Full wwPDB X-ray Structure Validation Report ⓘ

Jun 24, 2024 – 03:27 PM EDT

PDB ID	:	7AHX
Title	:	HIV-1 REVERSE TRANSCRIPTASE COMPLEX WITH DNA AND D-ASPARTATE TENOFOVIR WITH BOUND MANGANESE
Authors	:	Gu, W.; Martinez, S.E.; Nguyen, H.; Xu, H.; Herdewijn, P.; de Jonghe, S.; Das, K.
Deposited on	:	2020-09-25
Resolution	:	2.73 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

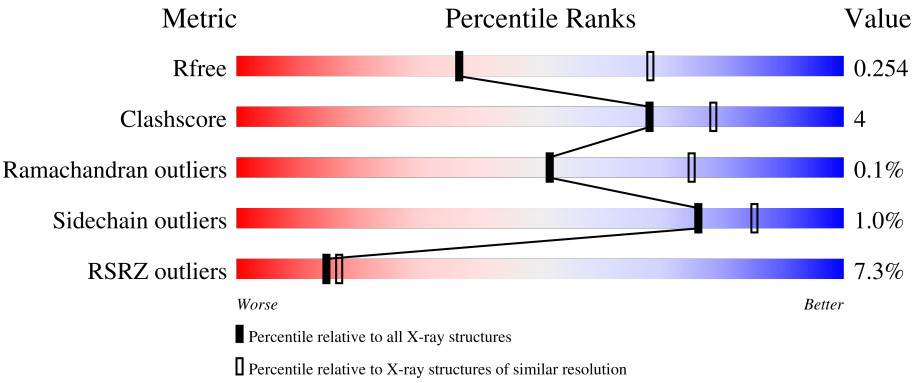
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (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 2.73 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.



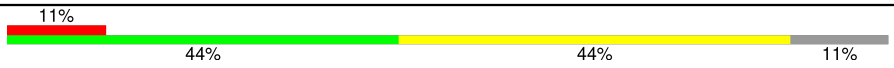
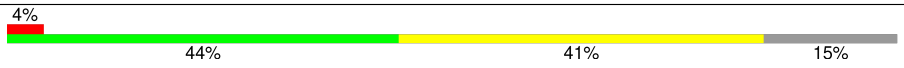
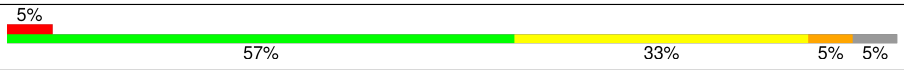
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1271 (2.76-2.72)
Clashscore	141614	1322 (2.76-2.72)
Ramachandran outliers	138981	1297 (2.76-2.72)
Sidechain outliers	138945	1298 (2.76-2.72)
RSRZ outliers	127900	1243 (2.76-2.72)

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	556	6% 89% 11%
1	C	556	12% 89% 11%
2	B	428	2% 88% 7% •
2	D	428	7% 86% 10% • •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	E	27	
3	T	27	
4	F	21	
4	P	21	

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 17778 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 Gag-Pol polyprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	555	Total	C	N	O	S	0	0	0
			4513	2922	751	832	8			
1	C	556	Total	C	N	O	S	0	0	0
			4521	2927	752	833	9			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	MET	-	initiating methionine	UNP P03366
A	0	VAL	-	expression tag	UNP P03366
A	258	CYS	GLN	engineered mutation	UNP P03366
A	280	SER	CYS	engineered mutation	UNP P03366
A	498	ASN	ASP	engineered mutation	UNP P03366
C	-1	MET	-	initiating methionine	UNP P03366
C	0	VAL	-	expression tag	UNP P03366
C	258	CYS	GLN	engineered mutation	UNP P03366
C	280	SER	CYS	engineered mutation	UNP P03366
C	498	ASN	ASP	engineered mutation	UNP P03366

- Molecule 2 is a protein called Gag-Pol polyprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	411	Total	C	N	O	S	0	0	0
			3401	2215	563	616	7			
2	D	412	Total	C	N	O	S	0	0	0
			3400	2212	563	619	6			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	280	SER	CYS	engineered mutation	UNP P03366
D	280	SER	CYS	engineered mutation	UNP P03366

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	T	23	Total	C	N	O	P	0	0	0
			477	223	95	136	23			
3	E	24	Total	C	N	O	P	0	0	0
			494	233	97	141	23			

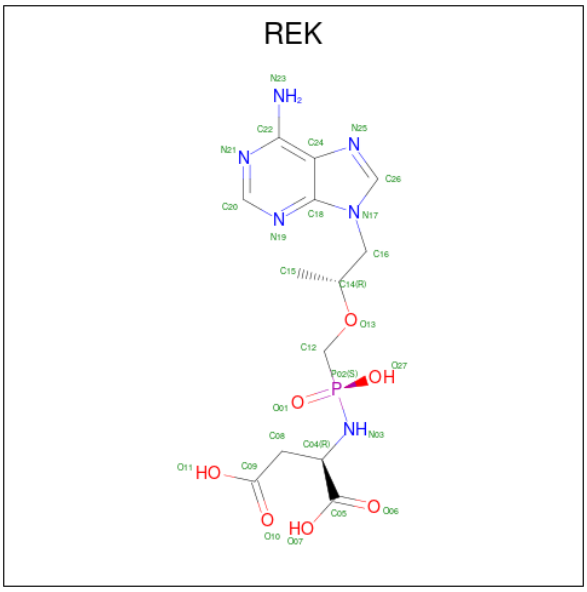
- Molecule 4 is a DNA chain called DNA (5'-D(*CP*AP*GP*TP*CP*CP*CP*TP*GP*TP*TP*CP*GP*GP*(MRG)*CP*GP*CP*CP*(DDG))-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	P	20	Total	C	N	O	P	0	0	0
			403	192	72	120	19			
4	F	20	Total	C	N	O	P	0	0	0
			403	192	72	120	19			

- Molecule 5 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	2	Total	Mn	0	0
			2	2		
5	C	2	Total	Mn	0	0
			2	2		

- Molecule 6 is D-Aspartate Tenofovir (three-letter code: REK) (formula: $C_{13}H_{19}N_6O_7P$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	A	1	Total	C	N	O	P	0	0
			27	13	6	7	1		
6	C	1	Total	C	N	O	P	0	0
			27	13	6	7	1		

- Molecule 7 is water.

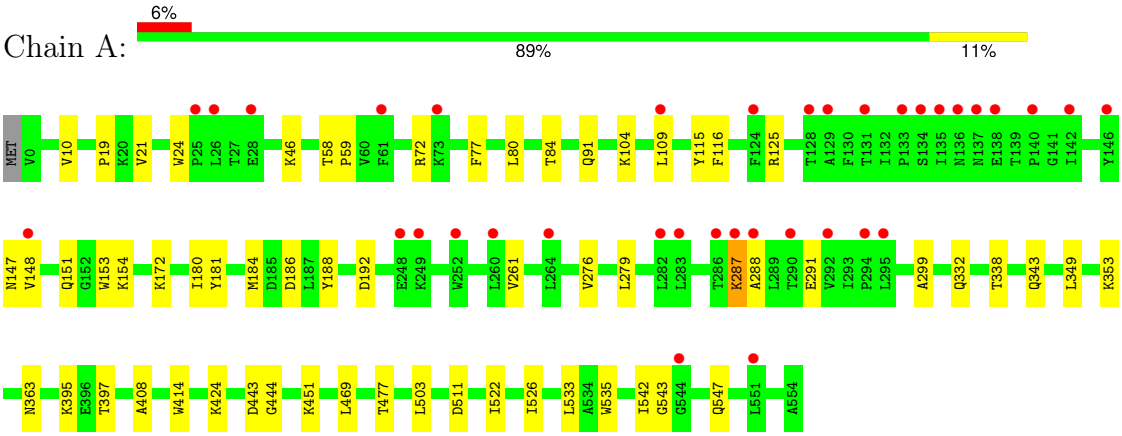
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	36	Total	O	0	0
			36	36		
7	B	31	Total	O	0	0
			31	31		
7	C	16	Total	O	0	0
			16	16		
7	D	15	Total	O	0	0
			15	15		
7	T	3	Total	O	0	0
			3	3		
7	P	4	Total	O	0	0
			4	4		
7	E	1	Total	O	0	0
			1	1		
7	F	2	Total	O	0	0
			2	2		

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.

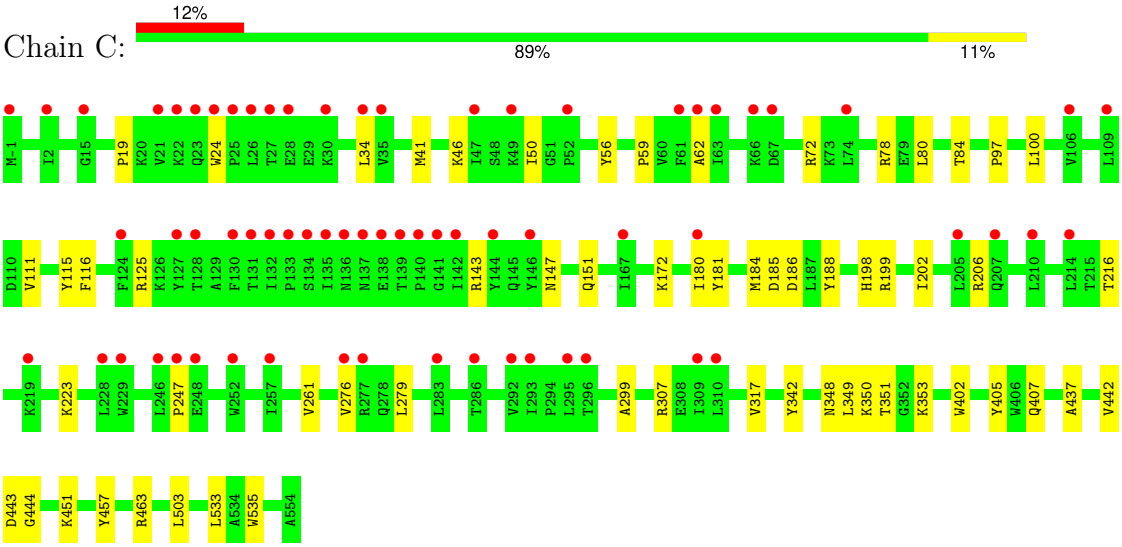
• Molecule 1: Gag-Pol polyprotein

Chain A:



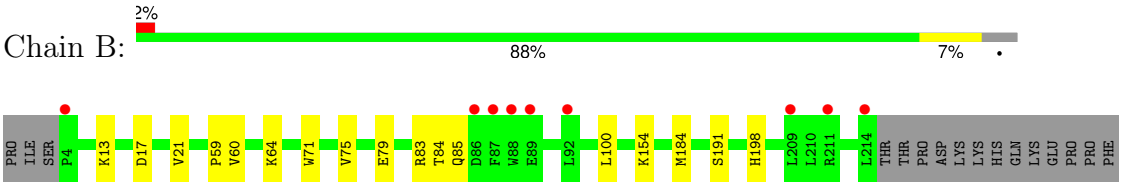
• Molecule 1: Gag-Pol polyprotein

Chain C:



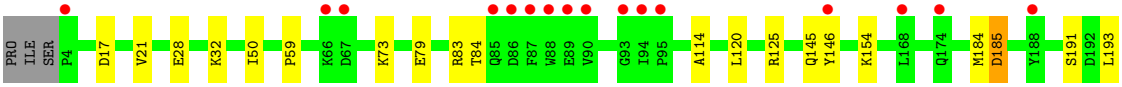
• Molecule 2: Gag-Pol polyprotein

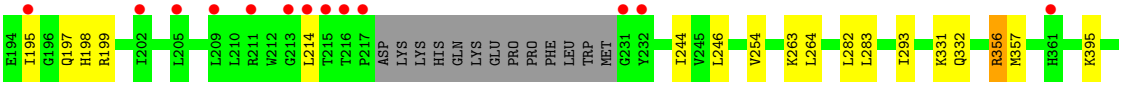
Chain B:

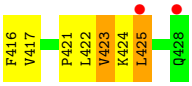


• Molecule 2: Gag-Pol polyprotein

Chain D:

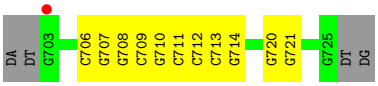






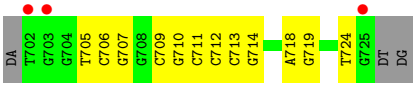
• Molecule 3: DNA (5'-D(*TP*GP*GP*TP*CP*GP*GP*CP*GP*CP*CP*CP*GP*AP*AP*CP *AP*GP*GP*GP*AP*CP*TP*G)-3')

Chain T:



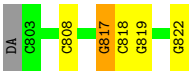
• Molecule 3: DNA (5'-D(*TP*GP*GP*TP*CP*GP*GP*CP*GP*CP*CP*CP*GP*AP*AP*CP *AP*GP*GP*GP*AP*CP*TP*G)-3')

Chain E:



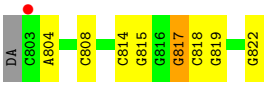
• Molecule 4: DNA (5'-D(*CP*AP*GP*TP*CP*CP*CP*TP*GP*TP*TP*CP*GP*GP*(MRG)* CP*GP*CP*CP*(DDG))-3')

Chain P:



• Molecule 4: DNA (5'-D(*CP*AP*GP*TP*CP*CP*CP*TP*GP*TP*TP*CP*GP*GP*(MRG)* CP*GP*CP*CP*(DDG))-3')

Chain F:


DA
C803
A804
C808
C814
G815
G816
G817
C818
G819
G822

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	89.42Å 132.37Å 138.74Å 90.00° 98.48° 90.00°	Depositor
Resolution (Å)	69.80 – 2.73 69.80 – 2.73	Depositor EDS
% Data completeness (in resolution range)	99.6 (69.80-2.73) 99.6 (69.80-2.73)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.34 (at 2.73Å)	Xtriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, R_{free}	0.214 , 0.254 0.214 , 0.254	Depositor DCC
R_{free} test set	2529 reflections (2.99%)	wwPDB-VP
Wilson B-factor (Å^2)	58.6	Xtriage
Anisotropy	0.114	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 57.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	17778	wwPDB-VP
Average B, all atoms (Å^2)	77.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.90% 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: MRG, DDG, REK, MN

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	0.24	0/4631	0.41	0/6293
1	C	0.23	0/4639	0.41	0/6303
2	B	0.24	0/3499	0.41	0/4752
2	D	0.23	0/3497	0.40	0/4751
3	E	0.49	0/555	0.79	0/856
3	T	0.48	0/536	0.78	0/826
4	F	0.54	0/400	0.92	0/612
4	P	0.54	0/400	0.88	0/612
All	All	0.28	0/18157	0.48	0/25005

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	4513	0	4578	36	0
1	C	4521	0	4587	36	0
2	B	3401	0	3431	18	0
2	D	3400	0	3433	23	0
3	E	494	0	269	8	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	T	477	0	256	8	0
4	F	403	0	224	5	0
4	P	403	0	224	4	0
5	A	2	0	0	0	0
5	C	2	0	0	0	0
6	A	27	0	0	1	0
6	C	27	0	0	0	0
7	A	36	0	0	0	0
7	B	31	0	0	0	0
7	C	16	0	0	1	0
7	D	15	0	0	1	0
7	E	1	0	0	0	0
7	F	2	0	0	0	0
7	P	4	0	0	0	0
7	T	3	0	0	0	0
All	All	17778	0	17002	129	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 4.

All (129) 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:181:TYR:HB2	1:C:188:TYR:HB3	1.74	0.70
1:A:125:ARG:HD3	1:A:147:ASN:HA	1.76	0.67
1:A:181:TYR:HB2	1:A:188:TYR:HB3	1.78	0.66
1:C:172:LYS:HE2	1:C:180:ILE:HB	1.80	0.63
2:D:244:ILE:HD13	2:D:425:LEU:HD11	1.80	0.63
1:A:543:GLY:N	2:B:283:LEU:O	2.33	0.61
1:A:84:THR:HB	1:A:154:LYS:HE2	1.83	0.61
1:A:343:GLN:HG3	1:A:349:LEU:HD11	1.83	0.61
1:A:451:LYS:NZ	4:P:808:DC:OP1	2.34	0.60
3:E:724:DT:H3	4:F:804:DA:H61	1.47	0.60
2:B:79:GLU:HG3	2:B:83:ARG:HE	1.66	0.60
2:D:421:PRO:HB2	2:D:423:VAL:HG22	1.84	0.59
1:A:46:LYS:HD3	1:A:116:PHE:HB3	1.83	0.59
2:B:21:VAL:HB	2:B:59:PRO:HD3	1.83	0.59
1:A:72:ARG:NE	6:A:603:REK:O06	2.35	0.59
2:D:263:LYS:HG3	2:D:425:LEU:HB2	1.85	0.58
1:C:451:LYS:NZ	4:F:808:DC:OP1	2.36	0.57
1:A:288:ALA:HB3	1:A:291:GLU:HB2	1.87	0.56

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:503:LEU:HD22	1:A:535:TRP:HB2	1.88	0.56
2:D:193:LEU:HB3	2:D:197:GLN:HG3	1.89	0.55
1:C:437:ALA:O	7:C:701:HOH:O	2.18	0.54
2:D:50:ILE:HD13	2:D:145:GLN:HB3	1.89	0.54
1:C:402:TRP:O	2:D:331:LYS:NZ	2.31	0.54
2:D:195:ILE:HG13	2:D:199:ARG:HE	1.72	0.54
2:B:332:GLN:HB2	2:B:336:GLN:HB2	1.90	0.53
1:A:443:ASP:OD1	1:A:444:GLY:N	2.40	0.53
2:D:21:VAL:HB	2:D:59:PRO:HD3	1.91	0.53
1:C:503:LEU:HD22	1:C:535:TRP:HB2	1.90	0.53
1:A:21:VAL:HB	1:A:59:PRO:HD3	1.91	0.53
1:C:111:VAL:HB	1:C:185:ASP:HB2	1.91	0.53
1:A:172:LYS:HE2	1:A:180:ILE:HB	1.91	0.53
2:D:191:SER:OG	2:D:198:HIS:ND1	2.36	0.52
1:C:78:ARG:HD2	3:E:705:DT:H5"	1.91	0.51
3:T:711:DC:H2'	3:T:712:DC:C6	2.46	0.51
2:B:390:LYS:NZ	2:B:415:GLU:OE2	2.35	0.51
1:A:469:LEU:HD12	1:A:477:THR:HG22	1.93	0.51
1:A:19:PRO:HD3	1:A:80:LEU:HD13	1.93	0.50
2:B:60:VAL:HG22	2:B:75:VAL:HG22	1.93	0.50
2:D:73:LYS:NZ	2:D:146:TYR:OH	2.32	0.50
2:D:114:ALA:HB2	2:D:214:LEU:HD13	1.93	0.50
1:C:247:PRO:O	1:C:307:ARG:NH2	2.44	0.50
1:C:503:LEU:HD12	1:C:533:LEU:HD13	1.93	0.50
1:C:342:TYR:HB3	1:C:348:ASN:HA	1.92	0.50
2:D:28:GLU:HG2	2:D:32:LYS:HE2	1.94	0.49
1:C:56:TYR:O	1:C:143:ARG:NH2	2.45	0.49
1:A:58:THR:HG21	1:A:77:PHE:CD1	2.48	0.49
1:A:276:VAL:HG23	1:A:353:LYS:HE2	1.94	0.49
1:A:104:LYS:HB2	1:A:192:ASP:HA	1.94	0.49
1:A:115:TYR:HD2	1:A:151:GLN:HG2	1.78	0.49
1:A:116:PHE:HA	1:A:148:VAL:HG21	1.95	0.48
2:D:154:LYS:HG2	2:D:184:MET:SD	2.54	0.48
2:D:254:VAL:HG13	2:D:283:LEU:HD22	1.96	0.48
3:E:718:DA:H1'	3:E:719:DG:C8	2.49	0.48
1:A:24:TRP:HD1	1:A:59:PRO:HB3	1.79	0.47
1:A:261:VAL:HG13	1:A:276:VAL:HG11	1.96	0.47
1:A:542:ILE:HG23	2:B:283:LEU:HD13	1.96	0.47
1:C:115:TYR:HD2	1:C:151:GLN:HG2	1.79	0.47
2:D:185:ASP:N	2:D:185:ASP:OD1	2.46	0.47
3:T:713:DC:H2'	3:T:714:DG:C8	2.50	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:91:GLN:O	3:T:708:DG:H4'	2.14	0.47
2:D:17:ASP:O	2:D:83:ARG:HD3	2.15	0.47
1:C:350:LYS:NZ	1:C:351:THR:O	2.45	0.47
1:C:125:ARG:HD3	1:C:147:ASN:HA	1.96	0.46
2:B:64:LYS:HE3	2:B:71:TRP:CE2	2.51	0.46
1:A:363:ASN:HA	1:A:511:ASP:OD1	2.16	0.46
3:E:711:DC:H2'	3:E:712:DC:C6	2.51	0.46
1:A:115:TYR:CE1	1:A:184:MET:HE1	2.50	0.46
1:C:115:TYR:CD2	1:C:151:GLN:HG2	2.51	0.45
3:E:706:DC:H2'	3:E:707:DG:C8	2.51	0.45
1:A:80:LEU:O	1:A:84:THR:OG1	2.27	0.45
3:T:720:DG:H2"	3:T:721:DG:C8	2.52	0.45
2:D:356:ARG:H	2:D:356:ARG:HG3	1.56	0.44
1:C:405:TYR:CE2	1:C:407:GLN:HB2	2.52	0.44
2:B:191:SER:OG	2:B:198:HIS:ND1	2.42	0.44
2:D:79:GLU:HG3	2:D:83:ARG:HE	1.83	0.44
2:B:365:VAL:HG11	2:B:401:TRP:HB2	1.99	0.44
1:C:50:ILE:H	1:C:50:ILE:HG13	1.65	0.44
1:C:442:VAL:HG12	1:C:457:TYR:HB3	2.00	0.44
4:P:818:DC:H2'	4:P:819:DG:C8	2.53	0.44
2:B:17:ASP:O	2:B:83:ARG:HD3	2.17	0.43
3:E:718:DA:H4'	3:E:719:DG:OP1	2.18	0.43
4:F:818:DC:H2'	4:F:819:DG:C8	2.53	0.43
1:C:276:VAL:HG23	1:C:353:LYS:HE2	2.00	0.43
1:C:443:ASP:OD1	1:C:444:GLY:N	2.49	0.43
1:C:198:HIS:O	1:C:202:ILE:HG12	2.18	0.43
2:D:246:LEU:HD11	2:D:264:LEU:HD21	1.99	0.43
1:A:279:LEU:HD23	1:A:299:ALA:HB1	2.01	0.43
1:C:34:LEU:HD21	1:C:62:ALA:HB2	2.01	0.43
3:E:713:DC:H2'	3:E:714:DG:C8	2.53	0.43
4:F:814:DC:H2"	4:F:815:DG:H8	1.83	0.43
2:B:154:LYS:HG2	2:B:184:MET:SD	2.58	0.43
1:C:199:ARG:HH12	1:C:223:LYS:HB3	1.83	0.43
1:C:24:TRP:HD1	1:C:59:PRO:HB3	1.83	0.43
1:C:41:MET:HB3	1:C:46:LYS:HB2	2.01	0.43
1:A:522:ILE:O	1:A:526:ILE:HG13	2.19	0.42
2:B:422:LEU:H	2:B:422:LEU:HG	1.45	0.42
1:C:279:LEU:HD23	1:C:299:ALA:HB1	2.00	0.42
1:C:46:LYS:HD3	1:C:116:PHE:HB3	2.01	0.42
1:A:397:THR:HG21	1:A:424:LYS:HA	2.00	0.42
1:C:115:TYR:CE1	1:C:184:MET:HE1	2.54	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:457:TYR:HE1	1:C:463:ARG:HG2	1.85	0.42
1:C:97:PRO:HA	1:C:100:LEU:HG	2.02	0.42
1:C:206:ARG:NH2	1:C:216:THR:O	2.51	0.42
2:B:282:LEU:HD21	2:B:295:LEU:HA	2.02	0.42
3:T:706:DC:H2'	3:T:707:DG:C8	2.55	0.42
2:D:282:LEU:HB3	2:D:293:ILE:HG21	2.01	0.41
1:A:10:VAL:HG11	1:A:153:TRP:CZ2	2.55	0.41
2:B:100:LEU:HG	2:B:381:VAL:HG13	2.02	0.41
1:C:80:LEU:O	1:C:84:THR:OG1	2.28	0.41
3:T:709:DC:H2'	3:T:710:DG:C8	2.55	0.41
4:P:817:MRG:H2'	4:P:818:DC:C6	2.56	0.41
1:A:503:LEU:HD12	1:A:533:LEU:HD13	2.02	0.41
1:C:261:VAL:HG13	1:C:276:VAL:HG11	2.02	0.41
2:D:120:LEU:HD23	2:D:125:ARG:HG2	2.01	0.41
2:D:395:LYS:HG3	2:D:416:PHE:CE2	2.55	0.41
1:A:408:ALA:O	2:B:393:ILE:HG13	2.20	0.41
1:A:547:GLN:OE1	1:A:547:GLN:N	2.54	0.41
1:A:395:LYS:NZ	1:A:414:TRP:O	2.54	0.41
2:B:317:VAL:HG22	2:B:347:LYS:HD2	2.03	0.41
4:P:818:DC:H2'	4:P:819:DG:H8	1.84	0.41
2:D:332:GLN:NE2	7:D:506:HOH:O	2.54	0.41
3:E:709:DC:H2'	3:E:710:DG:C8	2.55	0.41
1:C:19:PRO:HG3	1:C:80:LEU:HB2	2.04	0.40
3:T:710:DG:H2'	3:T:711:DC:C6	2.56	0.40
4:F:817:MRG:H2'	4:F:818:DC:C6	2.56	0.40
1:C:317:VAL:HG23	1:C:349:LEU:HD23	2.03	0.40
1:A:287:LYS:H	1:A:287:LYS:HD3	1.87	0.40
2:B:13:LYS:HD3	2:B:85:GLN:HB3	2.03	0.40
3:T:712:DC:H2'	3:T:713:DC:C6	2.57	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	553/556 (100%)	540 (98%)	12 (2%)	1 (0%)	47 69
1	C	554/556 (100%)	535 (97%)	19 (3%)	0	100 100
2	B	407/428 (95%)	389 (96%)	18 (4%)	0	100 100
2	D	408/428 (95%)	390 (96%)	17 (4%)	1 (0%)	47 69
All	All	1922/1968 (98%)	1854 (96%)	66 (3%)	2 (0%)	51 75

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	332	GLN
2	D	423	VAL

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	496/497 (100%)	492 (99%)	4 (1%)	81 89
1	C	497/497 (100%)	495 (100%)	2 (0%)	91 94
2	B	373/390 (96%)	369 (99%)	4 (1%)	73 84
2	D	374/390 (96%)	366 (98%)	8 (2%)	53 72
All	All	1740/1774 (98%)	1722 (99%)	18 (1%)	76 85

All (18) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	109	LEU
1	A	186	ASP
1	A	287	LYS
1	A	338	THR
2	B	84	THR
2	B	229	TRP
2	B	230	MET
2	B	422	LEU
1	C	72	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	186	ASP
2	D	84	THR
2	D	185	ASP
2	D	356	ARG
2	D	357	MET
2	D	417	VAL
2	D	422	LEU
2	D	424	LYS
2	D	425	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
2	D	182	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

4 non-standard protein/DNA/RNA residues 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	$\#\|Z\| > 2$
4	MRG	F	817	4,3	18,24,29	1.13	3 (16%)	19,35,42	0.72	0
4	DDG	P	822	4,3	17,23,24	4.19	10 (58%)	16,33,36	1.95	6 (37%)
4	DDG	F	822	4,3	17,23,24	4.20	10 (58%)	16,33,36	1.98	5 (31%)
4	MRG	P	817	4,3	18,24,29	1.12	3 (16%)	19,35,42	0.69	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	MRG	F	817	4,3	-	0/3/21/27	0/3/3/3
4	DDG	P	822	4,3	-	2/3/18/19	0/3/3/3
4	DDG	F	822	4,3	-	1/3/18/19	0/3/3/3
4	MRG	P	817	4,3	-	0/3/21/27	0/3/3/3

All (26) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	F	822	DDG	C2'-C3'	-9.72	1.27	1.54
4	P	822	DDG	C2'-C3'	-9.69	1.27	1.54
4	F	822	DDG	O4'-C4'	-8.13	1.29	1.44
4	P	822	DDG	O4'-C4'	-8.12	1.29	1.44
4	F	822	DDG	C1'-N9	-6.39	1.31	1.49
4	P	822	DDG	C1'-N9	-6.31	1.32	1.49
4	P	822	DDG	C2-N3	4.80	1.44	1.33
4	F	822	DDG	C2-N3	4.75	1.44	1.33
4	P	822	DDG	O4'-C1'	4.73	1.52	1.42
4	F	822	DDG	O4'-C1'	4.71	1.52	1.42
4	P	822	DDG	C4-N3	3.65	1.46	1.37
4	F	822	DDG	C4-N3	3.62	1.46	1.37
4	F	822	DDG	C2-N2	3.53	1.42	1.34
4	P	822	DDG	C2-N2	3.52	1.42	1.34
4	F	817	MRG	C5-C6	-2.78	1.41	1.47
4	P	817	MRG	C5-C6	-2.77	1.42	1.47
4	P	822	DDG	C5-C4	-2.67	1.36	1.43
4	F	822	DDG	C5-C4	-2.65	1.36	1.43
4	P	817	MRG	C8-N7	-2.49	1.30	1.34
4	F	817	MRG	C8-N7	-2.49	1.30	1.34
4	F	822	DDG	C2-N1	2.26	1.43	1.37
4	P	822	DDG	C2-N1	2.20	1.43	1.37
4	P	822	DDG	C3'-C4'	2.16	1.62	1.52
4	F	822	DDG	C3'-C4'	2.14	1.62	1.52
4	F	817	MRG	C5-C4	-2.04	1.38	1.43
4	P	817	MRG	C5-C4	-2.01	1.38	1.43

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	F	822	DDG	C4'-O4'-C1'	-3.67	106.35	109.81
4	P	822	DDG	C8-N7-C5	3.21	108.02	102.55

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	P	822	DDG	C5-C6-N1	3.14	120.06	114.07
4	F	822	DDG	C8-N7-C5	3.14	107.89	102.55
4	F	822	DDG	C5-C6-N1	3.08	119.95	114.07
4	P	822	DDG	C2-N1-C6	-2.98	119.66	125.11
4	F	822	DDG	C2-N1-C6	-2.94	119.74	125.11
4	P	822	DDG	C4'-O4'-C1'	-2.89	107.08	109.81
4	F	822	DDG	O6-C6-C5	-2.18	119.99	124.32
4	P	822	DDG	O6-C6-C5	-2.13	120.09	124.32
4	P	822	DDG	C3'-C2'-C1'	2.09	105.28	102.87

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	P	822	DDG	O4'-C4'-C5'-O5'
4	P	822	DDG	C3'-C4'-C5'-O5'
4	F	822	DDG	O4'-C4'-C5'-O5'

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	F	817	MRG	1	0
4	P	817	MRG	1	0

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 6 ligands modelled in this entry, 4 are monoatomic - leaving 2 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	REK	C	603	5	22,28,28	1.59	4 (18%)	18,40,40	1.29	2 (11%)
6	REK	A	603	5	22,28,28	1.63	6 (27%)	18,40,40	1.26	2 (11%)

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	REK	C	603	5	-	8/18/23/23	0/2/2/2
6	REK	A	603	5	-	8/18/23/23	0/2/2/2

All (10) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	C	603	REK	C20-N21	2.85	1.39	1.33
6	A	603	REK	C20-N21	2.83	1.39	1.33
6	A	603	REK	P02-O27	-2.56	1.50	1.56
6	C	603	REK	P02-O27	-2.55	1.50	1.56
6	C	603	REK	C04-C05	2.30	1.58	1.52
6	A	603	REK	C04-N03	2.16	1.49	1.47
6	C	603	REK	C16-N17	2.11	1.50	1.48
6	A	603	REK	C04-C05	2.10	1.58	1.52
6	A	603	REK	C08-C09	2.08	1.56	1.51
6	A	603	REK	C16-N17	2.01	1.50	1.48

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	C	603	REK	C24-C22-N23	2.40	123.97	120.31
6	A	603	REK	C24-C22-N23	2.33	123.86	120.31
6	C	603	REK	O27-P02-O01	2.10	117.80	112.36
6	A	603	REK	O27-P02-O01	2.09	117.78	112.36

There are no chirality outliers.

All (16) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	A	603	REK	N03-C04-C05-O07
6	A	603	REK	C05-C04-C08-C09

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
6	A	603	REK	O13-C12-P02-O27
6	C	603	REK	O13-C12-P02-O01
6	C	603	REK	O13-C12-P02-O27
6	C	603	REK	C15-C14-C16-N17
6	A	603	REK	N03-C04-C08-C09
6	A	603	REK	N03-C04-C05-O06
6	A	603	REK	O13-C12-P02-O01
6	A	603	REK	C05-C04-N03-P02
6	C	603	REK	O13-C14-C16-N17
6	A	603	REK	C08-C04-N03-P02
6	C	603	REK	C08-C04-N03-P02
6	C	603	REK	C08-C04-C05-O06
6	C	603	REK	C08-C04-C05-O07
6	C	603	REK	C05-C04-N03-P02

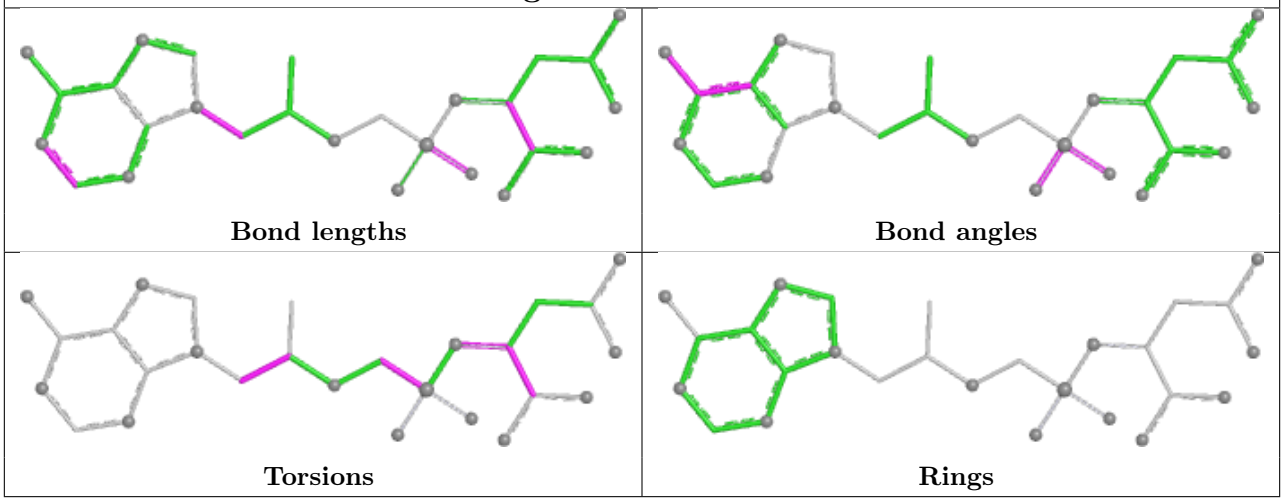
There are no ring outliers.

1 monomer is involved in 1 short contact:

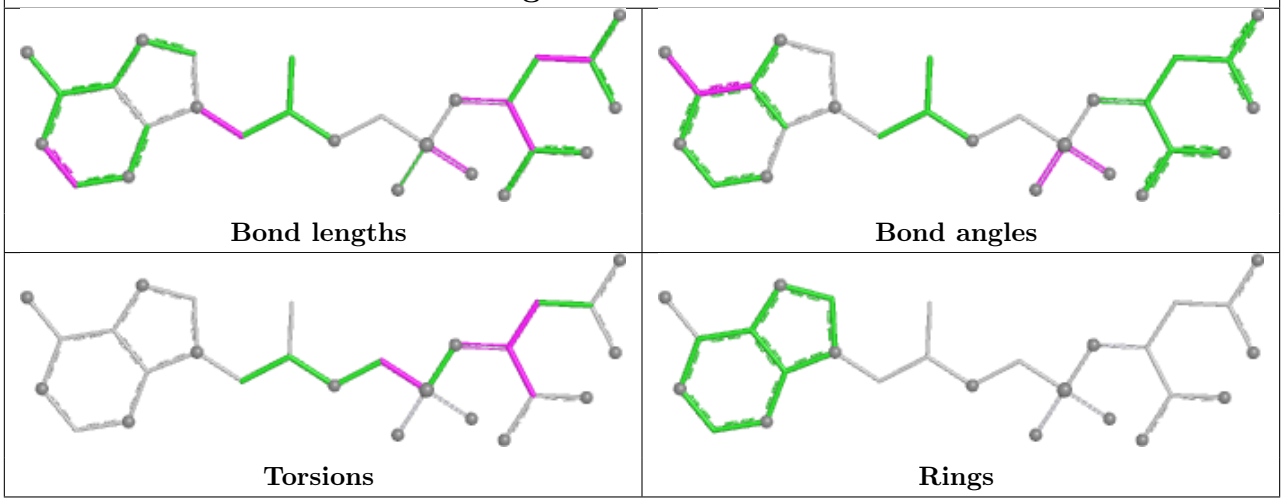
Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	603	REK	1	0

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 REK C 603



Ligand REK A 603



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	555/556 (99%)	0.32	36 (6%) 18 21	26, 63, 127, 178	0
1	C	556/556 (100%)	0.67	67 (12%) 4 4	26, 89, 150, 181	0
2	B	411/428 (96%)	0.17	9 (2%) 62 69	20, 54, 109, 143	0
2	D	412/428 (96%)	0.48	31 (7%) 14 16	28, 71, 124, 152	0
3	E	24/27 (88%)	-0.07	3 (12%) 3 4	71, 114, 190, 196	0
3	T	23/27 (85%)	-0.43	1 (4%) 35 38	53, 100, 150, 193	0
4	F	18/21 (85%)	-0.64	1 (5%) 24 27	59, 100, 161, 175	0
4	P	18/21 (85%)	-0.71	0 100 100	57, 79, 148, 153	0
All	All	2017/2064 (97%)	0.39	148 (7%) 15 17	20, 68, 139, 196	0

All (148) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	138	GLU	9.4
1	C	142	ILE	9.1
2	D	4	PRO	8.1
1	C	133	PRO	7.8
1	C	30	LYS	7.2
1	C	132	ILE	6.9
2	D	214	LEU	6.8
3	E	702	DT	6.7
2	B	214	LEU	6.4
1	C	134	SER	6.1
1	C	136	ASN	5.5
1	C	24	TRP	5.5
1	C	137	ASN	5.4
1	C	293	ILE	5.3
2	D	231	GLY	5.3
1	A	138	GLU	5.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	133	PRO	5.2
1	A	135	ILE	5.2
2	D	361	HIS	5.1
1	C	295	LEU	5.1
1	C	139	THR	5.0
1	C	205	LEU	5.0
2	B	4	PRO	5.0
1	C	52	PRO	4.9
1	C	135	ILE	4.9
2	D	168	LEU	4.9
1	C	214	LEU	4.8
1	C	49	LYS	4.7
1	A	282	LEU	4.7
1	C	74	LEU	4.7
2	B	88	TRP	4.7
1	A	249	LYS	4.6
1	A	134	SER	4.6
1	A	73	LYS	4.6
1	A	131	THR	4.6
1	A	128	THR	4.6
1	C	131	THR	4.5
1	C	252	TRP	4.5
1	C	34	LEU	4.4
3	T	703	DG	4.4
3	E	703	DG	4.3
1	C	109	LEU	4.3
1	C	141	GLY	4.2
1	A	28	GLU	4.2
2	D	215	THR	4.2
1	A	283	LEU	4.1
1	C	124	PHE	4.0
1	A	136	ASN	3.9
2	D	232	TYR	3.9
2	D	174	GLN	3.7
1	C	144	TYR	3.7
1	C	-1	MET	3.6
2	D	209	LEU	3.6
1	A	286	THR	3.6
1	A	140	PRO	3.5
1	C	61	PHE	3.5
1	C	28	GLU	3.5
1	C	27	THR	3.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	D	202	ILE	3.4
1	C	22	LYS	3.4
1	A	292	VAL	3.4
1	A	26	LEU	3.3
2	D	89	GLU	3.3
1	A	142	ILE	3.3
1	C	292	VAL	3.3
1	C	246	LEU	3.2
1	C	247	PRO	3.2
2	D	428	GLN	3.2
1	C	35	VAL	3.2
2	D	88	TRP	3.1
1	C	127	TYR	3.1
1	A	294	PRO	3.1
1	C	47	ILE	3.1
2	B	211	ARG	3.1
2	D	85	GLN	3.1
1	A	148	VAL	3.0
2	D	93	GLY	3.0
2	B	209	LEU	3.0
2	B	86	ASP	3.0
2	D	211	ARG	3.0
1	A	61	PHE	2.9
1	C	26	LEU	2.9
3	E	725	DG	2.9
1	A	124	PHE	2.9
1	A	544	GLY	2.9
2	D	216	THR	2.9
1	C	67	ASP	2.9
1	A	290	THR	2.9
1	A	551	LEU	2.9
1	A	264	LEU	2.9
1	C	257	ILE	2.9
1	C	210	LEU	2.9
1	C	309	ILE	2.9
1	C	128	THR	2.8
1	C	310	LEU	2.8
1	C	229	TRP	2.7
1	A	248	GLU	2.7
2	D	95	PRO	2.7
1	C	25	PRO	2.7
1	A	287	LYS	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	D	67	ASP	2.6
1	C	207	GLN	2.6
1	C	21	VAL	2.6
1	C	167	ILE	2.6
2	D	217	PRO	2.6
1	C	130	PHE	2.6
2	D	188	TYR	2.5
1	A	260	LEU	2.5
1	C	106	VAL	2.5
2	D	87	PHE	2.5
1	C	286	THR	2.5
1	C	140	PRO	2.5
1	A	295	LEU	2.5
1	C	62	ALA	2.4
2	D	86	ASP	2.4
1	C	248	GLU	2.4
1	A	146	TYR	2.4
1	C	66	LYS	2.4
1	C	180	ILE	2.4
2	D	213	GLY	2.4
1	C	146	TYR	2.4
2	D	146	TYR	2.4
1	C	63	ILE	2.4
1	A	288	ALA	2.3
1	C	283	LEU	2.3
1	C	2	ILE	2.3
2	D	425	LEU	2.2
2	D	205	LEU	2.2
2	D	66	LYS	2.2
1	C	296	THR	2.2
2	B	89	GLU	2.2
1	A	252	TRP	2.2
4	F	803	DC	2.2
1	C	228	LEU	2.1
1	A	109	LEU	2.1
2	B	87	PHE	2.1
2	D	94	ILE	2.1
1	C	219	LYS	2.1
1	C	277	ARG	2.1
2	D	90	VAL	2.1
2	D	195	ILE	2.1
1	A	137	ASN	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	25	PRO	2.1
2	B	92	LEU	2.0
1	C	276	VAL	2.0
1	C	15	GLY	2.0
1	C	23	GLN	2.0
1	A	129	ALA	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	MRG	F	817	22/27	0.88	0.11	81,115,142,148	0
4	MRG	P	817	22/27	0.93	0.11	57,84,110,126	0
4	DDG	F	822	21/22	0.95	0.12	71,85,100,107	0
4	DDG	P	822	21/22	0.96	0.15	44,54,63,69	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.

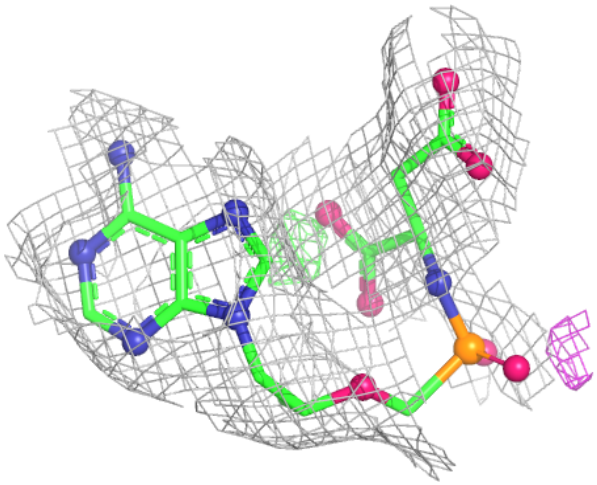
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
6	REK	C	603	27/27	0.86	0.15	88,103,126,129	0
5	MN	C	602	1/1	0.89	0.13	69,69,69,69	0
6	REK	A	603	27/27	0.91	0.20	54,70,97,105	0
5	MN	A	602	1/1	0.91	0.09	87,87,87,87	0
5	MN	C	601	1/1	0.95	0.07	124,124,124,124	0
5	MN	A	601	1/1	0.95	0.14	81,81,81,81	0

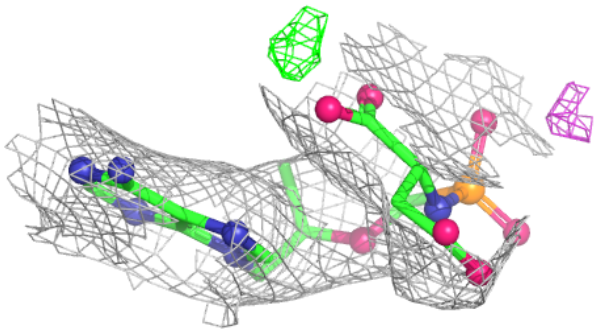
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers

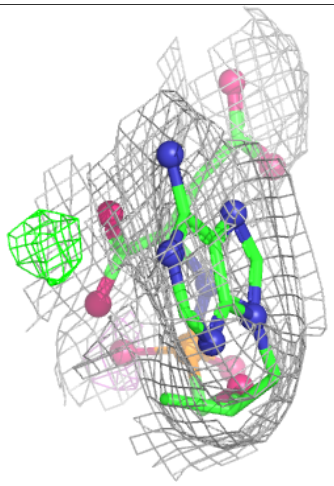
as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around REK C 603:

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

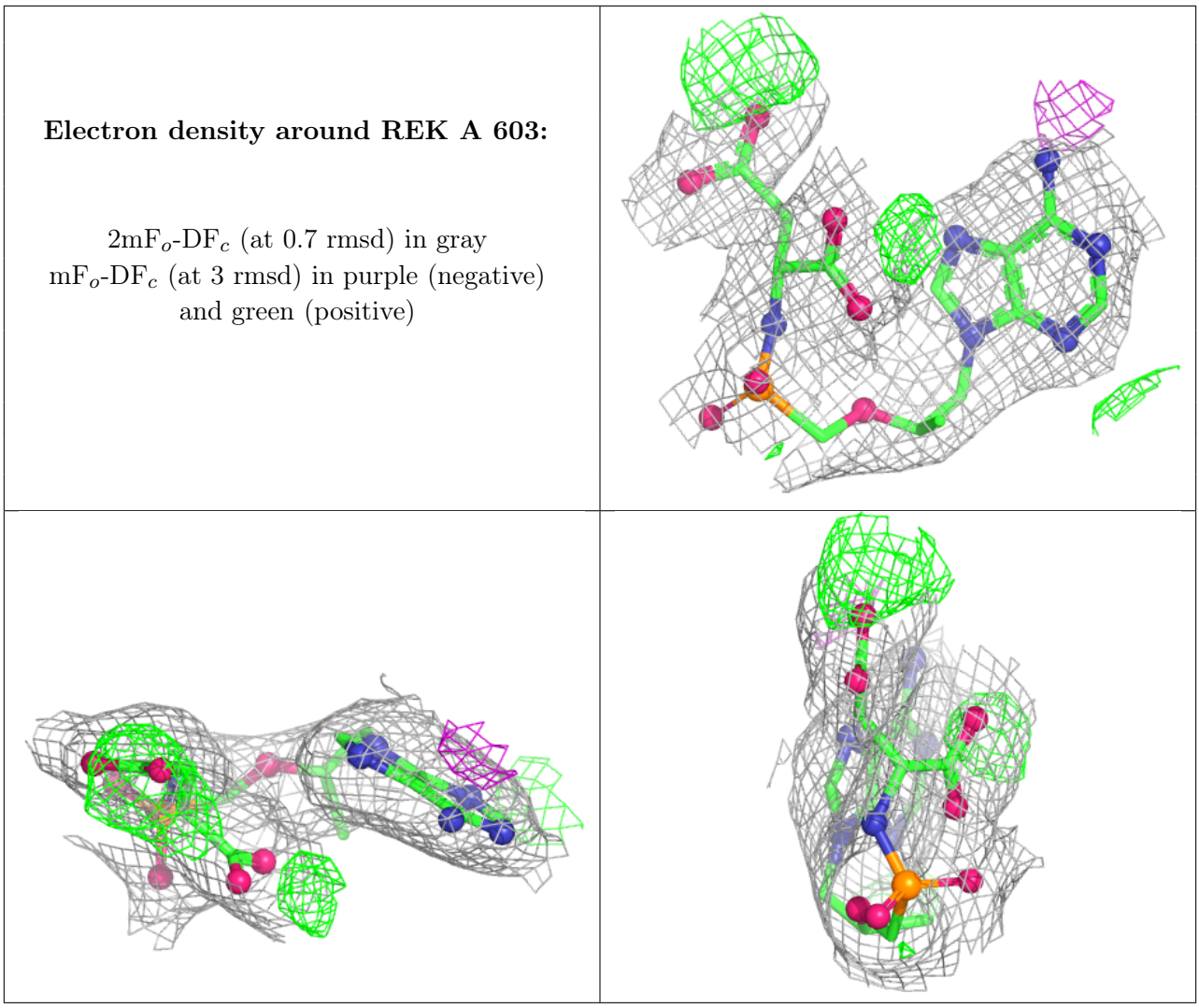






Electron density around REK A 603:

$2mF_o$-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.