Full wwPDB X-ray Structure Validation Report (i)

Mar 16, 2022 – 02:16 PM EDT

PDB ID	:	6BSJ
Title	:	Structure of HIV-1 RT complexed with an RNA/DNA hybrid sequence non-preferred for RNA hydrolysis
Authors	:	Tian, L.; Kim, M.; Yang, W.
Deposited on	:	2017-12-03
Resolution	:	2.89 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 following versions of software and data (see references (i)) were used in the production of this report:

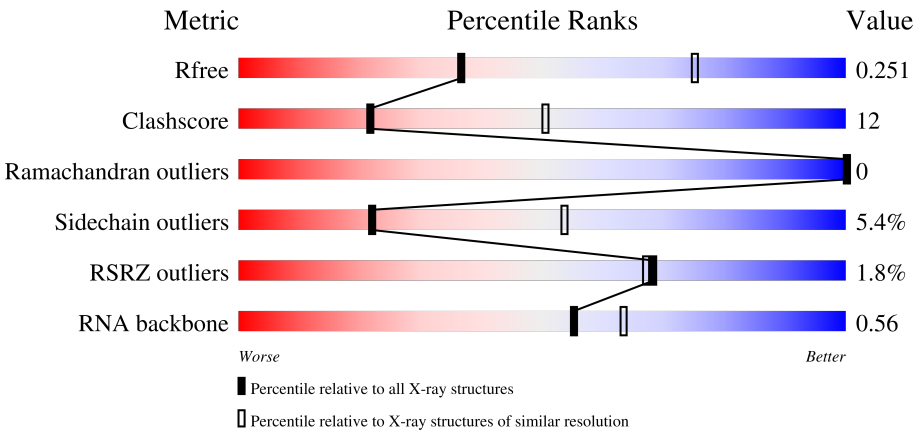
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.27
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.27

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 2.89 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	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)
RNA backbone	3102	1007 (3.16-2.64)

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	558	2% 68% 27% • •
2	B	441	% 68% 21% • 9%
3	D	23	39% 61%
4	R	25	4% 56% 40% •

2 Entry composition (i)

There are 9 unique types of molecules in this entry. The entry contains 8713 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 REVERSE TRANSCRIPTASE P66 SUBUNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	543	4389	2836	727	818	8	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	GLY	-	expression tag	UNP Q74085
A	68	GLY	SER	conflict	UNP Q74085
A	83	LYS	ARG	conflict	UNP Q74085
A	357	MET	THR	conflict	UNP Q74085
A	411	VAL	ILE	conflict	UNP Q74085
A	461	LYS	ARG	conflict	UNP Q74085
A	483	HIS	TYR	conflict	UNP Q74085
A	512	GLN	LYS	conflict	UNP Q74085

- Molecule 2 is a protein called REVERSE TRANSCRIPTASE P51 SUBUNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	400	3261	2127	528	600	6	0	1	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	0	GLY	-	expression tag	UNP A0A076Q3N8
B	68	GLY	SER	conflict	UNP A0A076Q3N8
B	83	LYS	ARG	conflict	UNP A0A076Q3N8
B	411	VAL	ILE	conflict	UNP A0A076Q3N8

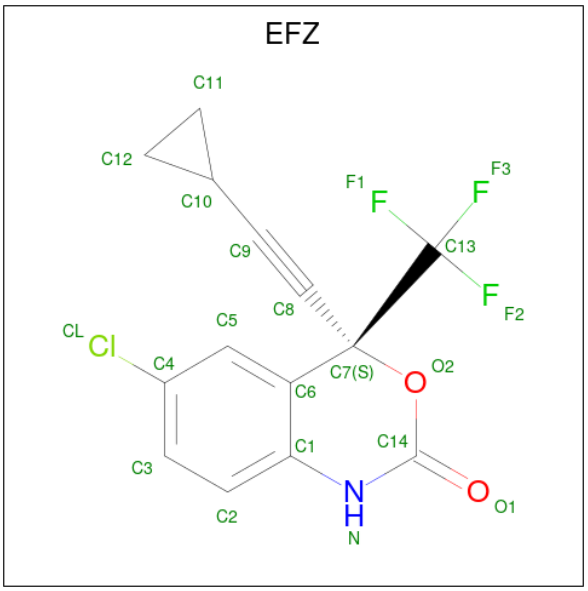
- Molecule 3 is a DNA chain called DNA (5'-D(*GP*TP*AP*TP*GP*CP*CP*TP*AP*TP*AP*GP*TP*TP*AP*TP*TP*GP*TP*GP*GP*CP*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	D	23	469	226	80	141	22	0	0	0

- Molecule 4 is a RNA chain called RNA (25-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
4	R	25	520	236	98	163	23	0	0	0

- Molecule 5 is (-)-6-CHLORO-4-CYCLOPROPYLETHYNYL-4-TRIFLUOROMETHYL-1,4-DIHYDRO-2H-3,1-BENZOXAZIN-2-ONE (three-letter code: EFZ) (formula: $C_{14}H_9ClF_3NO_2$).

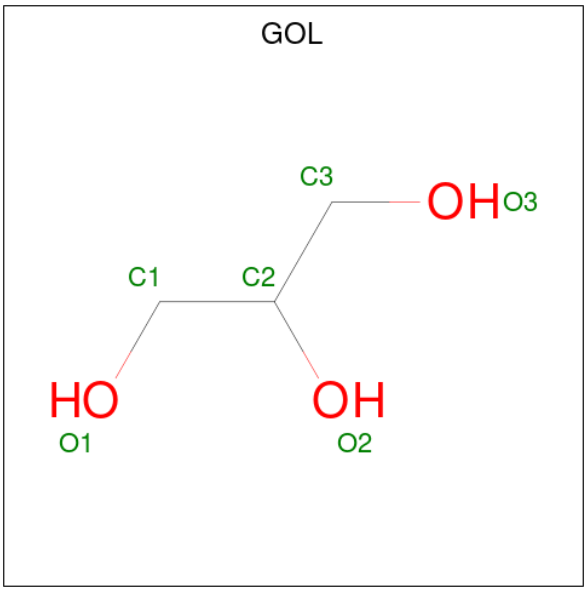


Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
			Total	C	Cl	F	N	O		
5	A	1	21	14	1	3	1	2	0	0

- Molecule 6 is CALCIUM ION (three-letter code: CA) (formula: Ca).

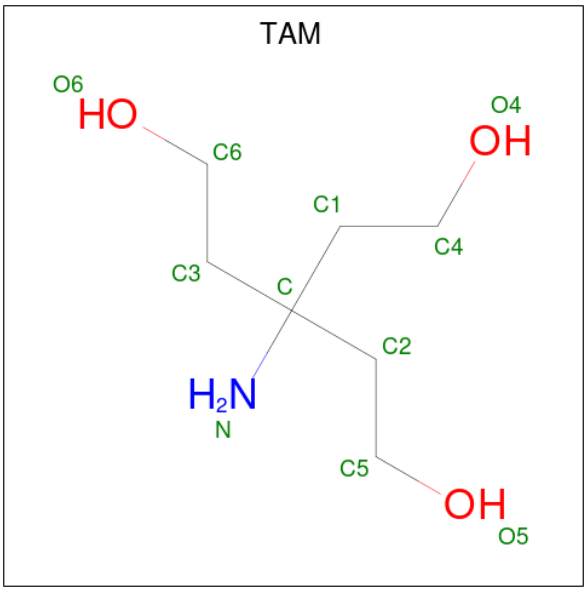
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Ca		
6	A	1	1	1	0	0

- Molecule 7 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
7	A	1	6	3	3	0	0
			Total	C	O		
7	A	1	6	3	3	0	0
			Total	C	O		
7	D	1	6	3	3	0	0
			Total	C	O		
7	D	1	6	3	3	0	0

- Molecule 8 is TRIS(HYDROXYETHYL)AMINOMETHANE (three-letter code: TAM) (formula: $C_7H_{17}NO_3$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
8	R	1	11	7	1	3	0	0

- Molecule 9 is water.

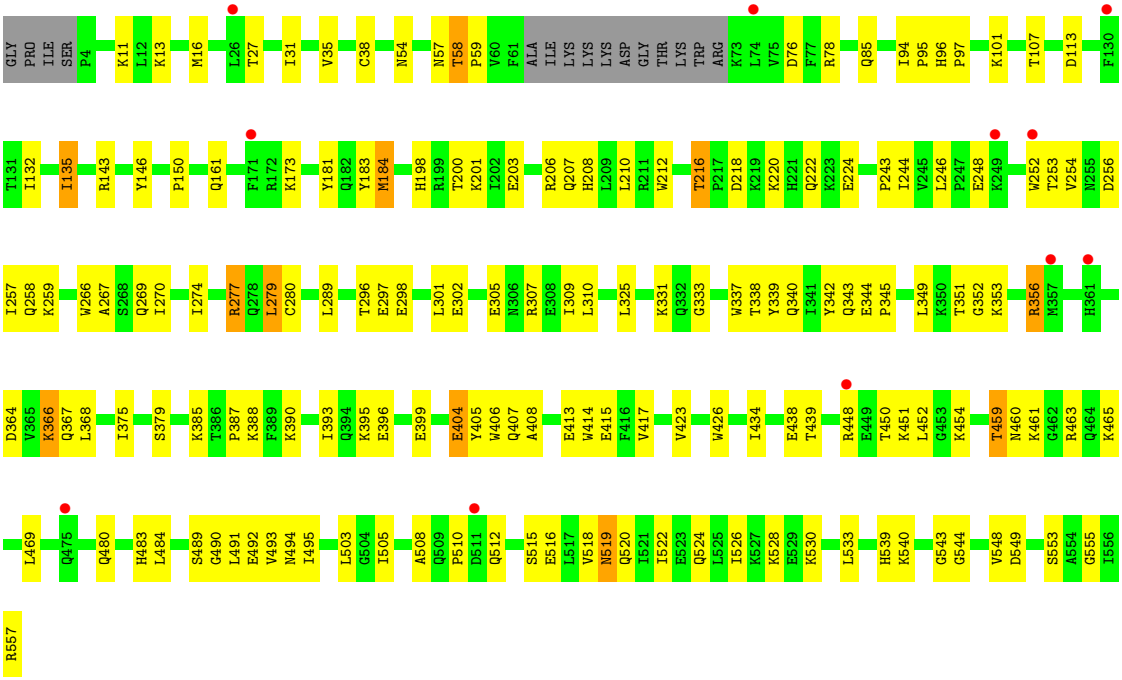
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
9	A	11	11	11	0	0
			Total	O		
9	B	5	5	5	0	0
			Total	O		
9	D	1	1	1	0	0

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.

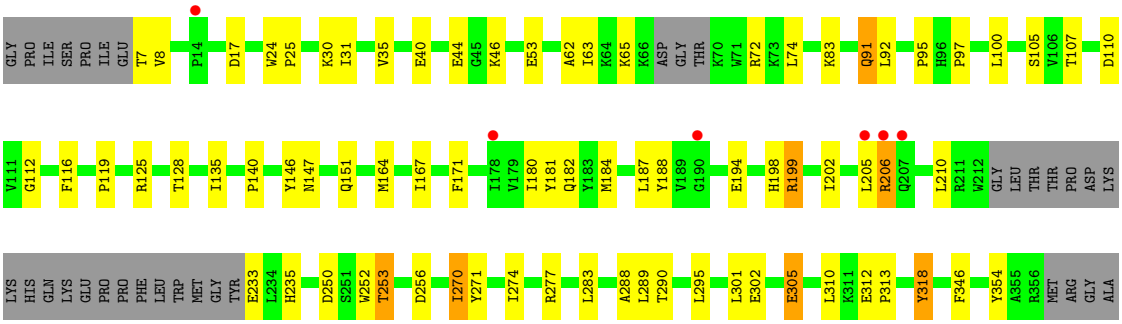
• Molecule 1: REVERSE TRANSCRIPTASE P66 SUBUNIT

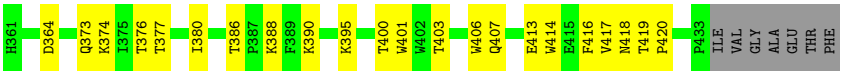
Chain A:



• Molecule 2: REVERSE TRANSCRIPTASE P51 SUBUNIT

Chain B:



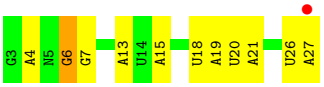


• Molecule 3: DNA (5'-D(*GP*TP*AP*TP*GP*CP*CP*TP*AP*TP*AP*GP*TP*TP*AP*TP *TP*GP*TP*GP*GP*CP*C)-3')

Chain D:

• Molecule 4: RNA (25-MER)

Chain R:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α, β, γ	163.24Å 163.24Å 129.40Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	41.20 – 2.89 41.20 – 2.89	Depositor EDS
% Data completeness (in resolution range)	100.0 (41.20-2.89) 100.0 (41.20-2.89)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.09	Depositor
$< I/\sigma(I) >$ [1]	2.13 (at 2.90Å)	Xtriage
Refinement program	PHENIX (1.10_2155: ???)	Depositor
R, R_{free}	0.201 , 0.251 0.201 , 0.251	Depositor DCC
R_{free} test set	1999 reflections (4.46%)	wwPDB-VP
Wilson B-factor (Å^2)	98.6	Xtriage
Anisotropy	0.289	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 55.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.026 for -h,-k,l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	8713	wwPDB-VP
Average B, all atoms (Å^2)	94.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.23% 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: 3DR, EFZ, GOL, CA, TAM

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	0.47	0/4502	0.59	0/6123
2	B	0.62	9/3358 (0.3%)	0.62	2/4580 (0.0%)
3	D	0.93	0/524	1.17	1/808 (0.1%)
4	R	0.61	0/569	1.16	1/881 (0.1%)
All	All	0.57	9/8953 (0.1%)	0.70	4/12392 (0.0%)

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	206	ARG	CG-CD	9.88	1.76	1.51
2	B	199	ARG	CZ-NH1	8.93	1.44	1.33
2	B	199	ARG	CZ-NH2	8.06	1.43	1.33
2	B	206	ARG	CZ-NH2	7.81	1.43	1.33
2	B	95	PRO	C-N	7.27	1.50	1.34
2	B	199	ARG	CG-CD	7.22	1.70	1.51
2	B	206	ARG	CB-CG	6.45	1.70	1.52
2	B	206	ARG	CZ-NH1	6.33	1.41	1.33
2	B	206	ARG	NE-CZ	5.77	1.40	1.33

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	199	ARG	NE-CZ-NH2	-8.94	115.83	120.30
2	B	206	ARG	NE-CZ-NH2	-8.06	116.27	120.30
4	R	15	A	OP2-P-O3'	5.77	117.90	105.20
3	D	9	DT	O4'-C4'-C3'	-5.38	102.35	104.50

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	A	4389	0	4397	101	0
2	B	3261	0	3203	70	0
3	D	469	0	264	15	0
4	R	520	0	272	23	0
5	A	21	0	9	1	0
6	A	1	0	0	0	0
7	A	12	0	16	1	0
7	D	12	0	16	0	0
8	R	11	0	17	2	0
9	A	11	0	0	0	0
9	B	5	0	0	0	0
9	D	1	0	0	0	0
All	All	8713	0	8194	196	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 12.

All (196) 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:B:206:ARG:CG	2:B:206:ARG:CD	1.76	1.58
1:A:426:TRP:HB3	1:A:526:ILE:HD11	1.52	0.91
4:R:6:G:H3'	4:R:6:G:P	2.15	0.86
1:A:448:ARG:NH2	3:D:6:DG:H21	1.76	0.83
4:R:6:G:H2'	4:R:7:G:H8	1.43	0.82
4:R:6:G:C2	4:R:7:G:C5	2.69	0.80
2:B:30:LYS:HD3	2:B:62:ALA:HB3	1.63	0.80
1:A:356:ARG:HH21	1:A:512:GLN:HG3	1.49	0.77
1:A:448:ARG:HH22	3:D:6:DG:H21	1.31	0.76
4:R:6:G:H2'	4:R:7:G:C8	2.22	0.74
1:A:277:ARG:HH22	1:A:356:ARG:HH12	1.34	0.73
2:B:125:ARG:O	2:B:128:THR:HG22	1.88	0.72
4:R:6:G:C2	4:R:7:G:C4	2.77	0.71
2:B:418:ASN:HD21	3:D:11:DT:H5'	1.56	0.70
4:R:6:G:N1	4:R:7:G:C6	2.60	0.69

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:D:5:DT:H2'	3:D:6:DG:H8	1.57	0.69
1:A:184:MET:HE1	3:D:23:DC:H2"	1.75	0.68
1:A:252:TRP:HB3	1:A:257:ILE:HD11	1.75	0.68
1:A:181:TYR:CE2	1:A:183:TYR:HB2	2.30	0.67
2:B:253:THR:HG22	2:B:256:ASP:H	1.59	0.66
2:B:128:THR:HG21	2:B:146:TYR:HB2	1.78	0.66
1:A:390:LYS:HE3	1:A:417:VAL:HG21	1.78	0.66
4:R:6:G:C4	4:R:7:G:C8	2.85	0.64
2:B:354:TYR:HB2	2:B:374:LYS:NZ	2.13	0.63
1:A:459:THR:HG23	1:A:461:LYS:H	1.65	0.61
1:A:405:TYR:CE2	1:A:407:GLN:HG2	2.35	0.60
2:B:206:ARG:CG	2:B:206:ARG:NE	2.62	0.60
1:A:356:ARG:NH2	1:A:512:GLN:HG3	2.16	0.60
2:B:128:THR:CG2	2:B:146:TYR:H	2.15	0.59
1:A:459:THR:HG22	1:A:463:ARG:H	1.68	0.58
1:A:342:TYR:HA	1:A:349:LEU:HD12	1.85	0.58
2:B:46:LYS:HE2	2:B:116:PHE:HB3	1.84	0.58
2:B:373:GLN:O	2:B:377:THR:HG23	2.05	0.57
1:A:277:ARG:HH22	1:A:356:ARG:NH1	2.01	0.57
2:B:270:ILE:HG13	2:B:271:TYR:CD1	2.39	0.57
1:A:519:ASN:HA	1:A:522:ILE:HD12	1.87	0.56
4:R:6:G:N2	4:R:7:G:C2	2.73	0.56
2:B:53:GLU:OE2	2:B:53:GLU:N	2.34	0.56
2:B:395:LYS:HG3	2:B:416:PHE:CE2	2.41	0.56
1:A:280:CYS:SG	4:R:13:A:H4'	2.45	0.56
2:B:7:THR:HG22	2:B:119:PRO:HG2	1.88	0.56
1:A:494:ASN:HB3	2:B:289:LEU:HD22	1.87	0.56
1:A:448:ARG:HH22	3:D:6:DG:N2	2.03	0.56
4:R:6:G:P	4:R:6:G:C3'	2.94	0.55
1:A:343:GLN:HG3	1:A:349:LEU:HD11	1.88	0.55
1:A:279:LEU:HD12	1:A:302:GLU:OE2	2.06	0.55
1:A:515:SER:HB3	1:A:518:VAL:HG12	1.88	0.55
1:A:367:GLN:CD	1:A:512:GLN:HE22	2.10	0.54
1:A:549:ASP:O	1:A:553:SER:OG	2.23	0.54
1:A:301:LEU:O	1:A:305:GLU:HG3	2.06	0.54
4:R:19:A:H2'	4:R:20:U:C6	2.43	0.54
1:A:483:HIS:CE1	1:A:520:GLN:HE21	2.25	0.54
2:B:128:THR:HG23	2:B:146:TYR:H	1.73	0.54
1:A:439:THR:HG21	2:B:289:LEU:HD13	1.91	0.53
2:B:274:ILE:HD11	2:B:310:LEU:HD21	1.91	0.53
4:R:6:G:C4	4:R:7:G:N7	2.76	0.53

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:107:THR:HG23	1:A:198:HIS:HE2	1.73	0.53
2:B:135:ILE:HD12	2:B:135:ILE:H	1.74	0.53
2:B:270:ILE:HG22	2:B:346:PHE:HB3	1.90	0.53
1:A:406:TRP:CZ3	1:A:407:GLN:HB3	2.44	0.53
2:B:65:LYS:HA	2:B:407:GLN:HE21	1.73	0.53
3:D:5:DT:H2'	3:D:6:DG:C8	2.42	0.53
1:A:38:CYS:SG	1:A:132:ILE:HD11	2.49	0.52
1:A:253:THR:HG23	1:A:256:ASP:H	1.74	0.52
1:A:337:TRP:CZ3	1:A:368:LEU:HD13	2.45	0.52
2:B:253:THR:CG2	2:B:256:ASP:H	2.22	0.52
2:B:125:ARG:HD3	2:B:147:ASN:HA	1.91	0.52
1:A:269:GLN:HA	1:A:351:THR:O	2.10	0.52
1:A:503:LEU:HD23	1:A:533:LEU:HG	1.92	0.52
1:A:544:GLY:O	1:A:548:VAL:HG12	2.11	0.51
2:B:72:ARG:NH2	2:B:110:ASP:OD2	2.43	0.51
2:B:206:ARG:CD	2:B:206:ARG:CB	2.86	0.51
2:B:180:ILE:HG12	2:B:187:LEU:HD12	1.92	0.51
1:A:366:LYS:NZ	1:A:405:TYR:OH	2.43	0.51
2:B:105:SER:OG	2:B:235:HIS:ND1	2.32	0.50
1:A:489:SER:OG	1:A:493:VAL:HG11	2.10	0.50
1:A:460:ASN:ND2	2:B:288:ALA:HB2	2.27	0.50
4:R:6:G:N1	4:R:7:G:C5	2.80	0.49
1:A:248:GLU:HA	1:A:307:ARG:NH1	2.27	0.49
2:B:395:LYS:NZ	3:D:13:DG:OP1	2.37	0.49
4:R:6:G:H2'	4:R:7:G:OP2	2.12	0.49
1:A:439:THR:CG2	2:B:289:LEU:HD13	2.43	0.49
1:A:454:LYS:NZ	1:A:555:GLY:H	2.11	0.49
1:A:379:SER:CB	1:A:387:PRO:HD3	2.42	0.48
2:B:390:LYS:HB3	2:B:417:VAL:HG11	1.95	0.48
1:A:339:TYR:CZ	1:A:352:GLY:HA3	2.49	0.48
1:A:325:LEU:HD12	1:A:385:LYS:HE2	1.95	0.48
2:B:125:ARG:HD3	2:B:146:TYR:O	2.13	0.48
3:D:6:DG:H2'	3:D:7:DC:C6	2.49	0.48
4:R:6:G:O2'	4:R:7:G:O4'	2.23	0.48
4:R:6:G:C2	4:R:7:G:C6	3.01	0.48
2:B:164:MET:HA	2:B:167:ILE:HG22	1.96	0.48
1:A:450:THR:O	1:A:451:LYS:HG2	2.15	0.47
2:B:302:GLU:HA	2:B:305:GLU:OE1	2.13	0.47
2:B:388:LYS:HG3	2:B:413:GLU:HG2	1.96	0.47
1:A:58:THR:HG22	1:A:59:PRO:HD2	1.97	0.47
1:A:393:ILE:HG13	1:A:423:VAL:HB	1.97	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:57:ASN:HB2	7:A:603:GOL:H11	1.97	0.47
1:A:76:ASP:OD1	1:A:78:ARG:HG3	2.15	0.47
2:B:301:LEU:HD12	2:B:301:LEU:HA	1.73	0.46
2:B:97:PRO:HG3	2:B:181:TYR:HB2	1.95	0.46
1:A:203:GLU:O	1:A:207:GLN:HG3	2.15	0.46
2:B:180:ILE:HA	2:B:188:TYR:O	2.16	0.46
1:A:11:LYS:O	1:A:85:GLN:HB3	2.15	0.46
3:D:14:DT:H2'	3:D:15:DT:C6	2.50	0.46
1:A:408:ALA:HB1	2:B:364:ASP:HB3	1.98	0.46
1:A:135:ILE:H	1:A:135:ILE:HG12	1.51	0.46
1:A:181:TYR:HE2	1:A:183:TYR:HB2	1.78	0.45
1:A:553:SER:HB2	1:A:557:ARG:HG3	1.98	0.45
2:B:312:GLU:HB2	2:B:313:PRO:HD2	1.98	0.45
2:B:171:PHE:CD2	2:B:205:LEU:HD13	2.50	0.45
1:A:13:LYS:HB2	1:A:16:MET:HG3	1.98	0.45
1:A:434:ILE:HD13	1:A:530:LYS:HB3	1.98	0.45
4:R:26:U:H2'	4:R:27:A:C8	2.51	0.45
2:B:354:TYR:HB2	2:B:374:LYS:HZ2	1.78	0.45
1:A:452:LEU:HD22	1:A:469:LEU:O	2.17	0.45
1:A:201:LYS:HD3	1:A:201:LYS:HA	1.71	0.45
2:B:63:ILE:HG23	2:B:406:TRP:O	2.17	0.45
1:A:340:GLN:HG3	1:A:351:THR:HG22	1.98	0.45
2:B:376:THR:O	2:B:380:ILE:HG13	2.17	0.45
1:A:218:ASP:O	1:A:222:GLN:HG3	2.17	0.45
1:A:259:LYS:HB3	3:D:19:DG:O3'	2.17	0.44
1:A:331:LYS:HE2	1:A:333:GLY:HA2	2.00	0.44
2:B:31:ILE:O	2:B:35:VAL:HG23	2.16	0.44
1:A:113:ASP:N	1:A:113:ASP:OD1	2.51	0.44
1:A:388:LYS:HE3	1:A:413:GLU:OE1	2.18	0.44
2:B:318[A]:TYR:HD2	2:B:318[A]:TYR:H	1.65	0.44
2:B:24:TRP:CG	2:B:25:PRO:HD2	2.52	0.44
2:B:40:GLU:O	2:B:44:GLU:HG3	2.18	0.44
2:B:354:TYR:HB2	2:B:374:LYS:HZ3	1.80	0.44
1:A:244:ILE:HD13	1:A:267:ALA:HB2	1.99	0.43
3:D:15:DT:C2	3:D:16:DA:C8	3.05	0.43
1:A:208:HIS:CD2	1:A:212:TRP:HZ3	2.36	0.43
2:B:107:THR:HG23	2:B:198:HIS:HE1	1.83	0.43
1:A:367:GLN:NE2	1:A:512:GLN:HE22	2.16	0.43
8:R:101:TAM:H12	8:R:101:TAM:H51	1.82	0.43
1:A:181:TYR:CD2	5:A:601:EFZ:H101	2.53	0.43
1:A:296:THR:OG1	1:A:298:GLU:OE1	2.24	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:254:VAL:O	1:A:258:GLN:HG3	2.18	0.43
1:A:395:LYS:HG2	1:A:414:TRP:CH2	2.54	0.43
1:A:483:HIS:ND1	1:A:524:GLN:OE1	2.50	0.43
2:B:91:GLN:HB3	2:B:92:LEU:HD12	1.99	0.43
2:B:199:ARG:O	2:B:202:ILE:HG12	2.18	0.43
3:D:17:DT:H2'	3:D:18:DT:C6	2.53	0.43
1:A:543:GLY:N	2:B:283:LEU:O	2.51	0.43
2:B:167:ILE:HG12	2:B:167:ILE:O	2.19	0.43
1:A:438:GLU:OE2	1:A:463:ARG:HD3	2.18	0.42
1:A:146:TYR:CD2	1:A:150:PRO:HB3	2.54	0.42
1:A:270:ILE:HD12	1:A:270:ILE:HA	1.79	0.42
2:B:252:TRP:CD1	2:B:295:LEU:HD11	2.54	0.42
4:R:6:G:H3'	4:R:6:G:OP2	2.19	0.42
4:R:6:G:C5	4:R:7:G:N7	2.88	0.42
1:A:54:ASN:HB3	1:A:143:ARG:NH2	2.34	0.42
1:A:224:GLU:N	1:A:224:GLU:OE2	2.53	0.42
1:A:31:ILE:O	1:A:35:VAL:HG12	2.19	0.42
1:A:101:LYS:HE3	1:A:101:LYS:HB2	1.82	0.42
1:A:404:GLU:O	1:A:404:GLU:HG3	2.20	0.42
1:A:331:LYS:NZ	1:A:364:ASP:OD2	2.37	0.42
1:A:465:LYS:HD3	1:A:484:LEU:HD13	2.01	0.42
2:B:180:ILE:HG13	2:B:188:TYR:O	2.20	0.42
1:A:266:TRP:O	1:A:269:GLN:HG2	2.19	0.42
3:D:5:DT:C2'	3:D:6:DG:H8	2.30	0.42
4:R:20:U:H2'	4:R:21:A:C8	2.55	0.42
1:A:492:GLU:HG2	1:A:530:LYS:HB2	2.01	0.41
2:B:182:GLN:HE21	2:B:182:GLN:HB3	1.57	0.41
8:R:101:TAM:H41	8:R:101:TAM:H32	1.74	0.41
1:A:146:TYR:CE2	1:A:150:PRO:HB3	2.55	0.41
1:A:210:LEU:HD12	1:A:210:LEU:HA	1.76	0.41
1:A:206:ARG:NH1	1:A:216:THR:O	2.53	0.41
1:A:508:ALA:O	1:A:510:PRO:HD3	2.20	0.41
2:B:17:ASP:O	2:B:83:LYS:HD3	2.21	0.41
2:B:400:THR:HG22	2:B:401:TRP:CD2	2.54	0.41
4:R:18:U:H2'	4:R:19:A:C8	2.55	0.41
1:A:344:GLU:HA	1:A:345:PRO:HD3	1.96	0.41
1:A:94:ILE:HD12	1:A:95:PRO:O	2.20	0.41
1:A:96:HIS:CG	1:A:97:PRO:HD2	2.56	0.41
1:A:339:TYR:CG	1:A:375:ILE:HD11	2.56	0.41
1:A:490:GLY:O	1:A:528:LYS:NZ	2.39	0.41
2:B:40:GLU:HG2	2:B:44:GLU:OE2	2.21	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:107:THR:HG22	2:B:233:GLU:HA	2.03	0.41
4:R:6:G:N3	4:R:7:G:C4	2.88	0.41
1:A:406:TRP:CE3	1:A:407:GLN:HB3	2.55	0.41
2:B:419:THR:HA	2:B:420:PRO:HD3	1.97	0.41
4:R:18:U:H2'	4:R:19:A:H8	1.85	0.41
1:A:539:HIS:O	1:A:540:LYS:HD2	2.21	0.40
2:B:198:HIS:O	2:B:202:ILE:HG23	2.21	0.40
2:B:250:ASP:OD1	2:B:250:ASP:N	2.38	0.40
1:A:161:GLN:HE21	2:B:140:PRO:HB3	1.86	0.40
1:A:254:VAL:HG13	1:A:289:LEU:HA	2.04	0.40
1:A:480:GLN:O	1:A:483:HIS:HB3	2.21	0.40
3:D:19:DG:H2'	3:D:20:DT:C6	2.56	0.40
2:B:270:ILE:HG22	2:B:346:PHE:O	2.22	0.40
1:A:495:ILE:HB	1:A:533:LEU:HD13	2.03	0.40
2:B:112:GLY:C	2:B:151:GLN:HE21	2.25	0.40

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	539/558 (97%)	514 (95%)	25 (5%)	0	100	100
2	B	393/441 (89%)	375 (95%)	18 (5%)	0	100	100
All	All	932/999 (93%)	889 (95%)	43 (5%)	0	100	100

There are no Ramachandran outliers to report.

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	479/496 (97%)	450 (94%)	29 (6%)	18	48
2	B	352/399 (88%)	335 (95%)	17 (5%)	25	58
All	All	831/895 (93%)	785 (94%)	46 (6%)	22	53

All (46) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	27	THR
1	A	58	THR
1	A	135	ILE
1	A	173	LYS
1	A	184	MET
1	A	200	THR
1	A	216	THR
1	A	220	LYS
1	A	243	PRO
1	A	246	LEU
1	A	274	ILE
1	A	277	ARG
1	A	279	LEU
1	A	297	GLU
1	A	309	ILE
1	A	310	LEU
1	A	338	THR
1	A	353	LYS
1	A	356	ARG
1	A	366	LYS
1	A	396	GLU
1	A	399	GLU
1	A	404	GLU
1	A	415	GLU
1	A	459	THR
1	A	491	LEU
1	A	505	ILE
1	A	516	GLU
1	A	519	ASN
2	B	8	VAL
2	B	74	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	91	GLN
2	B	100	LEU
2	B	184	MET
2	B	194	GLU
2	B	210	LEU
2	B	253	THR
2	B	270	ILE
2	B	277	ARG
2	B	290	THR
2	B	305	GLU
2	B	318[A]	TYR
2	B	318[B]	TYR
2	B	386	THR
2	B	403	THR
2	B	414	TRP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (12) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	235	HIS
1	A	330	GLN
1	A	332	GLN
1	A	361	HIS
1	A	464	GLN
1	A	520	GLN
2	B	151	GLN
2	B	182	GLN
2	B	198	HIS
2	B	407	GLN
2	B	418	ASN
2	B	428	GLN

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
4	R	22/25 (88%)	2 (9%)	0

All (2) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
4	R	4	A
4	R	6	G

There are no RNA pucker outliers to report.

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

1 non-standard protein/DNA/RNA residue is 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	$\#\|Z\| > 2$
4	3DR	R	5	4	8,11,12	0.47	0	9,14,17	1.20	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	3DR	R	5	4	-	3/3/15/16	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	R	5	3DR	O4'-C4'-C5'-O5'
4	R	5	3DR	C3'-C4'-C5'-O5'
4	R	5	3DR	C4'-C5'-O5'-P

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 7 ligands modelled in this entry, 1 is monoatomic - leaving 6 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	$\#\|Z\| > 2$
8	TAM	R	101	-	7,10,10	1.11	0	9,12,12	0.60	0
5	EFZ	A	601	-	23,23,23	5.94	9 (39%)	36,36,36	5.29	20 (55%)
7	GOL	D	102	-	5,5,5	0.43	0	5,5,5	0.38	0
7	GOL	A	604	-	5,5,5	0.36	0	5,5,5	0.34	0
7	GOL	A	603	-	5,5,5	0.33	0	5,5,5	0.36	0
7	GOL	D	101	-	5,5,5	0.32	0	5,5,5	0.52	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
8	TAM	R	101	-	-	7/12/12/12	-
5	EFZ	A	601	-	-	0/10/32/32	0/3/3/3
7	GOL	D	102	-	-	3/4/4/4	-
7	GOL	A	604	-	-	2/4/4/4	-
7	GOL	A	603	-	-	2/4/4/4	-
7	GOL	D	101	-	-	3/4/4/4	-

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	601	EFZ	C1-C6	19.62	1.62	1.40
5	A	601	EFZ	C12-C11	11.37	1.91	1.48

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	601	EFZ	C5-C4	-7.59	1.25	1.38
5	A	601	EFZ	C3-C4	-7.22	1.24	1.38
5	A	601	EFZ	C2-C1	7.19	1.51	1.39
5	A	601	EFZ	C7-C6	7.01	1.60	1.51
5	A	601	EFZ	C10-C9	5.89	1.66	1.46
5	A	601	EFZ	O2-C7	-5.32	1.37	1.45
5	A	601	EFZ	C7-C8	2.64	1.55	1.47

All (20) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	A	601	EFZ	C2-C1-C6	-16.42	103.73	119.88
5	A	601	EFZ	O2-C7-C6	-11.96	103.20	111.64
5	A	601	EFZ	C2-C1-N	11.06	138.75	119.84
5	A	601	EFZ	C12-C10-C11	10.71	89.60	59.23
5	A	601	EFZ	C7-O2-C14	10.16	139.91	121.44
5	A	601	EFZ	C11-C12-C10	-6.28	44.93	60.38
5	A	601	EFZ	C12-C11-C10	-6.06	45.47	60.38
5	A	601	EFZ	C3-C4-CL	-5.88	110.16	119.35
5	A	601	EFZ	C3-C4-C5	4.20	127.11	121.53
5	A	601	EFZ	C13-C7-C8	-3.98	102.48	108.89
5	A	601	EFZ	F3-C13-C7	-3.77	107.18	111.79
5	A	601	EFZ	C5-C6-C1	3.33	122.92	118.72
5	A	601	EFZ	O2-C14-O1	3.32	123.04	118.02
5	A	601	EFZ	C13-C7-C6	3.31	116.12	110.84
5	A	601	EFZ	C2-C3-C4	2.96	122.36	119.24
5	A	601	EFZ	O2-C7-C13	2.94	110.18	104.76
5	A	601	EFZ	C5-C6-C7	-2.94	119.30	122.71
5	A	601	EFZ	C5-C4-CL	2.86	122.73	119.15
5	A	601	EFZ	C3-C2-C1	2.04	123.59	119.64
5	A	601	EFZ	C11-C10-C9	-2.03	113.54	119.06

There are no chirality outliers.

All (17) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	A	604	GOL	O1-C1-C2-O2
7	A	604	GOL	O1-C1-C2-C3
7	D	101	GOL	O1-C1-C2-C3
7	D	102	GOL	O1-C1-C2-C3
8	R	101	TAM	C1-C-C2-C5
8	R	101	TAM	C3-C-C2-C5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
8	R	101	TAM	N-C-C2-C5
8	R	101	TAM	C1-C-C3-C6
8	R	101	TAM	C2-C-C3-C6
8	R	101	TAM	N-C-C3-C6
8	R	101	TAM	C-C2-C5-O5
7	D	101	GOL	O1-C1-C2-O2
7	D	102	GOL	O1-C1-C2-O2
7	A	603	GOL	O1-C1-C2-O2
7	D	102	GOL	C1-C2-C3-O3
7	A	603	GOL	O1-C1-C2-C3
7	D	101	GOL	O2-C2-C3-O3

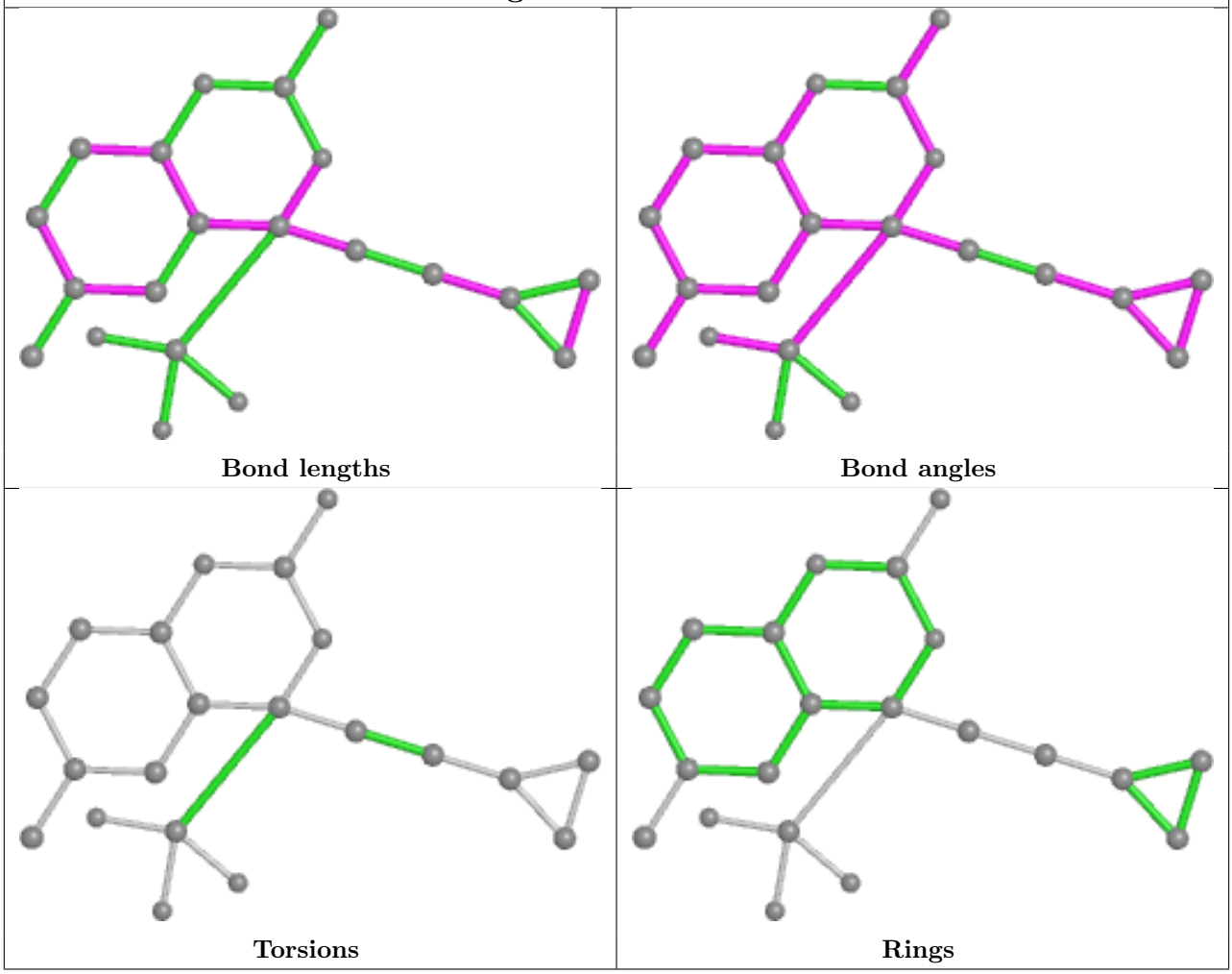
There are no ring outliers.

3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
8	R	101	TAM	2	0
5	A	601	EFZ	1	0
7	A	603	GOL	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 EFZ A 601



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	543/558 (97%)	0.12	11 (2%) 65 63	62, 87, 124, 155	0
2	B	400/441 (90%)	0.03	6 (1%) 73 73	64, 91, 146, 162	0
3	D	23/23 (100%)	-0.30	0 100 100	73, 88, 166, 180	0
4	R	24/25 (96%)	-0.05	1 (4%) 36 32	83, 92, 136, 161	0
All	All	990/1047 (94%)	0.07	18 (1%) 68 67	62, 89, 136, 180	0

All (18) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	252	TRP	4.5
2	B	178	ILE	3.6
1	A	26	LEU	3.6
4	R	27	A	3.3
1	A	511	ASP	3.3
2	B	206	ARG	2.7
1	A	475	GLN	2.6
1	A	448	ARG	2.6
1	A	357	MET	2.6
1	A	361	HIS	2.5
2	B	205	LEU	2.5
2	B	14	PRO	2.4
1	A	130	PHE	2.2
1	A	249	LYS	2.2
1	A	74	LEU	2.1
2	B	190	GLY	2.1
2	B	207	GLN	2.1
1	A	171	PHE	2.0

6.2 Non-standard residues in protein, DNA, RNA chains (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	3DR	R	5	11/12	0.88	0.16	112,124,142,145	0

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

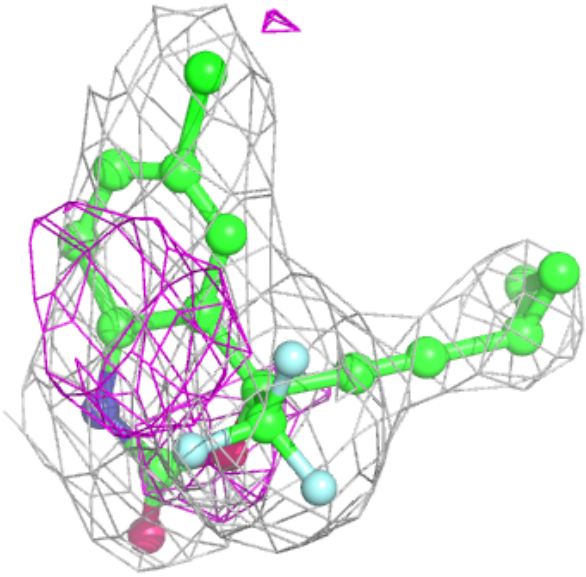
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

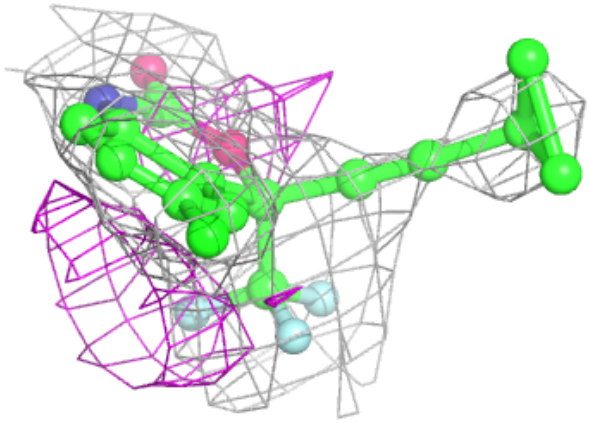
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
7	GOL	D	102	6/6	0.43	0.36	110,111,114,115	0
7	GOL	A	604	6/6	0.85	0.17	98,105,112,114	0
7	GOL	A	603	6/6	0.86	0.14	95,97,104,115	0
7	GOL	D	101	6/6	0.88	0.23	92,108,112,115	0
8	TAM	R	101	11/11	0.90	0.25	99,107,111,111	0
6	CA	A	602	1/1	0.93	0.08	120,120,120,120	0
5	EFZ	A	601	21/21	0.95	0.33	65,75,83,89	0

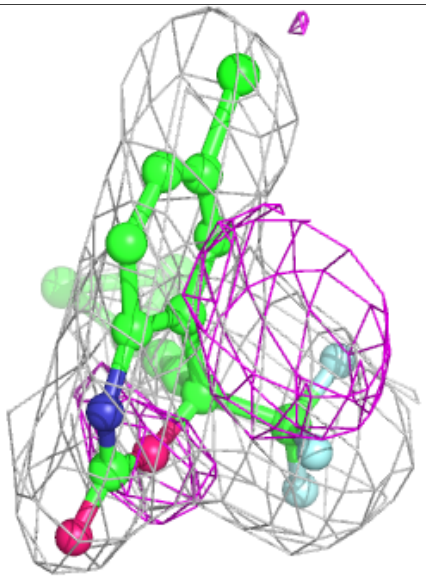
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around EFZ A 601:

$2mF_o$-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.