Full wwPDB X-ray Structure Validation Report ⓘ

Mar 16, 2022 – 02:23 PM EDT

PDB ID	:	6BSI
Title	:	Structure of HIV-1 RT complexed with an RNA/DNA hybrid containing the polypurine-tract sequence
Authors	:	Tian, L.; Kim, M.; Yang, W.
Deposited on	:	2017-12-03
Resolution	:	3.25 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 ⓘ symbol.

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

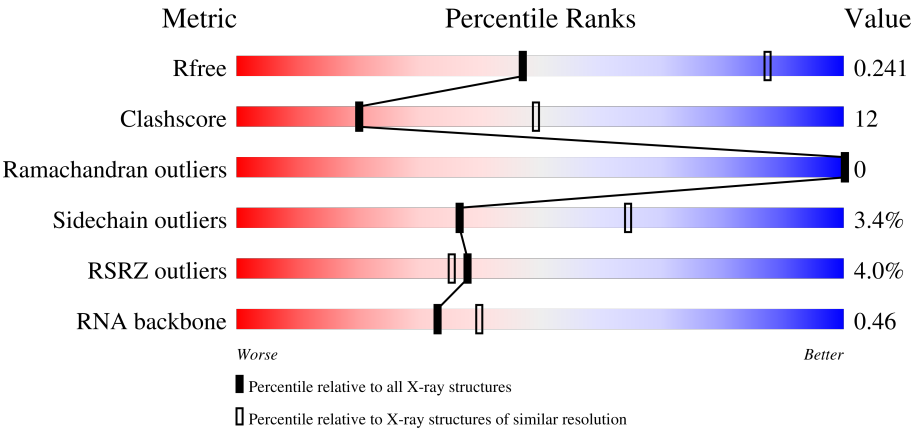
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.27
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (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 3.25 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)
RNA backbone	3102	1072 (3.62-2.90)

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	3% 67% 28%
2	B	441	5% 67% 24% 8%
3	D	23	52% 39% 9%
4	R	25	40% 28% 32%

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
5	EFZ	A	601	-	X	-	-

2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 8566 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	4307	2784	703	812	8	0	1	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	406	3227	2111	514	596	6	0	0	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

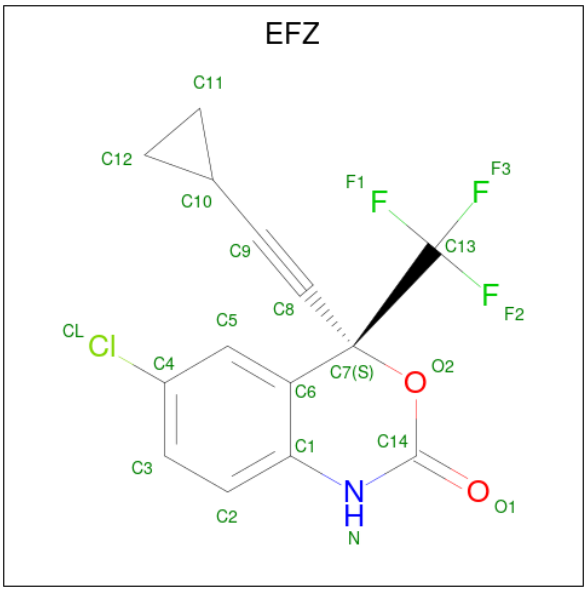
- Molecule 3 is a DNA chain called DNA (5'-D(*GP*TP*TP*TP*TP*TP*CP*TP*TP*TP*TP*GP*TP*TP*AP*TP*TP*GP*TP*GP*GP*CP*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	23	Total	C	N	O	P	0	0	0
			465	227	67	149	22			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	R	25	Total	C	N	O	P	0	0	0
			527	240	109	155	23			

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

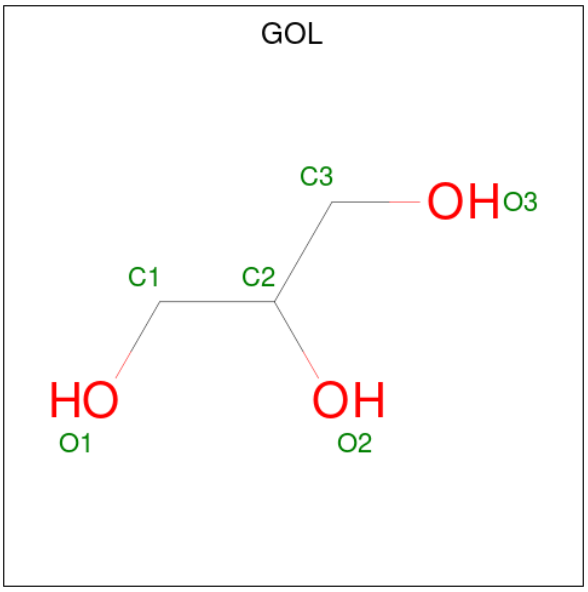


Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
5	A	1	Total	C	Cl	F	N	O	0	0
			21	14	1	3	1	2		

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

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

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



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

- Molecule 8 is water.

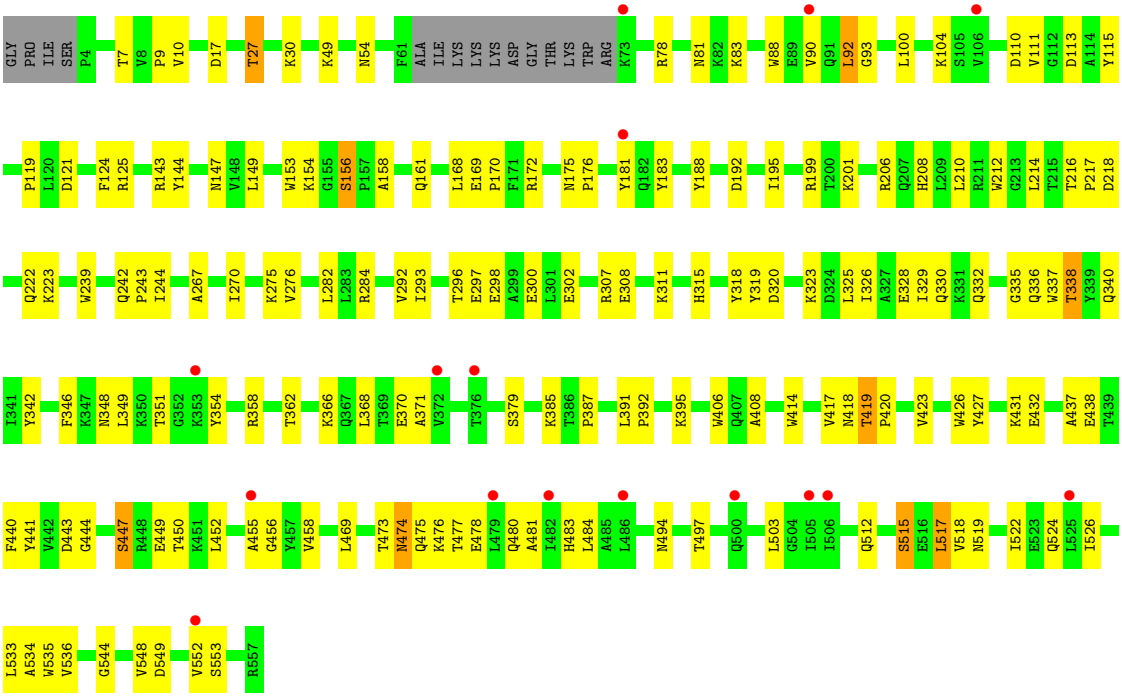
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	4	Total O 4 4	0	0
8	B	1	Total O 1 1	0	0

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.

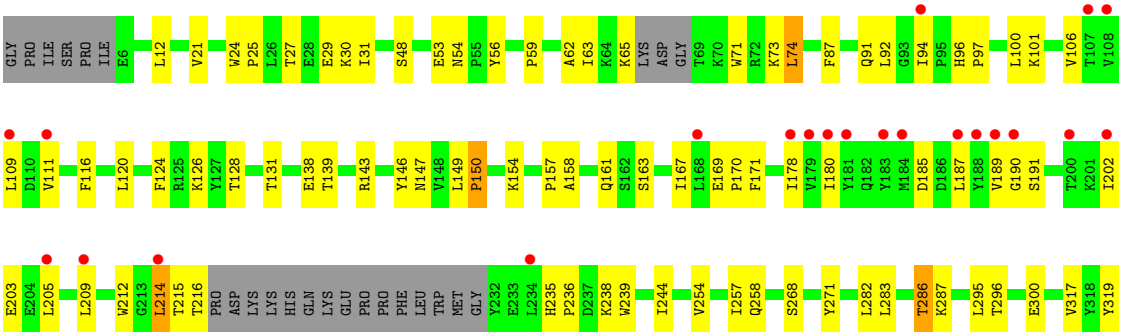
• Molecule 1: REVERSE TRANSCRIPTASE P66 SUBUNIT

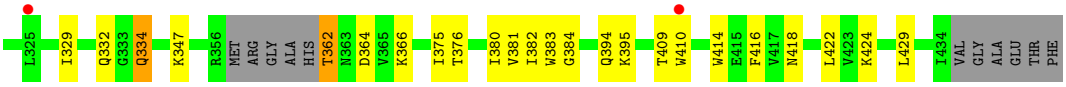
Chain A:



• Molecule 2: REVERSE TRANSCRIPTASE P51 SUBUNIT

Chain B:





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

• Molecule 4: RNA (25-MER)

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α, β, γ	163.85Å 163.85Å 129.57Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	47.30 – 3.25 95.68 – 3.25	Depositor EDS
% Data completeness (in resolution range)	99.7 (47.30-3.25) 92.8 (95.68-3.25)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.09	Depositor
$< I/\sigma(I) >$ [1]	1.27 (at 3.26Å)	Xtriage
Refinement program	PHENIX (1.10_2155: ???)	Depositor
R, R_{free}	0.193 , 0.241 0.193 , 0.241	Depositor DCC
R_{free} test set	1590 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	115.9	Xtriage
Anisotropy	0.274	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 83.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.032 for -h,-k,l	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	8566	wwPDB-VP
Average B, all atoms (Å^2)	118.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.19% 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: CA, EFZ, 3DR, GOL

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	0.54	1/4423 (0.0%)	0.66	1/6039 (0.0%)
2	B	0.47	0/3319	0.61	1/4538 (0.0%)
3	D	1.03	1/516 (0.2%)	1.36	2/796 (0.3%)
4	R	0.71	0/580	1.32	4/899 (0.4%)
All	All	0.57	2/8838 (0.0%)	0.78	8/12272 (0.1%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	223	LYS	C-N	-5.75	1.20	1.34
3	D	17	DT	C3'-O3'	-5.53	1.36	1.44

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	R	18	A	C8-N9-C4	7.92	108.97	105.80
3	D	13	DG	O4'-C4'-C3'	-6.56	101.88	104.50
4	R	18	A	N9-C4-C5	-6.34	103.26	105.80
4	R	21	A	O5'-P-OP1	-6.19	100.13	105.70
2	B	12	LEU	CA-CB-CG	6.03	129.17	115.30
3	D	15	DT	O5'-P-OP2	-5.90	100.39	105.70
4	R	17	C	C6-N1-C2	-5.80	117.98	120.30
1	A	92	LEU	CA-CB-CG	5.53	128.01	115.30

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	4307	0	4210	118	0
2	B	3227	0	3135	71	0
3	D	465	0	269	11	0
4	R	527	0	275	17	0
5	A	21	0	9	3	0
6	A	1	0	0	0	0
6	D	1	0	0	0	0
7	D	6	0	8	1	0
7	R	6	0	8	0	0
8	A	4	0	0	0	0
8	B	1	0	0	0	0
All	All	8566	0	7914	205	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 (205) 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:A:519:ASN:HA	1:A:522:ILE:HD12	1.52	0.89
3:D:11:DT:H3	4:R:20:A:H62	1.22	0.86
1:A:426:TRP:HB3	1:A:526:ILE:HD11	1.56	0.86
1:A:308:GLU:HA	1:A:311:LYS:HE2	1.68	0.74
2:B:111:VAL:HG21	2:B:187:LEU:HD13	1.70	0.73
1:A:427:TYR:HE2	1:A:522:ILE:HG23	1.56	0.70
2:B:380:ILE:O	2:B:384:GLY:N	2.22	0.70
1:A:92:LEU:HD12	1:A:93:GLY:H	1.57	0.69
1:A:335:GLY:H	1:A:512:GLN:HE22	1.41	0.68
1:A:115:TYR:CE1	1:A:156:SER:HB3	2.29	0.68
2:B:73:LYS:NZ	2:B:146:TYR:OH	2.24	0.67
1:A:244:ILE:HD13	1:A:267:ALA:HB2	1.76	0.66
1:A:100:LEU:O	1:A:318:TYR:HB3	1.96	0.66
1:A:181:TYR:CE2	1:A:183:TYR:HB2	2.31	0.65
2:B:131:THR:OG1	2:B:143:ARG:HD2	1.97	0.65
2:B:244:ILE:HG23	2:B:429:LEU:HB2	1.79	0.65

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:D:12:DT:H3	4:R:18:A:H62	1.44	0.64
2:B:120:LEU:HD21	2:B:124:PHE:HD2	1.62	0.64
1:A:438:GLU:HB3	1:A:440:PHE:HE1	1.63	0.63
1:A:335:GLY:N	1:A:512:GLN:HE22	1.96	0.63
4:R:17:C:HO2'	4:R:19:A:H2	1.46	0.63
1:A:474:ASN:OD1	1:A:474:ASN:N	2.32	0.62
1:A:474:ASN:O	1:A:478:GLU:HG2	2.00	0.62
2:B:63:ILE:HD12	2:B:65:LYS:H	1.65	0.61
1:A:282:LEU:HD21	1:A:296:THR:HG23	1.82	0.61
4:R:6:G:H2'	4:R:7:G:H8	1.66	0.60
1:A:104:LYS:HG2	1:A:192:ASP:HA	1.84	0.60
2:B:63:ILE:HG21	2:B:74:LEU:HD12	1.83	0.60
1:A:395:LYS:HD2	1:A:414:TRP:CZ2	2.38	0.59
2:B:21:VAL:HB	2:B:59:PRO:HD3	1.84	0.59
1:A:438:GLU:HB3	1:A:440:PHE:CE1	2.37	0.59
3:D:11:DT:H3	4:R:20:A:N6	1.97	0.59
1:A:549:ASP:O	1:A:553:SER:OG	2.15	0.58
1:A:188:TYR:HB3	5:A:601:EFZ:F3	1.94	0.58
1:A:395:LYS:HD2	1:A:414:TRP:CH2	2.39	0.58
4:R:20:A:H5'	4:R:20:A:N3	2.19	0.57
2:B:53:GLU:OE2	2:B:53:GLU:N	2.30	0.57
1:A:342:TYR:HB3	1:A:348:ASN:HA	1.87	0.57
1:A:458:VAL:HG13	1:A:548:VAL:HG12	1.87	0.56
1:A:474:ASN:ND2	4:R:25:A:O2'	2.38	0.56
1:A:449:GLU:HG2	1:A:450:THR:HG23	1.87	0.56
4:R:7:G:C8	4:R:7:G:H5"	2.41	0.56
1:A:536:VAL:HG11	2:B:258:GLN:HE21	1.71	0.55
1:A:149:LEU:HB3	1:A:156:SER:OG	2.06	0.55
1:A:329:ILE:O	1:A:392:PRO:HD3	2.07	0.55
1:A:337:TRP:HH2	1:A:423:VAL:HG21	1.71	0.55
1:A:408:ALA:HB1	2:B:364:ASP:HB3	1.87	0.55
2:B:48:SER:HG	2:B:147:ASN:HD21	1.54	0.55
2:B:56:TYR:HE2	2:B:126:LYS:HD2	1.72	0.54
2:B:30:LYS:HD2	2:B:62:ALA:O	2.07	0.54
1:A:358:ARG:NH1	2:B:394:GLN:OE1	2.40	0.54
3:D:17:DT:H2'	3:D:18:DT:C6	2.42	0.54
1:A:188:TYR:CG	5:A:601:EFZ:H112	2.42	0.54
1:A:337:TRP:CZ3	1:A:368:LEU:HD13	2.42	0.54
1:A:483:HIS:CE1	1:A:524:GLN:OE1	2.61	0.54
2:B:180:ILE:HD13	2:B:189:VAL:HG12	1.90	0.53
1:A:206:ARG:NH2	1:A:216:THR:O	2.37	0.53

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:209:LEU:HD22	2:B:214:LEU:HD12	1.90	0.53
2:B:235:HIS:O	2:B:238:LYS:N	2.37	0.53
1:A:78:ARG:NH1	4:R:4:A:OP2	2.41	0.53
1:A:208:HIS:NE2	1:A:212:TRP:HZ3	2.07	0.53
2:B:362:THR:HA	2:B:366:LYS:HD2	1.91	0.52
1:A:544:GLY:HA2	2:B:286:THR:HG22	1.90	0.52
1:A:188:TYR:CD2	5:A:601:EFZ:H112	2.44	0.52
1:A:503:LEU:HD13	1:A:533:LEU:HG	1.91	0.52
2:B:154:LYS:O	2:B:157:PRO:HD2	2.09	0.52
4:R:26:A:H2'	4:R:27:A:O4'	2.09	0.52
1:A:441:TYR:O	1:A:548:VAL:HG11	2.10	0.52
1:A:175:ASN:OD1	1:A:201:LYS:HE3	2.10	0.52
2:B:215:THR:HG22	2:B:216:THR:H	1.75	0.52
3:D:8:DC:H2'	3:D:9:DT:H6	1.74	0.52
1:A:452:LEU:HD22	1:A:469:LEU:O	2.10	0.51
2:B:395:LYS:HG3	2:B:416:PHE:CE2	2.46	0.51
2:B:106:VAL:HG22	2:B:190:GLY:HA2	1.94	0.50
1:A:447:SER:OG	1:A:450:THR:N	2.44	0.50
1:A:318:TYR:O	1:A:349:LEU:HD21	2.12	0.50
3:D:15:DT:H2"	3:D:16:DA:O5'	2.12	0.50
2:B:286:THR:O	2:B:286:THR:OG1	2.27	0.50
4:R:7:G:H8	4:R:7:G:H5"	1.75	0.49
1:A:406:TRP:CZ3	2:B:418:ASN:HA	2.47	0.49
3:D:8:DC:H2'	3:D:9:DT:C6	2.47	0.49
1:A:480:GLN:O	1:A:483:HIS:HB3	2.12	0.49
2:B:317:VAL:HG12	2:B:347:LYS:HB3	1.93	0.49
1:A:337:TRP:CH2	1:A:368:LEU:HB2	2.48	0.49
1:A:88:TRP:HA	1:A:88:TRP:CE3	2.46	0.49
1:A:533:LEU:HD12	1:A:534:ALA:H	1.77	0.49
1:A:391:LEU:HD12	1:A:414:TRP:CD2	2.47	0.49
2:B:376:THR:O	2:B:380:ILE:HG13	2.13	0.49
4:R:25:A:OP2	4:R:25:A:H8	1.96	0.49
1:A:431:LYS:HG3	1:A:432:GLU:HG2	1.95	0.49
1:A:337:TRP:CH2	1:A:423:VAL:HG21	2.48	0.48
1:A:418:ASN:O	1:A:420:PRO:HD3	2.13	0.48
1:A:9:PRO:HA	1:A:121:ASP:OD1	2.13	0.48
2:B:87:PHE:O	2:B:91:GLN:HB3	2.13	0.48
1:A:548:VAL:O	1:A:552:VAL:HG22	2.13	0.48
1:A:125:ARG:HD3	1:A:147:ASN:HA	1.96	0.48
2:B:254:VAL:HG21	2:B:287:LYS:HB2	1.94	0.48
1:A:455:ALA:HB1	1:A:481:ALA:HB2	1.94	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:218:ASP:O	1:A:222:GLN:HG3	2.14	0.48
2:B:101:LYS:O	2:B:236:PRO:HB2	2.12	0.47
2:B:54:ASN:O	2:B:143:ARG:NH2	2.47	0.47
2:B:100:LEU:HG	2:B:381:VAL:HG13	1.96	0.47
1:A:181:TYR:HD1	2:B:138:GLU:O	1.98	0.47
1:A:419:THR:O	1:A:419:THR:OG1	2.28	0.47
1:A:473:THR:HG23	1:A:476:LYS:H	1.80	0.47
2:B:254:VAL:HG12	2:B:258:GLN:OE1	2.14	0.47
4:R:6:G:H2'	4:R:7:G:C8	2.48	0.47
1:A:90:VAL:HG21	1:A:158:ALA:HA	1.97	0.46
1:A:275:LYS:HD3	1:A:332:GLN:HB3	1.97	0.46
1:A:391:LEU:HD12	1:A:414:TRP:CE3	2.50	0.46
1:A:340:GLN:HG3	1:A:351:THR:HG22	1.97	0.46
3:D:14:DT:H2'	3:D:15:DT:C6	2.51	0.46
1:A:319:TYR:OH	1:A:385:LYS:HD3	2.16	0.46
1:A:325:LEU:HD12	1:A:385:LYS:HE2	1.97	0.46
1:A:417:VAL:HG22	1:A:419:THR:HG23	1.97	0.46
2:B:257:ILE:HG22	2:B:283:LEU:HD11	1.98	0.46
2:B:30:LYS:HD2	2:B:62:ALA:HB3	1.98	0.46
1:A:298:GLU:H	1:A:298:GLU:CD	2.18	0.45
2:B:29:GLU:OE2	2:B:71:TRP:HZ2	1.99	0.45
2:B:334:GLN:H	2:B:334:GLN:HG3	1.32	0.45
1:A:443:ASP:OD1	1:A:444:GLY:N	2.47	0.45
3:D:15:DT:C2	3:D:16:DA:C8	3.04	0.45
1:A:473:THR:O	1:A:477:THR:HG23	2.17	0.45
2:B:239:TRP:CE3	2:B:382:ILE:HD11	2.52	0.45
3:D:12:DT:C2	4:R:18:A:N6	2.83	0.45
2:B:94:ILE:HG12	2:B:161:GLN:OE1	2.17	0.45
1:A:242:GLN:CD	1:A:243:PRO:HD2	2.37	0.45
1:A:379:SER:OG	1:A:387:PRO:HD3	2.17	0.45
1:A:92:LEU:HD12	1:A:93:GLY:N	2.30	0.44
2:B:178:ILE:HD12	2:B:191:SER:HA	1.99	0.44
1:A:54:ASN:HB3	1:A:143:ARG:HH21	1.82	0.44
1:A:307:ARG:O	1:A:311:LYS:HG3	2.17	0.44
1:A:336:GLN:C	1:A:337:TRP:CD1	2.91	0.44
2:B:329:ILE:HD11	2:B:375:ILE:HD12	1.99	0.44
3:D:18:DT:H2'	3:D:19:DG:C8	2.51	0.44
2:B:27:THR:O	2:B:31:ILE:HG13	2.18	0.44
2:B:382:ILE:HG22	2:B:383:TRP:CD2	2.53	0.44
2:B:332:GLN:OE1	2:B:424:LYS:HB2	2.17	0.44
1:A:328:GLU:HG2	1:A:330:GLN:HE21	1.82	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
7:D:102:GOL:H31	4:R:8:C:H41	1.82	0.44
2:B:185:ASP:HB2	2:B:409:THR:HG21	1.99	0.43
2:B:317:VAL:CG1	2:B:347:LYS:HB3	2.48	0.43
2:B:282:LEU:HD21	2:B:296:THR:HG23	2.00	0.43
1:A:7:THR:HG22	1:A:119:PRO:O	2.18	0.43
1:A:17:ASP:O	1:A:83:LYS:HD3	2.18	0.43
1:A:515:SER:OG	1:A:518:VAL:HG23	2.18	0.43
1:A:239:TRP:CH2	1:A:270:ILE:HD13	2.53	0.43
1:A:473:THR:OG1	1:A:474:ASN:OD1	2.33	0.43
2:B:169:GLU:N	2:B:170:PRO:HD2	2.33	0.43
1:A:337:TRP:HZ3	1:A:368:LEU:HD13	1.82	0.43
2:B:282:LEU:CD2	2:B:296:THR:HG23	2.48	0.43
1:A:81:ASN:OD1	1:A:154:LYS:N	2.41	0.43
1:A:168:LEU:O	1:A:172:ARG:HG3	2.19	0.43
1:A:328:GLU:HG2	1:A:330:GLN:NE2	2.33	0.43
2:B:92:LEU:HB2	2:B:158:ALA:HB1	2.00	0.43
2:B:178:ILE:HG23	2:B:190:GLY:O	2.18	0.43
2:B:295:LEU:HB2	2:B:300:GLU:HG2	2.01	0.43
2:B:422:LEU:HD23	2:B:422:LEU:HA	1.87	0.43
1:A:195:ILE:O	1:A:199:ARG:HG3	2.18	0.43
2:B:167:ILE:HG23	2:B:212:TRP:CE3	2.53	0.43
2:B:235:HIS:HB3	2:B:239:TRP:NE1	2.34	0.43
4:R:17:C:O2'	4:R:19:A:H2	2.00	0.43
1:A:49:LYS:HG3	1:A:144:TYR:CE1	2.54	0.43
1:A:175:ASN:N	1:A:176:PRO:HD3	2.34	0.43
1:A:27:THR:HB	1:A:30:LYS:H	1.84	0.43
1:A:354:TYR:CD2	1:A:371:ALA:HB2	2.54	0.42
2:B:382:ILE:HG22	2:B:383:TRP:CE2	2.54	0.42
1:A:456:GLY:HA2	1:A:484:LEU:HD12	2.00	0.42
1:A:517:LEU:H	1:A:517:LEU:HG	1.59	0.42
1:A:519:ASN:CA	1:A:522:ILE:HD12	2.36	0.42
2:B:202:ILE:HG13	2:B:203:GLU:N	2.33	0.42
1:A:533:LEU:HD12	1:A:534:ALA:N	2.33	0.42
1:A:210:LEU:HD12	1:A:210:LEU:HA	1.65	0.42
1:A:216:THR:HB	1:A:217:PRO:HD2	2.02	0.42
1:A:354:TYR:OH	1:A:370:GLU:HB3	2.19	0.42
1:A:172:ARG:HH11	1:A:172:ARG:HD2	1.71	0.42
2:B:96:HIS:ND1	2:B:97:PRO:HD2	2.34	0.42
2:B:163:SER:O	2:B:167:ILE:HG13	2.19	0.42
2:B:128:THR:OG1	2:B:146:TYR:HB2	2.20	0.42
2:B:171:PHE:CZ	2:B:205:LEU:HD11	2.54	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:113:ASP:OD1	1:A:113:ASP:N	2.50	0.42
2:B:282:LEU:HD21	2:B:296:THR:H	1.85	0.42
1:A:10:VAL:HG11	1:A:153:TRP:CH2	2.55	0.41
1:A:110:ASP:OD1	1:A:110:ASP:N	2.35	0.41
2:B:268:SER:HA	2:B:271:TYR:O	2.19	0.41
1:A:332:GLN:HG3	1:A:338:THR:HB	2.02	0.41
1:A:244:ILE:CD1	1:A:267:ALA:HB2	2.48	0.41
2:B:109:LEU:HD23	2:B:109:LEU:HA	1.91	0.41
1:A:366:LYS:HB3	1:A:366:LYS:HE2	1.79	0.41
4:R:12:A:H2'	4:R:13:A:C8	2.55	0.41
1:A:111:VAL:HG11	1:A:214:LEU:HD22	2.02	0.41
1:A:320:ASP:OD2	1:A:323:LYS:HD2	2.20	0.41
2:B:332:GLN:OE1	2:B:424:LYS:HE3	2.20	0.41
1:A:10:VAL:HG12	1:A:124:PHE:CD1	2.55	0.41
1:A:169:GLU:HB3	1:A:170:PRO:HD3	2.02	0.41
1:A:437:ALA:HB3	1:A:494:ASN:ND2	2.35	0.41
1:A:270:ILE:HD12	1:A:270:ILE:HA	1.73	0.41
1:A:325:LEU:C	1:A:326:ILE:HD12	2.41	0.41
1:A:473:THR:HG23	1:A:475:GLN:HG2	2.03	0.41
1:A:497:THR:O	1:A:535:TRP:HA	2.21	0.41
2:B:24:TRP:CG	2:B:25:PRO:HD2	2.56	0.41
2:B:319:TYR:CD1	2:B:383:TRP:CD1	3.08	0.41
2:B:149:LEU:HA	2:B:150:PRO:HD3	1.91	0.41
1:A:297:GLU:HA	1:A:300:GLU:HB2	2.02	0.40
1:A:458:VAL:HG13	1:A:548:VAL:CG1	2.51	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	540/558 (97%)	519 (96%)	21 (4%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
2	B	398/441 (90%)	369 (93%)	29 (7%)	0	100 100
All	All	938/999 (94%)	888 (95%)	50 (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	460/496 (93%)	443 (96%)	17 (4%)	34 62
2	B	341/399 (86%)	331 (97%)	10 (3%)	42 68
All	All	801/895 (90%)	774 (97%)	27 (3%)	37 64

All (27) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	27	THR
1	A	156	SER
1	A	161	GLN
1	A	276	VAL
1	A	284	ARG
1	A	292	VAL
1	A	293	ILE
1	A	302	GLU
1	A	315	HIS
1	A	338	THR
1	A	346	PHE
1	A	362	THR
1	A	419	THR
1	A	447	SER
1	A	474	ASN
1	A	515	SER
1	A	517	LEU
2	B	74	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	116	PHE
2	B	139	THR
2	B	150	PRO
2	B	214	LEU
2	B	286	THR
2	B	334	GLN
2	B	362	THR
2	B	410	TRP
2	B	414	TRP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	242	GLN
1	A	265	ASN
1	A	483	HIS
1	A	512	GLN
2	B	137	ASN

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
4	R	22/25 (88%)	6 (27%)	0

All (6) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
4	R	6	G
4	R	7	G
4	R	19	A
4	R	20	A
4	R	25	A
4	R	27	A

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	0.50	0	9,14,17	1.59	3 (33%)

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	-	0/3/15/16	0/1/1/1

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	R	5	3DR	O4'-C4'-C3'	2.80	107.85	103.73
4	R	5	3DR	C1'-C2'-C3'	2.48	106.00	103.20
4	R	5	3DR	C1'-O4'-C4'	2.02	111.73	108.48

There are no chirality outliers.

There are no torsion outliers.

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 5 ligands modelled in this entry, 2 are monoatomic - leaving 3 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	$\#\|Z\| > 2$
7	GOL	D	102	-	5,5,5	0.29	0	5,5,5	0.57	0
5	EFZ	A	601	-	23,23,23	5.94	11 (47%)	36,36,36	5.24	23 (63%)
7	GOL	R	101	-	5,5,5	0.43	0	5,5,5	0.60	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
7	GOL	D	102	-	-	2/4/4/4	-
5	EFZ	A	601	-	-	10/10/32/32	0/3/3/3
7	GOL	R	101	-	-	2/4/4/4	-

All (11) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	601	EFZ	C1-C6	19.56	1.62	1.40
5	A	601	EFZ	C12-C11	11.43	1.92	1.48
5	A	601	EFZ	C5-C4	-7.76	1.25	1.38
5	A	601	EFZ	C2-C1	7.28	1.52	1.39
5	A	601	EFZ	C7-C6	7.03	1.60	1.51
5	A	601	EFZ	C3-C4	-6.71	1.25	1.38
5	A	601	EFZ	C10-C9	5.79	1.65	1.46
5	A	601	EFZ	O2-C7	-5.43	1.36	1.45
5	A	601	EFZ	C7-C8	2.82	1.55	1.47
5	A	601	EFZ	O2-C14	-2.15	1.33	1.36
5	A	601	EFZ	C11-C10	-2.07	1.34	1.48

All (23) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	A	601	EFZ	C2-C1-C6	-16.32	103.83	119.88

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	A	601	EFZ	C2-C1-N	11.33	139.21	119.84
5	A	601	EFZ	C12-C10-C11	11.02	90.46	59.23
5	A	601	EFZ	O2-C7-C6	-10.51	104.22	111.64
5	A	601	EFZ	C7-O2-C14	10.05	139.71	121.44
5	A	601	EFZ	C11-C12-C10	-6.47	44.46	60.38
5	A	601	EFZ	C12-C11-C10	-6.21	45.08	60.38
5	A	601	EFZ	C3-C4-CL	-4.99	111.55	119.35
5	A	601	EFZ	F1-C13-C7	-4.11	106.77	111.79
5	A	601	EFZ	C2-C3-C4	3.63	123.07	119.24
5	A	601	EFZ	C3-C4-C5	3.60	126.31	121.53
5	A	601	EFZ	C5-C6-C1	3.58	123.24	118.72
5	A	601	EFZ	C13-C7-C8	-3.55	103.18	108.89
5	A	601	EFZ	F2-C13-C7	3.40	115.95	111.79
5	A	601	EFZ	O2-C14-O1	3.34	123.07	118.02
5	A	601	EFZ	C5-C6-C7	-3.26	118.92	122.71
5	A	601	EFZ	F3-C13-C7	-2.67	108.52	111.79
5	A	601	EFZ	C5-C4-CL	2.39	122.14	119.15
5	A	601	EFZ	C6-C1-N	-2.32	116.96	119.88
5	A	601	EFZ	O2-C7-C8	2.25	113.51	108.32
5	A	601	EFZ	C12-C10-C9	-2.23	113.02	119.06
5	A	601	EFZ	O2-C7-C13	2.20	108.83	104.76
5	A	601	EFZ	C13-C7-C6	2.04	114.09	110.84

There are no chirality outliers.

All (14) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	601	EFZ	F1-C13-C7-O2
5	A	601	EFZ	F2-C13-C7-O2
5	A	601	EFZ	F3-C13-C7-O2
5	A	601	EFZ	F1-C13-C7-C6
5	A	601	EFZ	F2-C13-C7-C6
5	A	601	EFZ	F3-C13-C7-C6
5	A	601	EFZ	F1-C13-C7-C8
5	A	601	EFZ	F2-C13-C7-C8
5	A	601	EFZ	F3-C13-C7-C8
7	R	101	GOL	O1-C1-C2-C3
7	D	102	GOL	O1-C1-C2-C3
7	D	102	GOL	O1-C1-C2-O2
7	R	101	GOL	O1-C1-C2-O2
5	A	601	EFZ	C7-C8-C9-C10

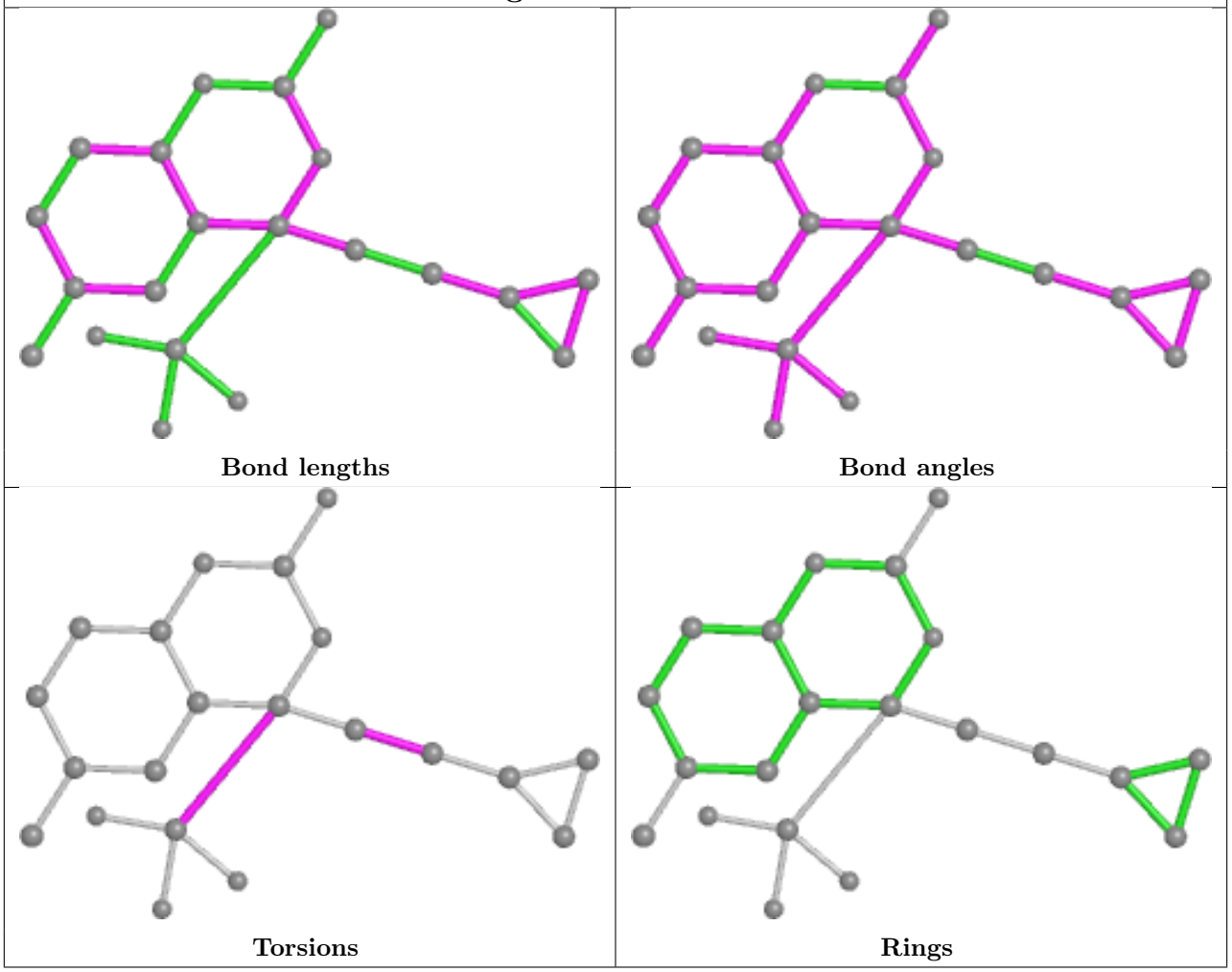
There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	D	102	GOL	1	0
5	A	601	EFZ	3	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

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, 95^{th} 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%)	0.22	16 (2%) 51 50	75, 112, 156, 184	0
2	B	406/441 (92%)	0.34	24 (5%) 22 21	76, 114, 182, 206	0
3	D	23/23 (100%)	-0.19	0 100 100	86, 110, 181, 186	0
4	R	24/25 (96%)	-0.27	0 100 100	97, 117, 162, 178	0
All	All	996/1047 (95%)	0.25	40 (4%) 38 35	75, 113, 173, 206	0

All (40) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	180	ILE	5.1
2	B	109	LEU	4.7
2	B	202	ILE	4.5
2	B	188	TYR	4.4
1	A	482	ILE	3.9
2	B	189	VAL	3.6
2	B	108	VAL	3.5
2	B	187	LEU	3.3
2	B	214	LEU	3.3
2	B	178	ILE	3.2
2	B	179	VAL	3.2
2	B	205	LEU	3.0
1	A	181	TYR	2.9
2	B	111	VAL	2.9
2	B	190	GLY	2.9
2	B	168	LEU	2.7
1	A	455	ALA	2.7
2	B	94	ILE	2.6
1	A	106	VAL	2.6
2	B	107	THR	2.6
2	B	209	LEU	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	B	200	THR	2.5
2	B	234	LEU	2.5
1	A	479	LEU	2.5
2	B	325	LEU	2.4
2	B	183	TYR	2.4
1	A	505	ILE	2.4
2	B	181	TYR	2.3
1	A	500	GLN	2.3
1	A	552	VAL	2.3
2	B	410	TRP	2.2
1	A	506	ILE	2.2
1	A	525	LEU	2.1
2	B	184	MET	2.1
1	A	73	LYS	2.1
1	A	90	VAL	2.1
1	A	376	THR	2.1
1	A	353	LYS	2.0
1	A	486	LEU	2.0
1	A	372	VAL	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, 95^{th} 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	0.92	0.16	117,153,170,173	0

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

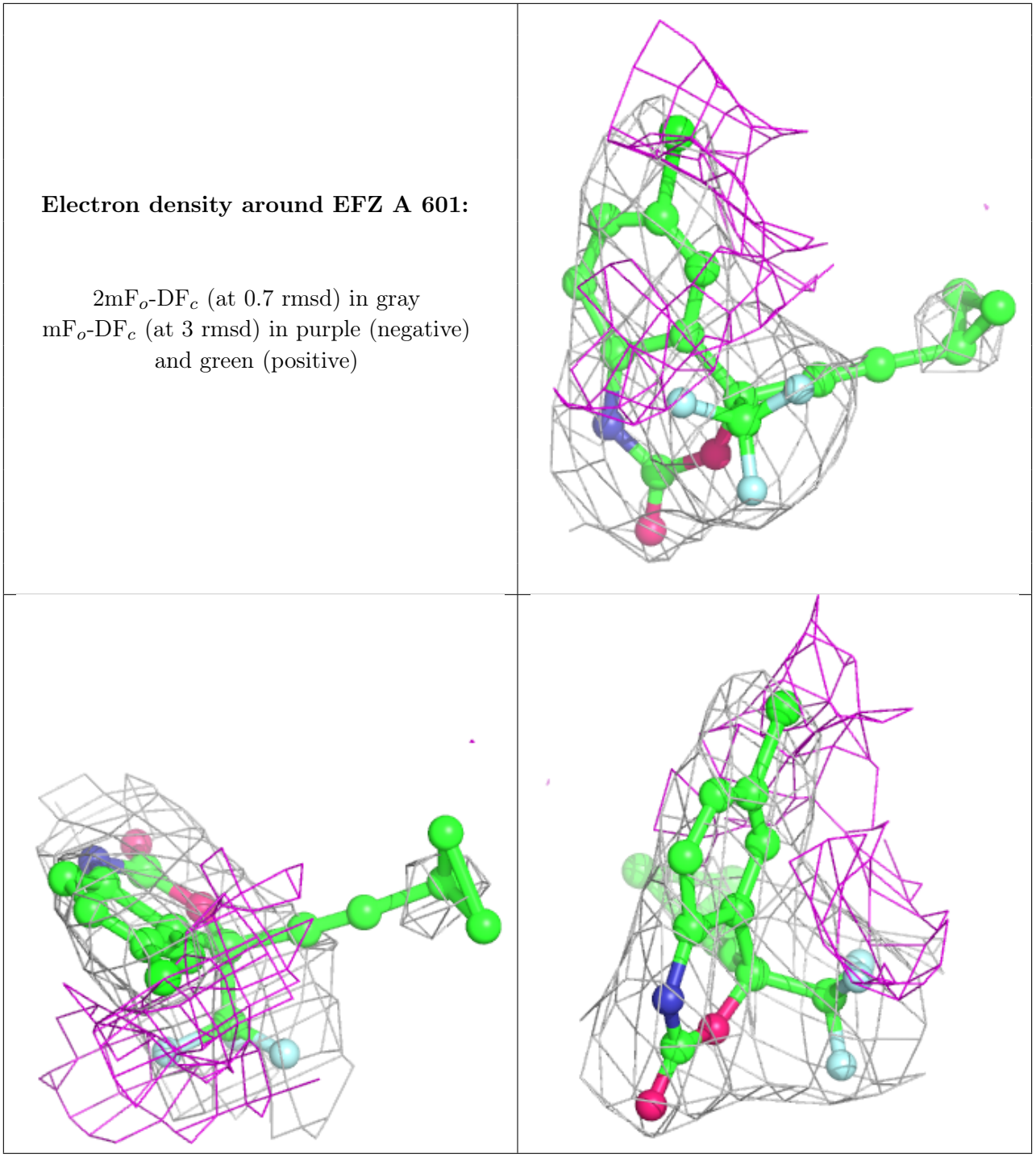
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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	R	101	6/6	0.57	0.32	119,125,126,131	0
6	CA	D	101	1/1	0.82	0.23	133,133,133,133	0
7	GOL	D	102	6/6	0.88	0.26	106,122,127,131	0
6	CA	A	602	1/1	0.92	0.17	140,140,140,140	0
5	EFZ	A	601	21/21	0.96	0.53	76,91,97,100	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

There are no such residues in this entry.