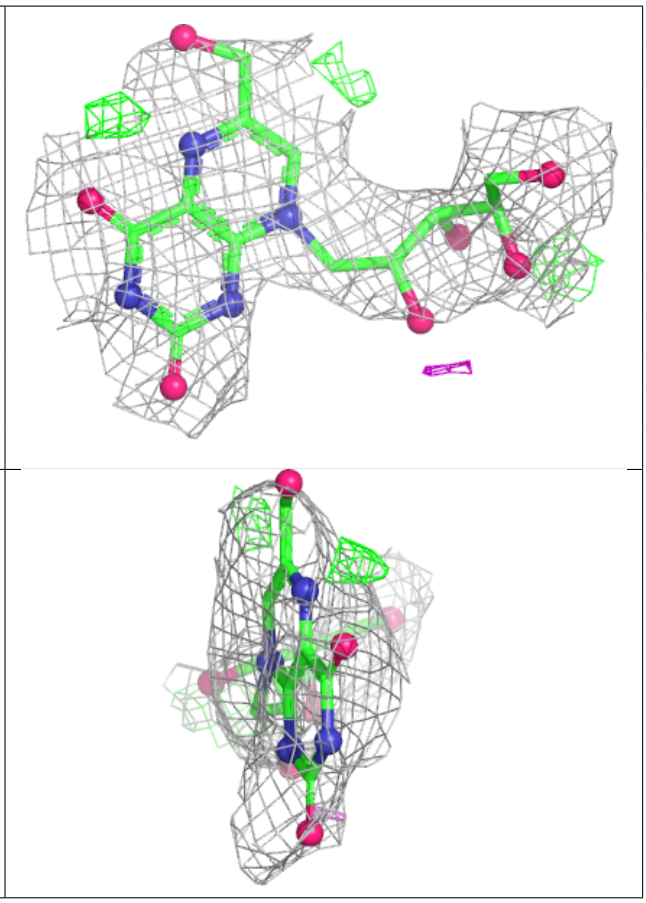
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	1XL	C	401	23/23	0.77	0.42	92,98,106,110	0
5	SO4	C	402	5/5	0.77	0.35	129,129,133,134	0
5	SO4	A	301	5/5	0.88	0.21	118,119,121,121	0
5	SO4	B	302	5/5	0.92	0.15	86,89,90,92	0
5	SO4	B	301	5/5	0.94	0.16	53,63,68,72	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 1XL C 401:

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
and green (positive)



6.5 Other polymers (i)

There are no such residues in this entry.