Full wwPDB X-ray Structure Validation Report ⓘ

Nov 19, 2023 – 05:59 PM JST

PDB ID	:	6LWI
Title	:	Crystal structure of human NEIL1(P2G, E3Q, R242) bound to duplex DNA containing dihydrothymine (DHT)
Authors	:	Liu, M.H.; Zhang, J.; Zhu, C.X.; Zhang, X.X.; Gao, Y.Q.; Yi, C.Q.
Deposited on	:	2020-02-07
Resolution	:	2.72 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:

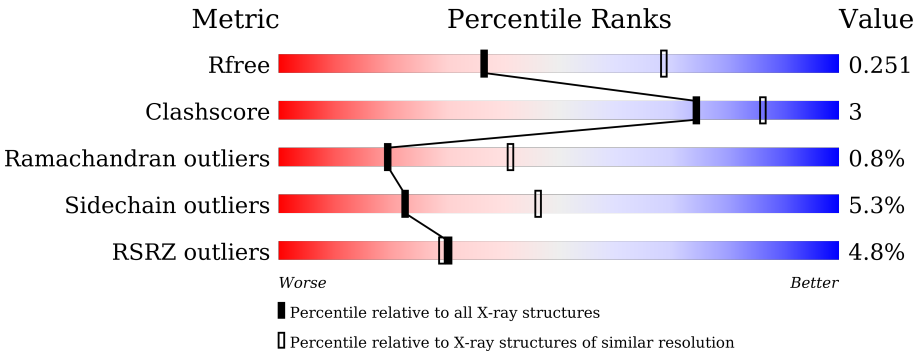
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.72 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	3359 (2.74-2.70)
Clashscore	141614	3686 (2.74-2.70)
Ramachandran outliers	138981	3622 (2.74-2.70)
Sidechain outliers	138945	3623 (2.74-2.70)
RSRZ outliers	127900	3276 (2.74-2.70)

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	295	80% 9% • 9%
1	D	295	81% 9% • 8%
1	G	295	13% 69% 7% 24%
2	B	13	62% 38%
2	E	13	85% 8% 8%
2	H	13	77% 23%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	C	13	62% 38%
3	F	13	62% 38%
3	I	13	69% 31%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 7762 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 Endonuclease 8-like 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	269	Total	C	N	O	S	0	0	0
			2131	1358	389	374	10			
1	D	270	Total	C	N	O	S	0	0	0
			2136	1359	391	376	10			
1	G	225	Total	C	N	O	S	0	0	0
			1608	998	312	290	8			

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	2	GLY	PRO	engineered mutation	UNP Q96FI4
A	3	GLN	GLU	engineered mutation	UNP Q96FI4
A	242	ARG	LYS	variant	UNP Q96FI4
D	2	GLY	PRO	engineered mutation	UNP Q96FI4
D	3	GLN	GLU	engineered mutation	UNP Q96FI4
D	242	ARG	LYS	variant	UNP Q96FI4
G	2	GLY	PRO	engineered mutation	UNP Q96FI4
G	3	GLN	GLU	engineered mutation	UNP Q96FI4
G	242	ARG	LYS	variant	UNP Q96FI4

- Molecule 2 is a DNA chain called DNA (5'-D(*CP*GP*TP*CP*CP*AP*(TDH)P*GP*TP*CP*TP*AP*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	13	Total	C	N	O	P	0	0	0
			258	125	43	78	12			
2	E	13	Total	C	N	O	P	0	0	0
			258	125	43	78	12			
2	H	13	Total	C	N	O	P	0	0	0
			258	125	43	78	12			

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

CP*GP*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	13	Total	C	N	O	P	0	0	0
			267	127	53	75	12			
3	F	13	Total	C	N	O	P	0	0	0
			267	127	53	75	12			
3	I	13	Total	C	N	O	P	0	0	0
			267	127	53	75	12			

- Molecule 4 is water.

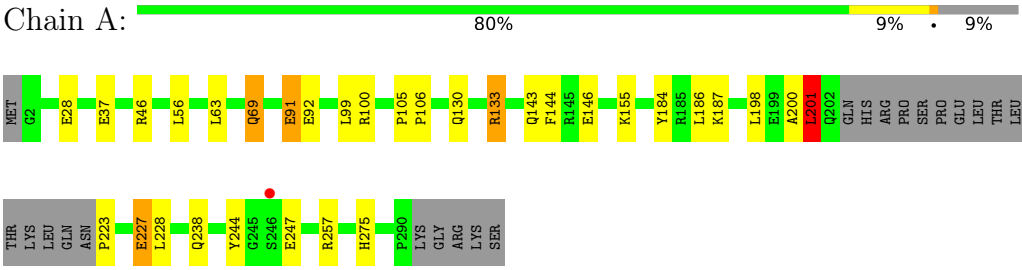
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	160	Total	O	0	0
			160	160		
4	B	16	Total	O	0	0
			16	16		
4	C	9	Total	O	0	0
			9	9		
4	D	90	Total	O	0	0
			90	90		
4	E	16	Total	O	0	0
			16	16		
4	F	16	Total	O	0	0
			16	16		
4	G	2	Total	O	0	0
			2	2		
4	H	1	Total	O	0	0
			1	1		
4	I	2	Total	O	0	0
			2	2		

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.

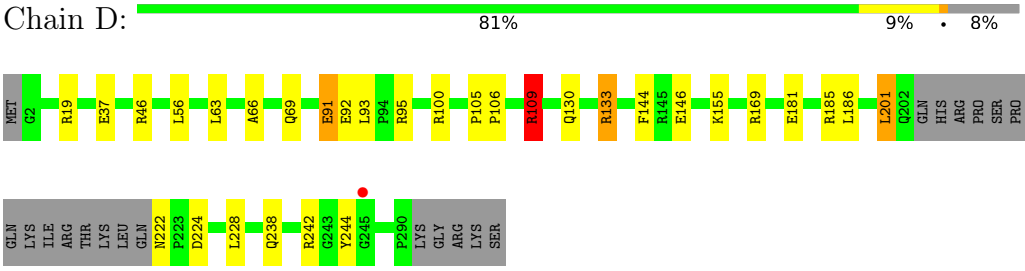
• Molecule 1: Endonuclease 8-like 1

Chain A:



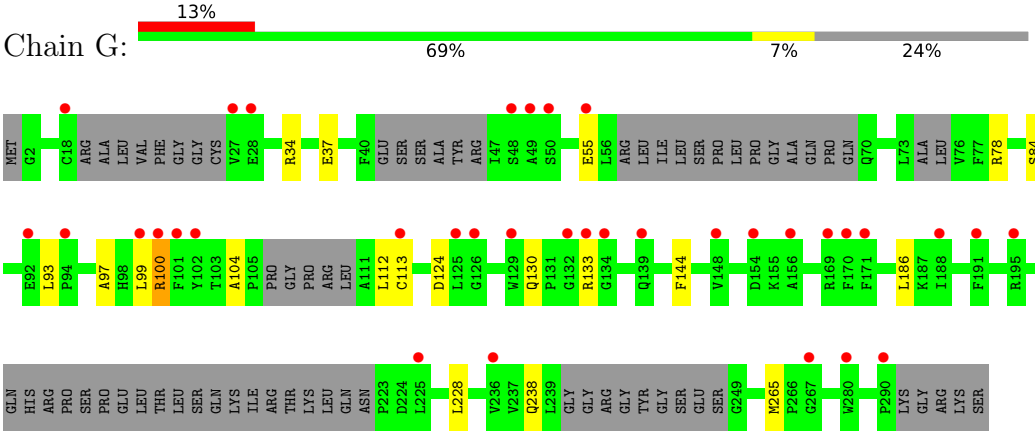
• Molecule 1: Endonuclease 8-like 1

Chain D:



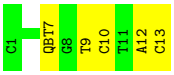
• Molecule 1: Endonuclease 8-like 1

Chain G:



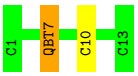
• Molecule 2: DNA (5'-D(*CP*GP*TP*CP*CP*AP*(TDH)P*GP*TP*CP*TP*AP*C)-3')

Chain B:



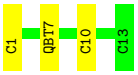
• Molecule 2: DNA (5'-D(*CP*GP*TP*CP*CP*AP*(TDH)P*GP*TP*CP*TP*AP*C)-3')

Chain E:



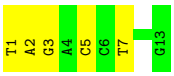
• Molecule 2: DNA (5'-D(*CP*GP*TP*CP*CP*AP*(TDH)P*GP*TP*CP*TP*AP*C)-3')

Chain H:



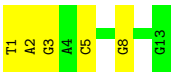
• Molecule 3: DNA (5'-D(*TP*AP*GP*AP*CP*CP*TP*GP*GP*AP*CP*GP*G)-3')

Chain C:



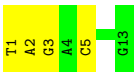
• Molecule 3: DNA (5'-D(*TP*AP*GP*AP*CP*CP*TP*GP*GP*AP*CP*GP*G)-3')

Chain F:



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

Chain I:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	74.01Å 108.61Å 170.14Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.71 – 2.72 49.66 – 2.70	Depositor EDS
% Data completeness (in resolution range)	93.4 (49.71-2.72) 93.5 (49.66-2.70)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.10 (at 2.69Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.205 , 0.248 0.207 , 0.251	Depositor DCC
R_{free} test set	1802 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	44.2	Xtriage
Anisotropy	0.116	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 53.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	7762	wwPDB-VP
Average B, all atoms (Å^2)	64.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.55% 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: QBT

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	0.85	4/2189 (0.2%)	0.92	3/2959 (0.1%)
1	D	0.83	5/2194 (0.2%)	0.91	3/2966 (0.1%)
1	G	0.75	1/1643 (0.1%)	0.81	0/2201
2	B	0.56	0/264	0.98	1/402 (0.2%)
2	E	0.55	0/264	0.90	1/402 (0.2%)
2	H	0.50	0/264	0.83	1/402 (0.2%)
3	C	0.65	0/300	0.93	2/462 (0.4%)
3	F	0.60	0/300	0.90	2/462 (0.4%)
3	I	0.42	0/300	0.77	1/462 (0.2%)
All	All	0.76	10/7718 (0.1%)	0.89	14/10718 (0.1%)

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
1	A	0	1
1	D	0	1
All	All	0	2

All (10) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	91	GLU	CG-CD	8.96	1.65	1.51
1	D	37	GLU	CG-CD	8.04	1.64	1.51
1	D	146	GLU	CD-OE1	-7.85	1.17	1.25
1	A	227	GLU	CD-OE1	6.64	1.32	1.25
1	G	37	GLU	CG-CD	6.62	1.61	1.51
1	D	146	GLU	CG-CD	6.03	1.60	1.51
1	D	91	GLU	CD-OE1	5.87	1.32	1.25

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	92	GLU	CG-CD	5.85	1.60	1.51
1	A	37	GLU	CG-CD	5.59	1.60	1.51
1	D	146	GLU	CD-OE2	5.26	1.31	1.25

All (14) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	E	10	DC	O5'-P-OP2	-7.96	98.54	105.70
3	C	7	DT	O5'-P-OP2	-6.81	99.57	105.70
2	B	10	DC	O5'-P-OP2	-6.30	100.03	105.70
2	H	10	DC	C5'-C4'-C3'	-6.26	102.84	114.10
3	C	5	DC	C1'-O4'-C4'	-6.16	103.94	110.10
1	A	91	GLU	OE1-CD-OE2	6.13	130.66	123.30
1	D	133	ARG	NE-CZ-NH1	6.00	123.30	120.30
1	A	133	ARG	NE-CZ-NH2	-5.97	117.31	120.30
3	F	8	DG	O5'-P-OP2	-5.85	100.44	105.70
3	F	5	DC	C1'-O4'-C4'	-5.67	104.43	110.10
1	D	109	ARG	NE-CZ-NH1	5.45	123.03	120.30
3	I	5	DC	C1'-O4'-C4'	-5.44	104.66	110.10
1	D	169	ARG	NE-CZ-NH2	-5.38	117.61	120.30
1	A	201	LEU	CB-CA-C	5.36	120.39	110.20

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	105	PRO	Peptide
1	D	105	PRO	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	2131	0	2102	16	0
1	D	2136	0	2099	10	0
1	G	1608	0	1411	8	0
2	B	258	0	151	4	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	E	258	0	151	1	0
2	H	258	0	151	2	0
3	C	267	0	147	2	0
3	F	267	0	147	2	0
3	I	267	0	147	2	0
4	A	160	0	0	4	0
4	B	16	0	0	0	0
4	C	9	0	0	0	0
4	D	90	0	0	0	0
4	E	16	0	0	0	0
4	F	16	0	0	0	0
4	G	2	0	0	0	0
4	H	1	0	0	0	0
4	I	2	0	0	0	0
All	All	7762	0	6506	44	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 3.

All (44) 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:143:GLN:HG3	4:A:385:HOH:O	1.76	0.84
1:G:55:GLU:OE1	1:G:130:GLN:HB2	1.90	0.72
1:A:198:LEU:O	1:A:201:LEU:HD22	1.92	0.69
1:G:97:ALA:O	1:G:100:ARG:NH1	2.33	0.61
1:A:184:TYR:O	1:A:187:LYS:HE3	2.01	0.60
3:I:1:DT:H2"	3:I:2:DA:C8	2.38	0.59
1:D:56:LEU:HD23	1:D:56:LEU:C	2.23	0.58
3:C:1:DT:H2"	3:C:2:DA:C8	2.40	0.57
1:A:186:LEU:HD11	1:A:228:LEU:HD12	1.87	0.56
1:A:56:LEU:C	1:A:56:LEU:HD23	2.26	0.56
3:F:1:DT:H2"	3:F:2:DA:C8	2.41	0.56
2:B:13:DC:H2"	2:H:1:DC:C6	2.42	0.54
1:G:86:GLN:O	1:G:112:LEU:HD12	2.06	0.54
1:G:130:GLN:NE2	1:G:133:ARG:HH11	2.06	0.53
1:A:275:HIS:HE1	2:B:9:DT:O4	1.92	0.53
1:A:227:GLU:HG2	4:A:407:HOH:O	2.08	0.53
2:B:12:DA:H2"	2:B:13:DC:OP2	2.08	0.52
3:I:2:DA:H2"	3:I:3:DG:OP2	2.09	0.52
3:F:2:DA:H2"	3:F:3:DG:OP2	2.10	0.52

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:C:2:DA:H2"	3:C:3:DG:OP2	2.09	0.51
1:D:130:GLN:NE2	1:D:133:ARG:HH11	2.07	0.51
1:A:130:GLN:NE2	1:A:133:ARG:HH11	2.08	0.51
1:A:223:PRO:HA	1:A:227:GLU:OE1	2.11	0.50
1:D:109:ARG:HG3	1:D:109:ARG:HH11	1.75	0.50
1:D:201:LEU:HD21	1:D:224:ASP:CG	2.32	0.50
1:A:186:LEU:CD1	1:A:228:LEU:HD12	2.42	0.49
1:D:109:ARG:HH11	1:D:109:ARG:CG	2.26	0.48
1:A:247:GLU:HB3	4:A:420:HOH:O	2.14	0.48
1:A:238:GLN:HG3	4:A:443:HOH:O	2.13	0.48
1:G:93:LEU:HD13	1:G:100:ARG:CD	2.45	0.47
1:A:69:GLN:NE2	1:A:69:GLN:HA	2.23	0.47
1:G:186:LEU:HD11	1:G:228:LEU:HD12	1.97	0.47
2:B:13:DC:H2"	2:H:1:DC:H6	1.79	0.46
1:A:99:LEU:C	1:A:100:ARG:HG3	2.35	0.45
1:A:28:GLU:OE1	1:A:100:ARG:NH2	2.49	0.45
1:G:55:GLU:OE1	1:G:133:ARG:HD2	2.17	0.45
1:A:46:ARG:HG3	1:A:63:LEU:HD11	1.99	0.44
1:D:186:LEU:HD11	1:D:228:LEU:HD12	2.00	0.44
2:E:7:QBT:H2'2	2:E:7:QBT:H61C	1.89	0.42
1:D:46:ARG:HG3	1:D:63:LEU:HD11	2.01	0.41
1:G:130:GLN:HE22	1:G:133:ARG:HH11	1.69	0.41
1:D:93:LEU:HD13	1:D:100:ARG:CD	2.51	0.41
1:D:186:LEU:CD1	1:D:228:LEU:HD12	2.50	0.41
1:D:181:GLU:O	1:D:185:ARG:HG2	2.21	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	265/295 (90%)	250 (94%)	13 (5%)	2 (1%)	19 41

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	D	266/295 (90%)	248 (93%)	16 (6%)	2 (1%)	19 41
1	G	209/295 (71%)	196 (94%)	11 (5%)	2 (1%)	15 35
All	All	740/885 (84%)	694 (94%)	40 (5%)	6 (1%)	19 41

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	106	PRO
1	A	106	PRO
1	A	200	ALA
1	D	66	ALA
1	G	104	ALA
1	G	200	ALA

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	222/247 (90%)	214 (96%)	8 (4%)	35 62
1	D	222/247 (90%)	209 (94%)	13 (6%)	19 41
1	G	144/247 (58%)	134 (93%)	10 (7%)	15 34
All	All	588/741 (79%)	557 (95%)	31 (5%)	22 46

All (31) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	69	GLN
1	A	91	GLU
1	A	144	PHE
1	A	146	GLU
1	A	155	LYS
1	A	201	LEU
1	A	244	TYR
1	A	257	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	19	ARG
1	D	69	GLN
1	D	91	GLU
1	D	92	GLU
1	D	95	ARG
1	D	109	ARG
1	D	144	PHE
1	D	155	LYS
1	D	201	LEU
1	D	222	ASN
1	D	238	GLN
1	D	242	ARG
1	D	244	TYR
1	G	34	ARG
1	G	78	ARG
1	G	84	SER
1	G	99	LEU
1	G	100	ARG
1	G	113	CYS
1	G	124	ASP
1	G	144	PHE
1	G	238	GLN
1	G	265	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	67	GLN
1	A	130	GLN
1	A	139	GLN
1	A	202	GLN
1	D	70	GLN
1	D	130	GLN
1	D	238	GLN
1	G	130	GLN
1	G	238	GLN

5.3.3 RNA

There are no RNA molecules in this entry.

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

3 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	$\#\|Z\| > 2$
2	QBT	E	7	2	17,21,22	1.05	2 (11%)	22,30,33	1.64	5 (22%)
2	QBT	B	7	2	17,21,22	1.02	1 (5%)	22,30,33	1.92	5 (22%)
2	QBT	H	7	2	17,21,22	0.90	1 (5%)	22,30,33	1.55	4 (18%)

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
2	QBT	E	7	2	-	1/7/37/38	0/2/2/2
2	QBT	B	7	2	-	2/7/37/38	0/2/2/2
2	QBT	H	7	2	-	1/7/37/38	0/2/2/2

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	7	QBT	C2-N3	-2.84	1.32	1.38
2	B	7	QBT	C4-N3	-2.52	1.32	1.37
2	H	7	QBT	C2-N3	-2.20	1.34	1.38
2	E	7	QBT	C4-N3	-2.02	1.33	1.37

All (14) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	7	QBT	O4'-C1'-N1	-4.21	103.04	108.41
2	E	7	QBT	C6-N1-C1'	-4.06	112.95	120.94
2	B	7	QBT	C6-N1-C1'	-4.03	113.02	120.94
2	H	7	QBT	C6-N1-C1'	-3.82	113.42	120.94
2	E	7	QBT	C4-N3-C2	-3.48	122.60	126.86

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	E	7	QBT	O4'-C1'-N1	-3.36	104.13	108.41
2	B	7	QBT	C4-N3-C2	-3.34	122.78	126.86
2	H	7	QBT	C4-N3-C2	-3.17	122.98	126.86
2	B	7	QBT	C5M-C5-C6	-2.97	106.01	112.34
2	H	7	QBT	O4'-C1'-N1	-2.90	104.71	108.41
2	H	7	QBT	O3'-C3'-C2'	-2.56	101.75	110.90
2	B	7	QBT	O2-C2-N1	-2.54	119.92	123.11
2	E	7	QBT	O3'-C3'-C2'	-2.38	102.39	110.90
2	E	7	QBT	N3-C2-N1	2.11	118.89	116.65

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	7	QBT	O4'-C4'-C5'-O5'
2	B	7	QBT	C3'-C4'-C5'-O5'
2	E	7	QBT	O4'-C4'-C5'-O5'
2	H	7	QBT	O4'-C4'-C5'-O5'

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	E	7	QBT	1	0

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	269/295 (91%)	-0.22	1 (0%) 92 93	17, 37, 68, 95	0
1	D	270/295 (91%)	-0.00	1 (0%) 92 93	20, 56, 98, 120	0
1	G	225/295 (76%)	1.05	38 (16%) 1 1	69, 99, 129, 160	0
2	B	12/13 (92%)	-0.42	0 100 100	35, 68, 102, 120	0
2	E	12/13 (92%)	-0.39	0 100 100	30, 64, 97, 111	0
2	H	12/13 (92%)	0.24	0 100 100	76, 100, 123, 124	0
3	C	13/13 (100%)	-0.46	0 100 100	46, 65, 97, 113	0
3	F	13/13 (100%)	-0.15	0 100 100	37, 62, 107, 120	0
3	I	13/13 (100%)	0.15	0 100 100	79, 94, 112, 113	0
All	All	839/963 (87%)	0.20	40 (4%) 30 29	17, 62, 117, 160	0

All (40) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	125	LEU	7.3
1	G	100	ARG	6.6
1	G	113	CYS	4.0
1	G	27	VAL	4.0
1	G	99	LEU	3.9
1	G	94	PRO	3.8
1	G	148	VAL	3.7
1	G	28	GLU	3.6
1	G	133	ARG	3.5
1	G	101	PHE	3.5
1	G	48	SER	3.4
1	G	86	GLN	3.4
1	G	132	GLY	3.1
1	G	126	GLY	3.0
1	G	92	GLU	3.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	G	139	GLN	3.0
1	G	88	VAL	2.7
1	G	191	PHE	2.7
1	G	171	PHE	2.6
1	G	18	CYS	2.6
1	G	156	ALA	2.6
1	G	169	ARG	2.6
1	G	280	TRP	2.5
1	G	267	GLY	2.5
1	G	129	TRP	2.5
1	G	102	TYR	2.5
1	G	199	GLU	2.4
1	G	188	ILE	2.4
1	G	50	SER	2.3
1	G	170	PHE	2.3
1	G	225	LEU	2.3
1	G	134	GLY	2.3
1	G	236	VAL	2.3
1	G	49	ALA	2.2
1	G	290	PRO	2.2
1	A	246	SER	2.2
1	D	245	GLY	2.2
1	G	195	ARG	2.0
1	G	55	GLU	2.0
1	G	154	ASP	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
2	QBT	H	7	20/21	0.96	0.15	76,81,90,90	0
2	QBT	B	7	20/21	0.97	0.15	32,36,44,47	0
2	QBT	E	7	20/21	0.98	0.16	33,44,49,53	0

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.