Full wwPDB X-ray Structure Validation Report ⓘ

Aug 30, 2021 – 12:55 PM EDT

PDB ID	:	6UG1
Title	:	Sequence impact in DNA duplex opening by the Rad4/XPC nucleotide excision repair complex
Authors	:	Paul, D.; Min, J.-H.
Deposited on	:	2019-09-25
Resolution	:	2.83 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:

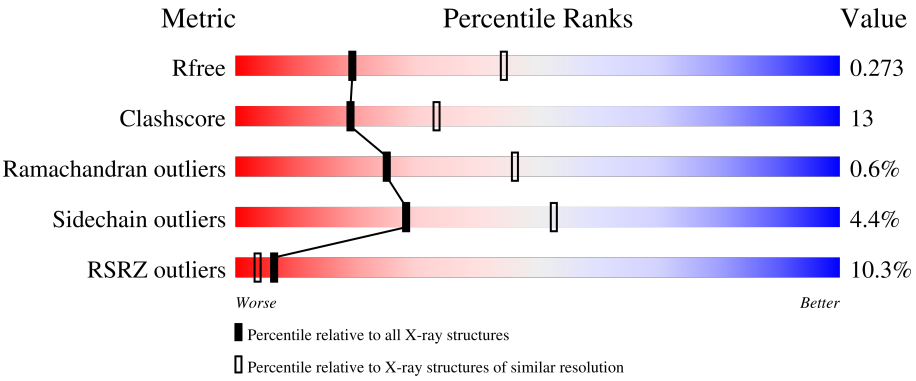
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.23.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.1

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 2.83 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	1031 (2.86-2.82)
Clashscore	141614	1078 (2.86-2.82)
Ramachandran outliers	138981	1050 (2.86-2.82)
Sidechain outliers	138945	1051 (2.86-2.82)
RSRZ outliers	127900	1019 (2.86-2.82)

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	504	11% 65% 26% · 5%
2	X	57	4% 63% 33% ·
3	Y	21	24% 29% 48%
4	W	21	10% 43% 48%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 10173 atoms, of which 4942 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 DNA repair protein RAD4.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	477	7964	2520	4027	706	684	27	0	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	131	CYS	VAL	conflict	UNP P14736
A	132	SER	CYS	conflict	UNP P14736
A	223	GLU	VAL	conflict	UNP P14736
A	427	ARG	GLN	conflict	UNP P14736
A	527	ASP	GLU	conflict	UNP P14736
A	528	ALA	ASP	conflict	UNP P14736

- Molecule 2 is a protein called UV excision repair protein RAD23.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
2	X	57	872	274	439	73	84	2	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
X	255	ALA	-	expression tag	UNP P32628

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

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	P			
3	Y	21	671	205	236	83	126	21	0	0	0

- Molecule 4 is a DNA chain called DNA (5'-D(*TP*TP*GP*AP*CP*TP*CP*(G47)P*AP*

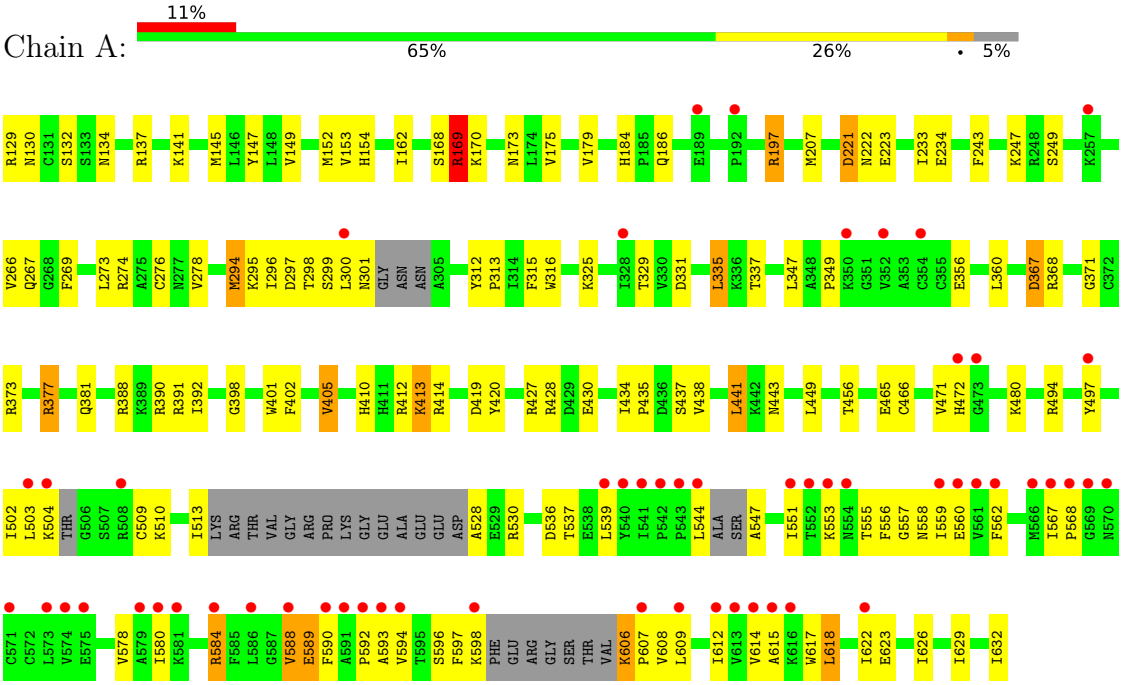
CP*AP*TP*CP*CP*CP*CP*TP*AP*CP*AP*A)-3').

Mol	Chain	Residues	Atoms							ZeroOcc	AltConf	Trace
			Total	C	H	N	O	P	S			
4	W	21	666	204	240	77	124	20	1	0	0	0

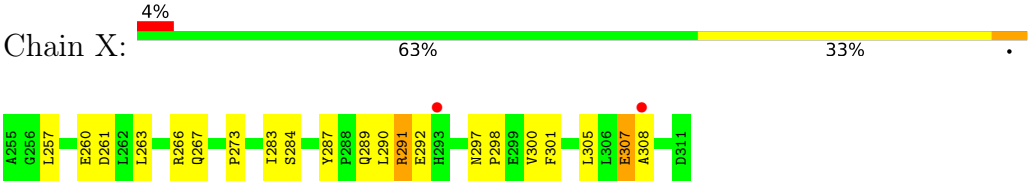
3 Residue-property plots (i)

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.

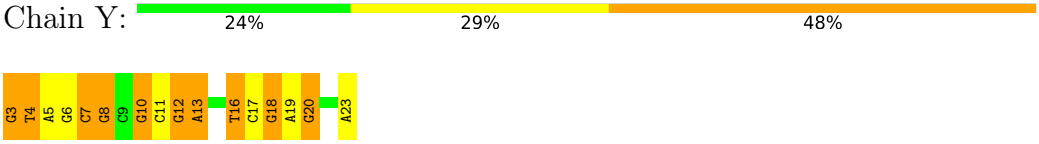
• Molecule 1: DNA repair protein RAD4



• Molecule 2: UV excision repair protein RAD23



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



• Molecule 4: DNA (5'-D(*TP*TP*GP*AP*CP*TP*CP*(G47)P*AP*CP*AP*TP*CP*CP*CP*CP*TP*AP*CP*AP*A)-3')

Chain W:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	53.24Å 59.57Å 78.24Å 105.51° 97.88° 107.14°	Depositor
Resolution (Å)	39.41 – 2.83 39.41 – 2.83	Depositor EDS
% Data completeness (in resolution range)	87.7 (39.41-2.83) 88.8 (39.41-2.83)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.02 (at 2.81Å)	Xtriage
Refinement program	PHENIX 1.12_2829	Depositor
R, R_{free}	0.224 , 0.272 0.225 , 0.273	Depositor DCC
R_{free} test set	905 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å^2)	60.2	Xtriage
Anisotropy	0.378	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 48.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.81	EDS
Total number of atoms	10173	wwPDB-VP
Average B, all atoms (Å^2)	84.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.28% 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: G47

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	0.89	4/4021 (0.1%)	1.10	17/5397 (0.3%)
2	X	0.80	0/439	1.16	2/598 (0.3%)
3	Y	1.77	7/488 (1.4%)	1.91	18/752 (2.4%)
4	W	1.67	5/447 (1.1%)	1.77	14/684 (2.0%)
All	All	1.07	16/5395 (0.3%)	1.28	51/7431 (0.7%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
2	X	0	1

All (16) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	Y	10	DG	C3'-O3'	8.00	1.54	1.44
3	Y	3	DG	C3'-O3'	-7.69	1.33	1.44
4	W	9	DA	C3'-O3'	-7.26	1.34	1.44
1	A	405	VAL	CB-CG2	-7.25	1.37	1.52
4	W	22	DC	C1'-N1	7.12	1.58	1.49
3	Y	16	DT	C3'-O3'	7.02	1.53	1.44
1	A	147	TYR	CB-CG	-6.56	1.41	1.51
4	W	13	DC	C3'-O3'	-6.16	1.35	1.44
1	A	466	CYS	CB-SG	-5.90	1.72	1.81
3	Y	20	DG	C3'-O3'	5.71	1.51	1.44
4	W	19	DC	C3'-O3'	-5.43	1.36	1.44
1	A	443	ASN	C-N	-5.40	1.21	1.34
3	Y	18	DG	C3'-O3'	5.40	1.50	1.44
3	Y	19	DA	C5-C6	-5.06	1.36	1.41

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	W	12	DT	C1'-N1	5.05	1.55	1.49
3	Y	19	DA	C3'-O3'	-5.02	1.37	1.44

All (51) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	Y	10	DG	O4'-C1'-N9	-15.48	97.16	108.00
3	Y	11	DC	O5'-P-OP1	-11.70	95.17	105.70
3	Y	13	DA	O5'-P-OP1	-11.37	95.46	105.70
4	W	9	DA	O4'-C1'-N9	-10.12	100.92	108.00
4	W	12	DT	O4'-C1'-N1	8.98	114.28	108.00
1	A	390	ARG	NE-CZ-NH1	8.66	124.63	120.30
3	Y	8	DG	O5'-P-OP1	-8.31	98.22	105.70
4	W	16	DG	O5'-P-OP2	-7.93	98.56	105.70
4	W	6	DT	O4'-C1'-N1	-7.88	102.49	108.00
4	W	11	DA	O5'-P-OP2	-7.45	99.00	105.70
1	A	152	MET	CG-SD-CE	7.41	112.06	100.20
4	W	3	DG	O4'-C1'-N9	-7.38	102.84	108.00
1	A	373	ARG	NE-CZ-NH2	-7.29	116.66	120.30
3	Y	19	DA	O5'-P-OP2	-7.28	99.15	105.70
4	W	11	DA	C5'-C4'-C3'	7.09	126.86	114.10
1	A	584	ARG	NE-CZ-NH1	6.82	123.71	120.30
1	A	584	ARG	NE-CZ-NH2	-6.78	116.91	120.30
4	W	10	DC	O4'-C1'-N1	6.74	112.72	108.00
1	A	294	MET	CG-SD-CE	6.73	110.96	100.20
3	Y	7	DC	P-O3'-C3'	6.68	127.72	119.70
2	X	291	ARG	CG-CD-NE	6.67	125.80	111.80
4	W	18	DG	O4'-C1'-N9	6.61	112.62	108.00
1	A	391	ARG	NE-CZ-NH1	6.57	123.59	120.30
3	Y	23	DA	O4'-C1'-N9	6.48	112.53	108.00
3	Y	3	DG	O4'-C1'-N9	-6.46	103.47	108.00
3	Y	4	DT	O4'-C1'-C2'	-6.46	100.73	105.90
1	A	169	ARG	CG-CD-NE	-6.05	99.09	111.80
3	Y	7	DC	OP1-P-O3'	6.04	118.49	105.20
1	A	367	ASP	CB-CG-OD1	-6.03	112.87	118.30
4	W	7	DC	OP1-P-OP2	5.96	128.53	119.60
1	A	390	ARG	NE-CZ-NH2	-5.86	117.37	120.30
3	Y	7	DC	O5'-P-OP1	-5.78	100.50	105.70
4	W	11	DA	OP2-P-O3'	5.76	117.86	105.20
2	X	290	LEU	CB-CG-CD1	-5.68	101.34	111.00
1	A	556	PHE	CB-CG-CD2	-5.67	116.83	120.80
1	A	584	ARG	CD-NE-CZ	5.64	131.50	123.60

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	466	CYS	CA-CB-SG	5.62	124.11	114.00
1	A	388	ARG	NE-CZ-NH2	5.53	123.06	120.30
3	Y	13	DA	O4'-C1'-N9	-5.48	104.16	108.00
4	W	11	DA	O5'-P-OP1	5.38	117.15	110.70
1	A	360	LEU	CB-CG-CD2	5.36	120.10	111.00
3	Y	4	DT	N3-C4-O4	5.35	123.11	119.90
1	A	413	LYS	CA-CB-CG	5.35	125.17	113.40
1	A	367	ASP	CB-CG-OD2	5.34	123.11	118.30
3	Y	16	DT	P-O3'-C3'	5.25	126.00	119.70
4	W	11	DA	P-O3'-C3'	5.24	125.99	119.70
4	W	20	DT	P-O3'-C3'	-5.18	113.48	119.70
3	Y	16	DT	N3-C4-O4	5.16	123.00	119.90
3	Y	12	DG	OP2-P-O3'	5.15	116.53	105.20
3	Y	5	DA	P-O3'-C3'	5.11	125.83	119.70
3	Y	5	DA	O4'-C1'-N9	5.10	111.57	108.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	X	307	GLU	Peptide

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	3937	4027	4026	100	2
2	X	433	439	439	16	0
3	Y	435	236	236	13	0
4	W	426	240	240	13	2
All	All	5231	4942	4941	130	2

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 13.

All (130) 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:137:ARG:NE	1:A:294:MET:SD	2.51	0.83
1:A:295:LYS:HG3	4:W:9:DA:H5"	1.73	0.71
1:A:130:ASN:OD1	1:A:296:ILE:HG22	1.93	0.69
1:A:594:VAL:HA	1:A:608:VAL:O	1.94	0.68
1:A:593:ALA:O	1:A:609:LEU:HA	1.96	0.65
1:A:129:ARG:NH1	3:Y:18:DG:O4'	2.29	0.64
1:A:615:ALA:HB3	1:A:618:LEU:HD12	1.81	0.61
2:X:289:GLN:N	2:X:289:GLN:OE1	2.31	0.61
1:A:223:GLU:OE1	1:A:223:GLU:N	2.33	0.61
4:W:11:DA:H2"	4:W:12:DT:H5'	1.83	0.61
1:A:377:ARG:HD2	1:A:428:ARG:HB3	1.83	0.61
1:A:367:ASP:OD1	1:A:371:GLY:N	2.33	0.61
1:A:503:LEU:O	1:A:504:LYS:HG3	2.03	0.59
2:X:297:ASN:O	2:X:300:VAL:HG22	2.03	0.58
2:X:260:GLU:HG3	2:X:261:ASP:N	2.18	0.58
1:A:567:ILE:HD12	1:A:568:PRO:O	2.04	0.57
1:A:414:ARG:NH1	1:A:419:ASP:OD1	2.36	0.57
2:X:263:LEU:O	2:X:267:GLN:HG3	2.05	0.57
1:A:588:VAL:HG12	1:A:589:GLU:N	2.20	0.57
1:A:553:LYS:HG3	1:A:557:GLY:HA2	1.87	0.56
1:A:544:LEU:HD12	1:A:544:LEU:O	2.05	0.56
3:Y:6:DG:H2"	3:Y:7:DC:H5'	1.88	0.56
1:A:169:ARG:HB2	1:A:169:ARG:HH11	1.71	0.56
1:A:465:GLU:HG2	1:A:480:LYS:HD3	1.87	0.56
1:A:606:LYS:N	1:A:607:PRO:HD3	2.21	0.56
1:A:580:ILE:O	1:A:584:ARG:HG3	2.07	0.55
1:A:578:VAL:CG2	1:A:626:ILE:CD1	2.84	0.55
1:A:137:ARG:HG2	1:A:294:MET:SD	2.47	0.55
1:A:465:GLU:CG	1:A:480:LYS:HD3	2.37	0.55
4:W:9:DA:H2"	4:W:10:DC:C6	2.43	0.54
4:W:19:DC:H2'	4:W:20:DT:C6	2.43	0.54
1:A:300:LEU:HD13	1:A:356:GLU:HB3	1.90	0.53
1:A:221:ASP:O	1:A:222:ASN:C	2.44	0.53
4:W:14:DG:H2"	4:W:15:DC:C6	2.44	0.53
1:A:597:PHE:HB3	1:A:606:LYS:HG2	1.92	0.52
1:A:622:ILE:O	1:A:626:ILE:HG22	2.10	0.52
1:A:590:PHE:HA	1:A:614:VAL:HG12	1.92	0.52
1:A:553:LYS:HB2	1:A:558:ASN:O	2.09	0.52
1:A:266:VAL:HG21	1:A:316:TRP:HA	1.91	0.52
3:Y:7:DC:H2"	3:Y:8:DG:C8	2.46	0.51
1:A:141:LYS:O	1:A:145:MET:HG3	2.10	0.51
1:A:502:ILE:CG2	1:A:503:LEU:N	2.73	0.51

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:184:HIS:CE1	1:A:276:CYS:O	2.64	0.50
1:A:437:SER:HB2	3:Y:13:DA:OP1	2.11	0.50
1:A:331:ASP:C	1:A:331:ASP:OD1	2.50	0.50
1:A:588:VAL:O	1:A:589:GLU:HG2	2.12	0.50
3:Y:16:DT:H4'	3:Y:17:DC:OP1	2.12	0.50
1:A:392:ILE:O	1:A:398:GLY:HA3	2.12	0.50
1:A:547:ALA:O	1:A:617:TRP:CZ2	2.65	0.49
1:A:170:LYS:O	1:A:173:ASN:OD1	2.30	0.49
1:A:325:LYS:HE2	1:A:420:TYR:CE2	2.48	0.49
1:A:588:VAL:HG13	1:A:618:LEU:HD22	1.95	0.49
4:W:9:DA:H2"	4:W:10:DC:H6	1.79	0.47
1:A:297:ASP:O	1:A:298:THR:OG1	2.25	0.47
1:A:606:LYS:O	1:A:606:LYS:CG	2.62	0.47
1:A:402:PHE:O	1:A:405:VAL:HG22	2.13	0.47
1:A:494:ARG:NH1	1:A:497:TYR:HD2	2.12	0.47
1:A:592:PRO:HB3	1:A:609:LEU:HD23	1.96	0.47
1:A:598:LYS:HG3	1:A:608:VAL:CG2	2.45	0.47
2:X:266:ARG:HD3	2:X:308:ALA:HB3	1.96	0.47
3:Y:3:DG:C2'	3:Y:4:DT:H71	2.44	0.47
1:A:295:LYS:CG	4:W:9:DA:H5"	2.43	0.47
1:A:175:VAL:HG12	1:A:179:VAL:HG23	1.97	0.47
1:A:427:ARG:O	1:A:430:GLU:HG2	2.15	0.47
2:X:283:ILE:HG23	2:X:284:SER:N	2.30	0.47
1:A:441:LEU:HB3	1:A:449:LEU:CD2	2.46	0.46
1:A:154:HIS:CE1	2:X:273:PRO:HG3	2.50	0.46
1:A:560:GLU:HB3	1:A:562:PHE:CE2	2.50	0.46
1:A:273:LEU:HB3	1:A:278:VAL:CG1	2.45	0.46
1:A:145:MET:HB3	1:A:401:TRP:CZ2	2.51	0.46
2:X:301:PHE:O	2:X:305:LEU:CD1	2.64	0.45
1:A:175:VAL:CG1	1:A:179:VAL:HG23	2.46	0.45
2:X:260:GLU:CG	2:X:261:ASP:N	2.80	0.45
1:A:149:VAL:O	1:A:153:VAL:HG23	2.16	0.45
4:W:16:DG:H1'	4:W:17:DC:C6	2.52	0.45
1:A:551:ILE:O	1:A:553:LYS:HE2	2.16	0.44
2:X:257:LEU:HA	2:X:287:TYR:CE2	2.52	0.44
1:A:129:ARG:CZ	3:Y:18:DG:O4'	2.65	0.44
1:A:503:LEU:CD2	1:A:537:THR:HG22	2.47	0.44
1:A:300:LEU:CD1	1:A:356:GLU:HB3	2.48	0.44
1:A:401:TRP:NE1	2:X:298:PRO:HG3	2.31	0.44
2:X:283:ILE:CG2	2:X:284:SER:N	2.81	0.44
1:A:438:VAL:HG22	3:Y:13:DA:P	2.58	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:629:ILE:HD12	1:A:632:ILE:HD12	2.00	0.44
2:X:291:ARG:HG2	2:X:292:GLU:N	2.32	0.44
2:X:301:PHE:O	2:X:305:LEU:HD13	2.17	0.44
1:A:134:ASN:HB2	3:Y:20:DG:H5"	1.99	0.44
1:A:335:LEU:O	1:A:337:THR:HG23	2.18	0.43
1:A:162:ILE:HD13	1:A:162:ILE:HA	1.88	0.43
2:X:307:GLU:O	2:X:307:GLU:CG	2.65	0.43
1:A:502:ILE:HG22	1:A:503:LEU:N	2.34	0.43
1:A:371:GLY:HA2	1:A:410:HIS:ND1	2.33	0.43
1:A:137:ARG:CG	1:A:294:MET:SD	3.06	0.43
1:A:145:MET:O	1:A:149:VAL:HG23	2.18	0.43
1:A:137:ARG:CD	1:A:294:MET:SD	3.06	0.43
1:A:513:ILE:O	1:A:528:ALA:HA	2.19	0.43
4:W:5:DC:H2"	4:W:6:DT:H5'	2.01	0.43
1:A:434:ILE:HG22	1:A:435:PRO:O	2.19	0.42
1:A:509:CYS:HB3	1:A:530:ARG:HB3	2.01	0.42
3:Y:12:DG:H1'	3:Y:13:DA:C8	2.54	0.42
1:A:502:ILE:CG2	1:A:503:LEU:H	2.32	0.42
1:A:299:SER:O	1:A:300:LEU:HD23	2.19	0.42
3:Y:10:DG:O6	4:W:14:DG:C6	2.72	0.42
1:A:274:ARG:HD3	1:A:368:ARG:HG2	2.01	0.42
1:A:313:PRO:HG3	1:A:331:ASP:OD2	2.19	0.42
4:W:16:DG:H1'	4:W:17:DC:C5	2.55	0.42
2:X:307:GLU:O	2:X:307:GLU:HG3	2.19	0.42
1:A:300:LEU:HD13	1:A:356:GLU:CG	2.50	0.42
1:A:557:GLY:O	1:A:592:PRO:HD2	2.20	0.42
1:A:596:SER:O	1:A:597:PHE:HD1	2.03	0.42
1:A:312:TYR:CE1	1:A:349:PRO:HA	2.55	0.41
1:A:559:ILE:N	1:A:592:PRO:O	2.44	0.41
1:A:266:VAL:HG11	1:A:315:PHE:HB2	2.02	0.41
1:A:269:PHE:CE1	1:A:273:LEU:HD11	2.56	0.41
1:A:367:ASP:OD2	1:A:412:ARG:NE	2.54	0.41
1:A:471:VAL:HG13	3:Y:13:DA:H5"	2.03	0.41
1:A:266:VAL:HG13	1:A:267:GLN:N	2.36	0.41
1:A:606:LYS:O	1:A:606:LYS:HG3	2.20	0.41
1:A:598:LYS:C	1:A:606:LYS:HG3	2.42	0.41
1:A:623:GLU:HA	1:A:626:ILE:HG22	2.03	0.41
1:A:438:VAL:HG22	3:Y:13:DA:O5'	2.21	0.40
1:A:612:ILE:CD1	1:A:622:ILE:HD13	2.51	0.40
1:A:197:ARG:O	1:A:197:ARG:HD2	2.20	0.40
4:W:20:DT:H2'	4:W:21:DA:C8	2.56	0.40

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:169:ARG:CG	1:A:170:LYS:N	2.84	0.40
1:A:233:ILE:O	1:A:234:GLU:C	2.57	0.40
1:A:329:THR:HG21	1:A:347:LEU:HD12	2.02	0.40
1:A:536:ASP:OD1	1:A:536:ASP:N	2.54	0.40
4:W:18:DG:C4	4:W:19:DC:C4	3.09	0.40
1:A:594:VAL:HG13	1:A:608:VAL:H	1.86	0.40

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:186:GLN:NE2	4:W:22:DC:O3'[1_545]	2.08	0.12
1:A:186:GLN:HE22	4:W:22:DC:O3'[1_545]	1.50	0.10

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	465/504 (92%)	433 (93%)	29 (6%)	3 (1%)	25 46
2	X	55/57 (96%)	48 (87%)	7 (13%)	0	100 100
All	All	520/561 (93%)	481 (92%)	36 (7%)	3 (1%)	25 46

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	588	VAL
1	A	589	GLU
1	A	247	LYS

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	430/451 (95%)	409 (95%)	21 (5%)	25 47
2	X	48/48 (100%)	48 (100%)	0	100 100
All	All	478/499 (96%)	457 (96%)	21 (4%)	28 53

All (21) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	132	SER
1	A	168	SER
1	A	169	ARG
1	A	197	ARG
1	A	207	MET
1	A	221	ASP
1	A	243	PHE
1	A	249	SER
1	A	301	ASN
1	A	335	LEU
1	A	377	ARG
1	A	381	GLN
1	A	413	LYS
1	A	441	LEU
1	A	456	THR
1	A	472	HIS
1	A	510	LYS
1	A	539	LEU
1	A	555	THR
1	A	606	LYS
1	A	618	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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	G47	W	8	4,3,1	20,27,28	4.45	13 (65%)	21,38,41	1.94	9 (42%)

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	G47	W	8	4,3,1	-	3/7/25/26	0/3/3/3

All (13) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	W	8	G47	O4'-C1'	8.61	1.61	1.42
4	W	8	G47	C4-N3	8.23	1.48	1.35
4	W	8	G47	O4'-C4'	-6.49	1.30	1.45
4	W	8	G47	C6-N1	6.40	1.44	1.33
4	W	8	G47	C6-C5	6.27	1.52	1.41
4	W	8	G47	C2'-C1'	-6.22	1.35	1.52
4	W	8	G47	C2-N2	6.06	1.43	1.34
4	W	8	G47	C2-N1	3.03	1.43	1.34
4	W	8	G47	O3'-C3'	-2.90	1.37	1.43
4	W	8	G47	C2'-C3'	-2.86	1.45	1.52
4	W	8	G47	C2-N3	2.83	1.43	1.34
4	W	8	G47	C3'-C4'	2.24	1.59	1.53
4	W	8	G47	C5-C4	-2.06	1.35	1.40

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	W	8	G47	C5-C6-N1	-3.57	118.55	123.43
4	W	8	G47	O3'-C3'-C4'	-2.92	98.93	110.10
4	W	8	G47	N3-C2-N1	-2.80	121.80	126.23
4	W	8	G47	C6-N1-C2	2.74	120.08	115.18
4	W	8	G47	O4'-C1'-C2'	2.48	110.94	106.25
4	W	8	G47	O3'-C3'-C2'	2.30	119.13	110.90
4	W	8	G47	O4'-C4'-C5'	-2.30	101.81	109.37
4	W	8	G47	C2'-C3'-C4'	2.19	107.32	102.76
4	W	8	G47	C6-C5-C4	-2.05	118.84	120.80

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	W	8	G47	N2-C6A-C7A-SG
4	W	8	G47	O4'-C4'-C5'-O5'
4	W	8	G47	C3'-C4'-C5'-O5'

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)

There are no ligands in this entry.

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	477/504 (94%)	0.70	57 (11%) 4 2	30, 69, 123, 166	0
2	X	57/57 (100%)	0.32	2 (3%) 44 36	41, 69, 100, 111	0
3	Y	21/21 (100%)	-0.02	0 100 100	56, 79, 135, 174	0
4	W	20/21 (95%)	-0.20	0 100 100	64, 88, 146, 153	0
All	All	575/603 (95%)	0.60	59 (10%) 6 4	30, 70, 125, 174	0

All (59) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	559	ILE	9.7
1	A	560	GLU	8.5
1	A	614	VAL	7.2
1	A	573	LEU	6.1
1	A	615	ALA	5.4
1	A	568	PRO	5.2
1	A	571	CYS	5.0
1	A	540	TYR	4.9
1	A	612	ILE	4.9
1	A	561	VAL	4.9
1	A	613	VAL	4.9
1	A	590	PHE	4.6
1	A	354	CYS	4.6
1	A	588	VAL	4.6
1	A	562	PHE	4.4
1	A	570	ASN	4.4
1	A	508	ARG	4.3
1	A	567	ILE	4.1
1	A	553	LYS	4.1
1	A	551	ILE	4.0
1	A	579	ALA	3.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	566	MET	3.8
1	A	598	LYS	3.7
1	A	189	GLU	3.6
1	A	622	ILE	3.6
1	A	543	PRO	3.6
1	A	472	HIS	3.5
1	A	616	LYS	3.5
1	A	581	LYS	3.3
1	A	497	TYR	3.3
1	A	542	PRO	3.3
1	A	593	ALA	3.2
1	A	300	LEU	3.1
1	A	569	GLY	3.1
1	A	594	VAL	3.0
1	A	591	ALA	3.0
1	A	350	LYS	3.0
1	A	541	ILE	3.0
1	A	609	LEU	2.9
1	A	552	THR	2.8
1	A	580	ILE	2.8
1	A	607	PRO	2.8
1	A	504	LYS	2.7
1	A	584	ARG	2.6
1	A	503	LEU	2.5
1	A	592	PRO	2.5
1	A	539	LEU	2.5
1	A	586	LEU	2.4
1	A	257	LYS	2.4
1	A	574	VAL	2.3
1	A	352	VAL	2.3
1	A	554	ASN	2.2
2	X	293	HIS	2.1
1	A	473	GLY	2.1
1	A	192	PRO	2.1
2	X	308	ALA	2.1
1	A	544	LEU	2.1
1	A	575	GLU	2.1
1	A	328	ILE	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	G47	W	8	25/26	0.85	0.24	58,75,103,116	0

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.