Full wwPDB X-ray Structure Validation Report ⓘ

Nov 19, 2023 – 06:30 PM JST

PDB ID	:	6LWG
Title	:	Crystal structure of human NEIL1(P2G, E3Q, R242) bound to duplex DNA containing guanidinohydantoin (Gh)
Authors	:	Liu, M.H.; Zhang, J.; Zhu, C.X.; Zhang, X.X.; Gao, Y.Q.; Yi, C.Q.
Deposited on	:	2020-02-07
Resolution	:	2.53 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:

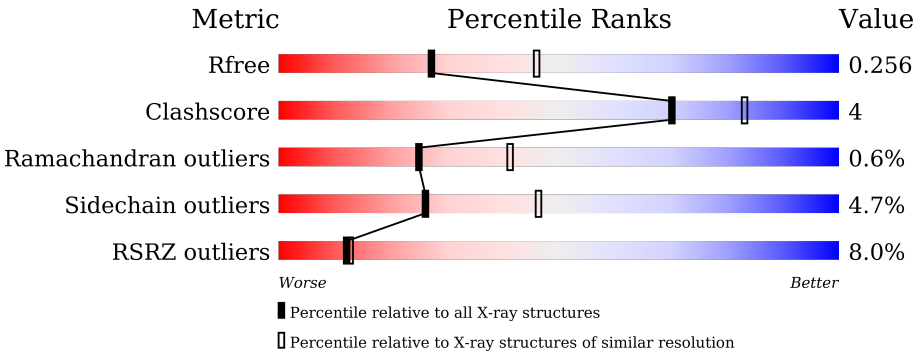
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, 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 2.53 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	5743 (2.54-2.50)
Clashscore	141614	6463 (2.54-2.50)
Ramachandran outliers	138981	6335 (2.54-2.50)
Sidechain outliers	138945	6337 (2.54-2.50)
RSRZ outliers	127900	5630 (2.54-2.50)

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	6% 81% 9% • 9%
1	D	295	3% 80% 8% • 11%
1	G	295	7% 62% 6% • 31%
2	B	13	23% 92% 8%
2	E	13	15% 92% 8%
2	H	13	46% 92% 8%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	C	13	85% 15%
3	F	13	85% 15%
3	I	13	38% 69% 31%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 7518 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	262	Total	C	N	O	S	0	0	0
			2065	1318	382	355	10			
1	G	205	Total	C	N	O	S	0	0	0
			1432	886	275	263	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*(DGH)P*GP*TP *CP*TP*AP*C)-3').

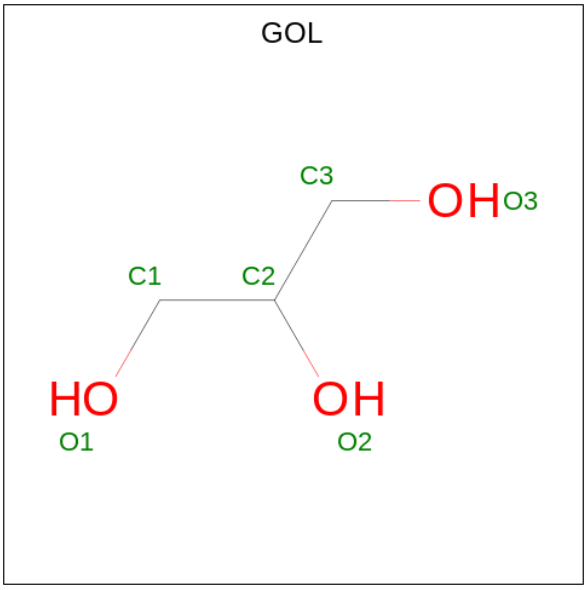
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	13	Total	C	N	O	P	0	0	0
			260	124	46	78	12			
2	E	13	Total	C	N	O	P	0	0	0
			260	124	46	78	12			
2	H	13	Total	C	N	O	P	0	0	0
			260	124	46	78	12			

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



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	B	1	Total	C	O		0	0
			6	3	3			
4	D	1	Total	C	O		0	0
			6	3	3			
4	E	1	Total	C	O		0	0
			6	3	3			

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	130	Total	O	0	0
			130	130		
5	B	18	Total	O	0	0
			18	18		
5	C	7	Total	O	0	0
			7	7		

Continued on next page...

Continued from previous page...

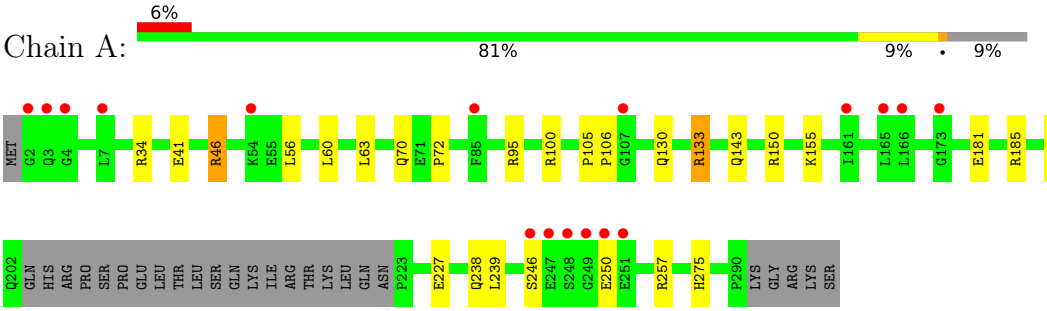
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	D	100	Total O 100 100	0	0
5	E	15	Total O 15 15	0	0
5	F	10	Total O 10 10	0	0
5	G	6	Total O 6 6	0	0
5	H	2	Total O 2 2	0	0
5	I	3	Total O 3 3	0	0

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.

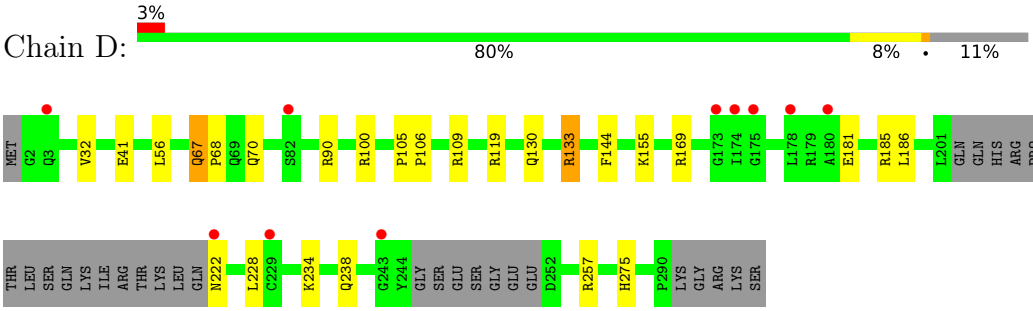
• Molecule 1: Endonuclease 8-like 1

Chain A:



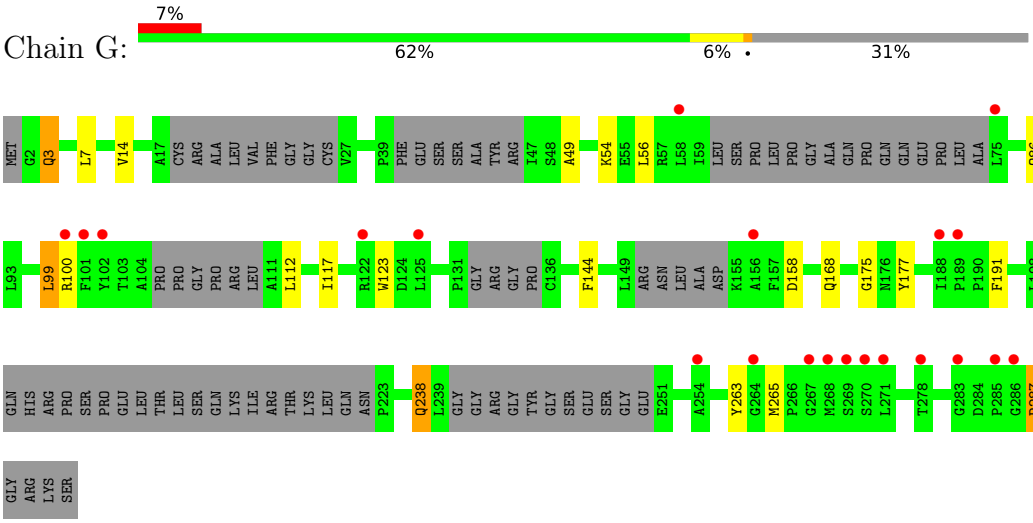
• Molecule 1: Endonuclease 8-like 1

Chain D:



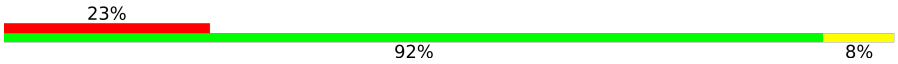
• Molecule 1: Endonuclease 8-like 1

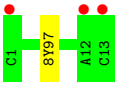
Chain G:



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

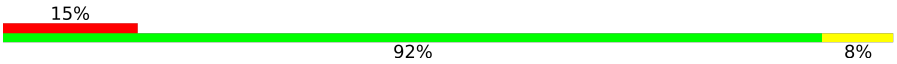
Chain B:





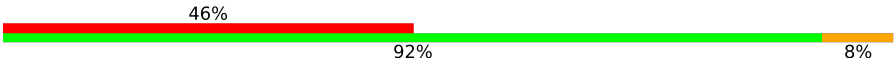
• Molecule 2: DNA (5'-D(*CP*GP*TP*CP*CP*AP*(DGH)P*GP*TP*CP*TP*AP*C)-3')

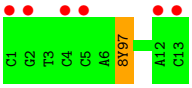
Chain E:



• Molecule 2: DNA (5'-D(*CP*GP*TP*CP*CP*AP*(DGH)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')

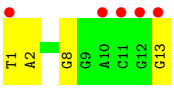
Chain C:

• 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	73.80Å 109.25Å 168.25Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.69 – 2.53 29.67 – 2.53	Depositor EDS
% Data completeness (in resolution range)	99.2 (29.69-2.53) 99.3 (29.67-2.53)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.57 (at 2.54Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.207 , 0.254 0.211 , 0.256	Depositor DCC
R_{free} test set	2326 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å^2)	57.1	Xtriage
Anisotropy	0.506	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 57.3	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	0.94	EDS
Total number of atoms	7518	wwPDB-VP
Average B, all atoms (Å^2)	80.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.33% 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: 8Y9, 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	0.78	0/2189	0.92	2/2959 (0.1%)
1	D	0.75	0/2119	0.92	3/2864 (0.1%)
1	G	0.79	0/1455	0.85	0/1935
2	B	0.48	0/264	0.76	0/402
2	E	0.50	0/264	0.76	0/402
2	H	0.36	0/264	0.74	0/402
3	C	0.53	0/300	0.82	0/462
3	F	0.56	0/300	0.84	0/462
3	I	0.48	0/300	0.80	0/462
All	All	0.72	0/7455	0.87	5/10350 (0.0%)

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	2
All	All	0	3

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	133	ARG	NE-CZ-NH2	-6.28	117.16	120.30
1	D	133	ARG	NE-CZ-NH2	-6.19	117.20	120.30
1	A	133	ARG	NE-CZ-NH1	5.84	123.22	120.30
1	D	169	ARG	NE-CZ-NH2	-5.63	117.48	120.30
1	D	119	ARG	CG-CD-NE	-5.59	100.07	111.80

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	105	PRO	Peptide
1	D	67	GLN	Peptide,Mainchain

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	12	0
1	D	2065	0	2038	13	0
1	G	1432	0	1221	19	0
2	B	260	0	137	0	0
2	E	260	0	137	0	0
2	H	260	0	137	1	0
3	C	267	0	147	2	0
3	F	267	0	147	1	0
3	I	267	0	147	4	0
4	B	6	0	8	0	0
4	D	6	0	8	0	0
4	E	6	0	8	0	0
5	A	130	0	0	1	0
5	B	18	0	0	1	0
5	C	7	0	0	0	0
5	D	100	0	0	1	0
5	E	15	0	0	0	0
5	F	10	0	0	0	0
5	G	6	0	0	0	0
5	H	2	0	0	0	0
5	I	3	0	0	0	0
All	All	7518	0	6237	48	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 (48) 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:G:3:GLN:OE1	1:G:175:GLY:CA	2.12	0.97
1:G:3:GLN:OE1	1:G:175:GLY:HA3	1.72	0.86
1:G:3:GLN:OE1	1:G:175:GLY:N	2.14	0.79
1:A:46:ARG:HD3	1:A:63:LEU:HD21	1.76	0.67
1:A:275:HIS:HB3	5:B:212:HOH:O	1.98	0.62
1:G:54:LYS:HE3	1:G:168:GLN:CD	2.25	0.57
1:G:263:TYR:OH	2:H:7:8Y9:OP1	2.12	0.56
1:D:275:HIS:HB3	5:D:475:HOH:O	2.08	0.54
3:C:1:DT:H2'	3:C:2:DA:C8	2.42	0.54
1:A:197:VAL:O	1:A:201:LEU:HD13	2.08	0.54
1:D:67:GLN:HA	1:D:67:GLN:OE1	2.08	0.54
1:G:99:LEU:HG	1:G:123:TRP:CZ2	2.45	0.52
1:G:191:PHE:HB3	1:G:288:LEU:H	1.75	0.52
1:G:14:VAL:HG11	1:G:49:ALA:HB2	1.92	0.51
1:D:130:GLN:OE1	1:D:133:ARG:HD2	2.11	0.51
3:I:1:DT:H2'	3:I:2:DA:C8	2.46	0.50
1:D:41:GLU:O	1:D:70:GLN:NE2	2.45	0.49
1:D:130:GLN:HB3	1:D:133:ARG:HD2	1.94	0.49
1:A:130:GLN:NE2	1:A:133:ARG:HH11	2.11	0.48
1:G:86:GLN:O	1:G:112:LEU:HD12	2.13	0.48
1:G:238:GLN:HE21	1:G:238:GLN:HA	1.78	0.48
1:D:130:GLN:HB3	1:D:133:ARG:CD	2.43	0.48
1:A:41:GLU:O	1:A:70:GLN:NE2	2.47	0.47
1:G:54:LYS:HE3	1:G:168:GLN:NE2	2.29	0.47
3:F:1:DT:H2'	3:F:2:DA:C8	2.50	0.47
1:G:238:GLN:HE21	1:G:238:GLN:CA	2.27	0.47
1:A:130:GLN:HB3	1:A:133:ARG:CD	2.45	0.46
1:A:46:ARG:CD	1:A:63:LEU:HD21	2.45	0.46
1:G:7:LEU:HD12	1:G:56:LEU:HB2	1.96	0.46
1:G:54:LYS:HE3	1:G:168:GLN:OE1	2.16	0.46
1:A:181:GLU:O	1:A:185:ARG:HG2	2.16	0.46
1:D:56:LEU:HD23	1:D:56:LEU:C	2.36	0.46
1:D:238:GLN:NE2	1:D:238:GLN:HA	2.31	0.45
1:G:117:ILE:HG21	3:I:8:DG:H5'	1.99	0.45
1:G:117:ILE:CG2	3:I:8:DG:H5'	2.46	0.45
1:D:181:GLU:O	1:D:185:ARG:HG2	2.16	0.45
1:D:105:PRO:HA	1:D:106:PRO:HA	1.67	0.43
1:G:191:PHE:O	1:G:287:PRO:N	2.51	0.43
1:A:56:LEU:C	1:A:56:LEU:HD23	2.39	0.42
1:A:130:GLN:HB3	1:A:133:ARG:HD2	2.01	0.42
1:D:186:LEU:HG	1:D:228:LEU:CD1	2.50	0.42
1:D:100:ARG:HD2	1:D:100:ARG:N	2.35	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:60:LEU:O	1:A:72:PRO:HA	2.21	0.41
1:A:143:GLN:HG3	5:A:347:HOH:O	2.21	0.40
1:G:99:LEU:HD23	1:G:99:LEU:HA	1.87	0.40
1:G:3:GLN:CD	1:G:175:GLY:HA3	2.38	0.40
3:C:1:DT:C6	3:I:13:DG:C4	3.10	0.40
1:D:90:ARG:HD3	1:D:109:ARG:NH1	2.36	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%)	254 (96%)	9 (3%)	2 (1%)	19	33
1	D	256/295 (87%)	245 (96%)	10 (4%)	1 (0%)	34	53
1	G	185/295 (63%)	162 (88%)	22 (12%)	1 (0%)	29	47
All	All	706/885 (80%)	661 (94%)	41 (6%)	4 (1%)	25	41

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	200	ALA
1	G	287	PRO
1	A	106	PRO
1	D	68	PRO

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%)	210 (95%)	12 (5%)	22	40
1	D	212/247 (86%)	206 (97%)	6 (3%)	43	68
1	G	124/247 (50%)	116 (94%)	8 (6%)	17	31
All	All	558/741 (75%)	532 (95%)	26 (5%)	26	46

All (26) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	34	ARG
1	A	46	ARG
1	A	95	ARG
1	A	100	ARG
1	A	150	ARG
1	A	155	LYS
1	A	227	GLU
1	A	238	GLN
1	A	239	LEU
1	A	246	SER
1	A	250	GLU
1	A	257	ARG
1	D	32	VAL
1	D	144	PHE
1	D	155	LYS
1	D	222	ASN
1	D	234	LYS
1	D	257	ARG
1	G	3	GLN
1	G	99	LEU
1	G	100	ARG
1	G	144	PHE
1	G	158	ASP
1	G	177	TYR
1	G	238	GLN
1	G	265	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	8	HIS
1	A	12	GLN
1	A	70	GLN
1	A	130	GLN
1	A	139	GLN
1	D	69	GLN
1	D	238	GLN
1	G	238	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)

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	8Y9	E	7	2	18,23,24	1.60	7 (38%)	20,33,36	3.61	5 (25%)
2	8Y9	B	7	2	18,23,24	1.55	4 (22%)	20,33,36	3.68	6 (30%)
2	8Y9	H	7	2	18,23,24	1.60	6 (33%)	20,33,36	3.62	5 (25%)

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	8Y9	E	7	2	-	2/10/41/42	0/2/2/2
2	8Y9	B	7	2	-	2/10/41/42	0/2/2/2
2	8Y9	H	7	2	-	2/10/41/42	0/2/2/2

All (17) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	7	8Y9	O5-C5	3.14	1.29	1.23
2	B	7	8Y9	C1'-N9	2.86	1.49	1.45
2	H	7	8Y9	C1'-N9	2.81	1.49	1.45
2	E	7	8Y9	C1'-N9	2.71	1.49	1.45
2	E	7	8Y9	O8-C8	2.67	1.27	1.23
2	H	7	8Y9	O5-C5	2.66	1.28	1.23
2	H	7	8Y9	O8-C8	2.60	1.27	1.23
2	E	7	8Y9	O5-C5	2.56	1.28	1.23
2	H	7	8Y9	C8-N9	-2.54	1.33	1.37
2	E	7	8Y9	C8-N9	-2.31	1.34	1.37
2	H	7	8Y9	C5-N7	-2.29	1.34	1.37
2	B	7	8Y9	C8-N9	-2.28	1.34	1.37
2	E	7	8Y9	C2-N11	-2.19	1.25	1.34
2	E	7	8Y9	C5-N7	-2.18	1.34	1.37
2	B	7	8Y9	O8-C8	2.12	1.26	1.23
2	H	7	8Y9	C2-N11	-2.04	1.26	1.34
2	E	7	8Y9	C4-N3	2.04	1.46	1.44

All (16) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	E	7	8Y9	C4-C5-N7	12.77	122.52	106.88
2	B	7	8Y9	C4-C5-N7	12.63	122.35	106.88
2	H	7	8Y9	C4-C5-N7	12.60	122.32	106.88
2	B	7	8Y9	C5-C4-N9	-8.65	90.78	102.28
2	H	7	8Y9	C5-C4-N9	-8.46	91.04	102.28
2	E	7	8Y9	C5-C4-N9	-8.22	91.35	102.28
2	B	7	8Y9	C5-C4-N3	-3.40	105.55	112.76
2	H	7	8Y9	O5-C5-N7	-3.23	121.03	124.94
2	E	7	8Y9	O5-C5-N7	-2.93	121.39	124.94
2	E	7	8Y9	C5-C4-N3	-2.80	106.80	112.76
2	H	7	8Y9	N3-C2-N12	2.63	125.18	120.59
2	H	7	8Y9	C5-C4-N3	-2.59	107.26	112.76
2	B	7	8Y9	O5-C5-N7	-2.52	121.89	124.94
2	B	7	8Y9	C2'-C1'-N9	2.35	118.77	115.59
2	B	7	8Y9	N3-C2-N12	2.05	124.17	120.59
2	E	7	8Y9	N3-C2-N12	2.00	124.09	120.59

There are no chirality outliers.

All (6) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	E	7	8Y9	O4'-C4'-C5'-O5'

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
2	B	7	8Y9	O4'-C4'-C5'-O5'
2	H	7	8Y9	O4'-C4'-C5'-O5'
2	E	7	8Y9	C3'-C4'-C5'-O5'
2	H	7	8Y9	C3'-C4'-C5'-O5'
2	B	7	8Y9	C3'-C4'-C5'-O5'

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	H	7	8Y9	1	0

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

3 ligands 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	GOL	B	101	-	5,5,5	0.13	0	5,5,5	0.32	0
4	GOL	E	101	-	5,5,5	0.13	0	5,5,5	0.34	0
4	GOL	D	301	-	5,5,5	0.21	0	5,5,5	0.51	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	GOL	B	101	-	-	2/4/4/4	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	GOL	E	101	-	-	4/4/4/4	-
4	GOL	D	301	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (8) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	E	101	GOL	O1-C1-C2-C3
4	B	101	GOL	C1-C2-C3-O3
4	D	301	GOL	C1-C2-C3-O3
4	E	101	GOL	C1-C2-C3-O3
4	B	101	GOL	O2-C2-C3-O3
4	D	301	GOL	O2-C2-C3-O3
4	E	101	GOL	O1-C1-C2-O2
4	E	101	GOL	O2-C2-C3-O3

There are no ring outliers.

No monomer is involved in short contacts.

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%)	0.33	17 (6%) 20 21	34, 57, 96, 131	0
1	D	262/295 (88%)	0.33	10 (3%) 40 44	38, 65, 110, 128	0
1	G	205/295 (69%)	0.71	22 (10%) 6 6	73, 106, 133, 150	0
2	B	12/13 (92%)	0.91	3 (25%) 0 0	58, 96, 131, 139	0
2	E	12/13 (92%)	0.85	2 (16%) 1 1	59, 89, 118, 136	0
2	H	12/13 (92%)	1.75	6 (50%) 0 0	94, 119, 142, 146	0
3	C	13/13 (100%)	0.55	0 100 100	75, 90, 138, 159	0
3	F	13/13 (100%)	0.68	0 100 100	54, 85, 134, 140	0
3	I	13/13 (100%)	1.38	5 (38%) 0 0	78, 118, 140, 151	0
All	All	811/963 (84%)	0.49	65 (8%) 12 13	34, 75, 128, 159	0

All (65) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	125	LEU	6.5
1	A	246	SER	4.8
2	E	1	DC	4.8
1	D	243	GLY	4.6
1	A	248	SER	4.5
1	G	102	TYR	4.3
1	G	264	GLY	3.9
1	G	269	SER	3.9
1	G	285	PRO	3.7
1	A	249	GLY	3.6
1	A	166	LEU	3.6
1	A	2	GLY	3.5
2	H	13	DC	3.5
2	H	1	DC	3.3
2	H	2	DG	3.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	B	13	DC	3.3
1	A	247	GLU	3.3
3	I	11	DC	3.1
1	A	7	LEU	3.1
2	H	12	DA	3.1
1	G	101	PHE	3.1
2	B	1	DC	3.0
1	G	58	LEU	2.9
2	H	4	DC	2.9
3	I	13	DG	2.9
1	D	178	LEU	2.9
1	G	122	ARG	2.9
1	G	286	GLY	2.8
1	G	271	LEU	2.8
1	G	75	LEU	2.8
1	A	165	LEU	2.8
1	G	267	GLY	2.7
1	G	254	ALA	2.7
1	A	85	PHE	2.7
1	G	288	LEU	2.7
1	D	222	ASN	2.6
1	G	278	THR	2.6
2	B	12	DA	2.5
1	A	3	GLN	2.5
1	G	270	SER	2.5
1	G	268	MET	2.4
1	G	100	ARG	2.4
1	A	161	ILE	2.4
3	I	12	DG	2.4
1	D	3	GLN	2.3
3	I	1	DT	2.3
1	G	156	ALA	2.3
1	A	4	GLY	2.3
1	D	174	ILE	2.3
1	A	250	GLU	2.2
1	G	283	GLY	2.2
1	A	107	GLY	2.2
1	G	189	PRO	2.2
2	E	13	DC	2.2
1	A	173	GLY	2.2
1	D	173	GLY	2.2
1	D	175	GLY	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	229	CYS	2.1
1	G	188	ILE	2.1
1	D	180	ALA	2.1
3	I	10	DA	2.1
2	H	5	DC	2.1
1	A	251	GLU	2.1
1	A	54	LYS	2.1
1	D	82	SER	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	8Y9	H	7	22/23	0.96	0.13	94,99,104,105	0
2	8Y9	E	7	22/23	0.97	0.20	57,65,84,88	0
2	8Y9	B	7	22/23	0.97	0.19	52,59,79,87	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
4	GOL	D	301	6/6	0.64	0.36	66,71,75,75	0
4	GOL	E	101	6/6	0.75	0.34	81,86,95,100	0
4	GOL	B	101	6/6	0.79	0.31	93,97,101,102	0

6.5 Other polymers (i)

There are no such residues in this entry.