Full wwPDB X-ray Structure Validation Report (i)

Jul 27, 2016 – 04:02 PM EDT

PDB ID	:	5ITX
Title	:	Crystal Structure of Human NEIL1(P2G R242K) bound to duplex DNA containing Thymine Glycol
Authors	:	Zhu, C.; Lu, L.; Zhang, J.; Yue, Z.; Song, J.; Zong, S.; Liu, M.; Stovicek, O.; Gao, Y.; Yi, C.
Deposited on	:	2016-03-17
Resolution	:	2.65 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
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

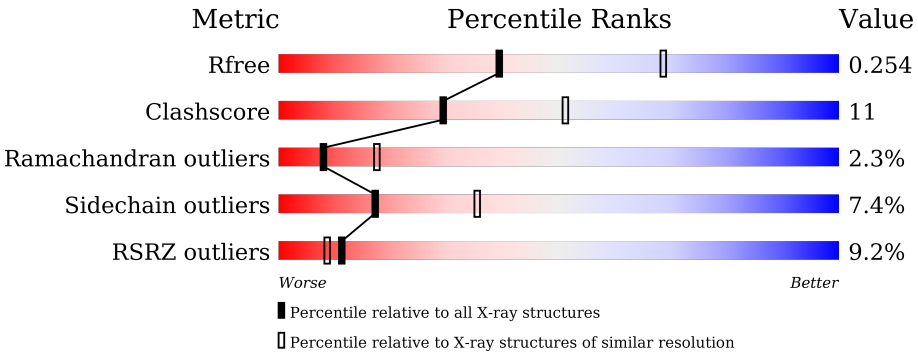
MolProbity	:	4.02b-467
Mogul	:	1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20027939
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20027939

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 2.65 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}	91344	3152 (2.70-2.62)
Clashscore	102246	3524 (2.70-2.62)
Ramachandran outliers	100387	3469 (2.70-2.62)
Sidechain outliers	100360	3469 (2.70-2.62)
RSRZ outliers	91569	3161 (2.70-2.62)

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 on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. 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	390	57% 9% •• 32%
1	B	390	% 47% 15% 5% 33%
2	C	26	8% 58% 38% •
2	D	26	4% 65% 27% 8%
2	F	26	12% 73% 27%
3	E	400	17% 47% 14% • • 35%

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 8021 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	264	Total	C	N	O	S	0	0	0
			2099	1342	383	364	10			
1	B	263	Total	C	N	O	S	0	1	0
			2089	1335	382	363	9			

- Molecule 2 is a DNA chain called DNA (26-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	26	Total	C	N	O	P	0	0	0
			527	252	96	155	24			
2	C	26	Total	C	N	O	P	0	0	0
			527	252	96	155	24			
2	F	26	Total	C	N	O	P	0	0	0
			527	252	96	155	24			

- Molecule 3 is a protein called Endonuclease 8-like 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	260	Total	C	N	O	S	0	1	0
			2069	1321	379	359	10			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	2	GLY	PRO	engineered mutation	UNP Q96FI4
E	242	ARG	LYS	engineered mutation	UNP Q96FI4
E	391	ALA	-	expression tag	UNP Q96FI4
E	392	ALA	-	expression tag	UNP Q96FI4
E	393	LEU	-	expression tag	UNP Q96FI4
E	394	GLY	-	expression tag	UNP Q96FI4
E	395	HIS	-	expression tag	UNP Q96FI4
E	396	HIS	-	expression tag	UNP Q96FI4

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	397	HIS	-	expression tag	UNP Q96FI4
E	398	HIS	-	expression tag	UNP Q96FI4
E	399	HIS	-	expression tag	UNP Q96FI4
E	400	HIS	-	expression tag	UNP Q96FI4

- Molecule 4 is water.

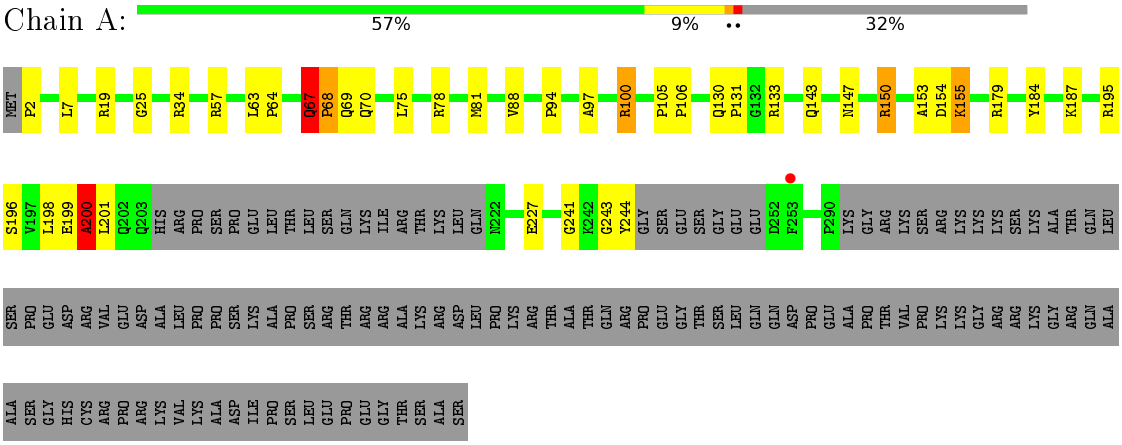
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	113	Total O 113 113	0	0
4	D	11	Total O 11 11	0	0
4	B	49	Total O 49 49	0	0
4	C	6	Total O 6 6	0	0
4	E	3	Total O 3 3	0	0
4	F	1	Total O 1 1	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

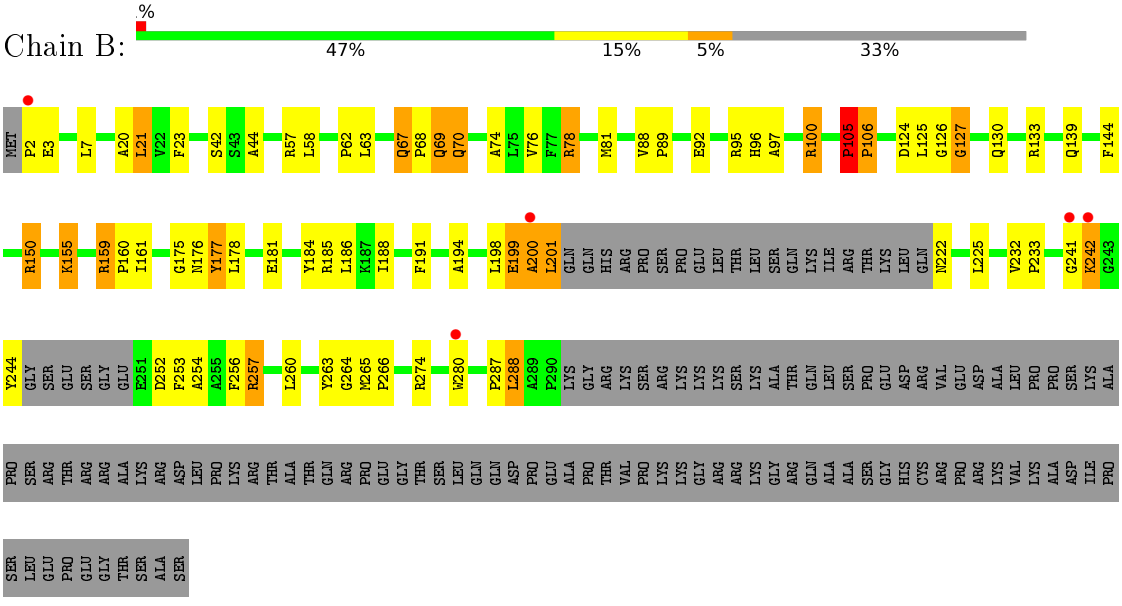
- Molecule 1: Endonuclease 8-like 1

Chain A:



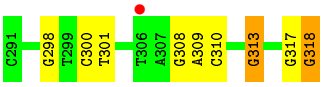
- Molecule 1: Endonuclease 8-like 1

Chain B:



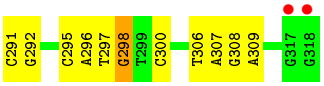
- Molecule 2: DNA (26-MER)

Chain D:



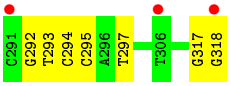
• Molecule 2: DNA (26-MER)

Chain C:



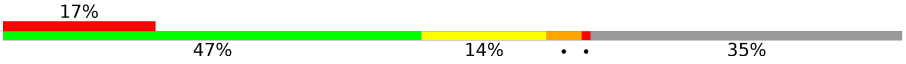
• Molecule 2: DNA (26-MER)

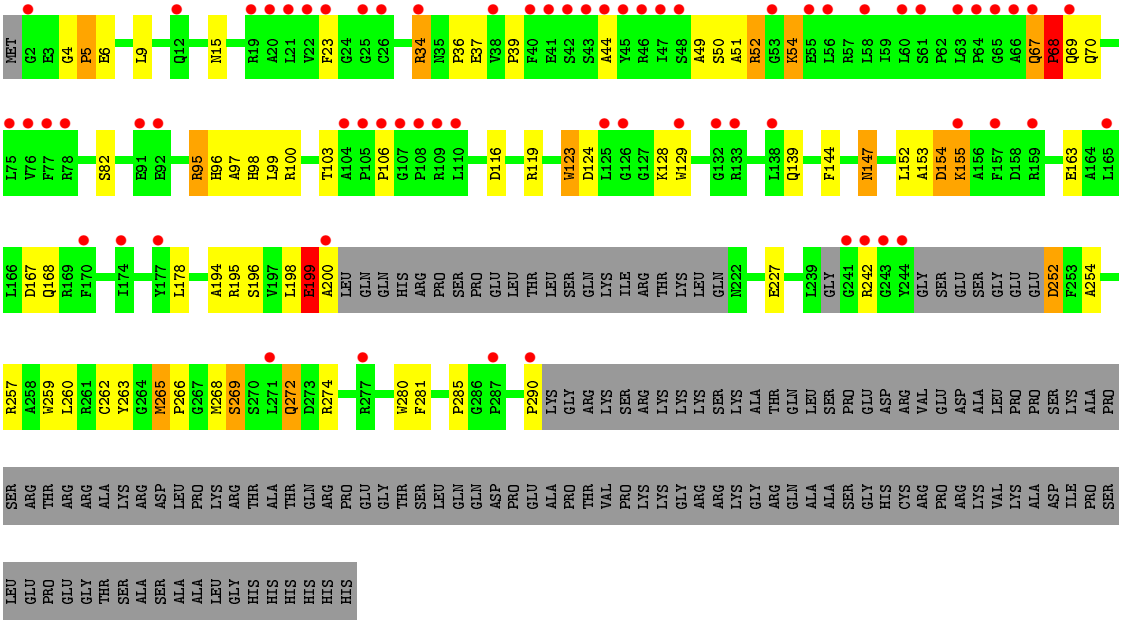
Chain F:



• Molecule 3: Endonuclease 8-like 1

Chain E:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	73.99Å 109.38Å 170.97Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	92.14 – 2.65 49.81 – 2.65	Depositor EDS
% Data completeness (in resolution range)	99.2 (92.14-2.65) 99.2 (49.81-2.65)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	0.06	Depositor
$< I/\sigma(I) >$ [1]	2.26 (at 2.65Å)	Xtriage
Refinement program	REFMAC 5.8.0073	Depositor
R, R_{free}	0.192 , 0.235 0.232 , 0.254	Depositor DCC
R_{free} test set	2022 reflections (4.96%)	DCC
Wilson B-factor (Å^2)	62.9	Xtriage
Anisotropy	0.240	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 46.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	8021	wwPDB-VP
Average B, all atoms (Å^2)	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.25% 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: CTG

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	0.97	1/2156 (0.0%)	1.03	5/2915 (0.2%)
1	B	0.90	0/2145	1.01	6/2901 (0.2%)
2	C	0.70	1/564 (0.2%)	1.22	2/864 (0.2%)
2	D	0.59	0/564	1.12	6/864 (0.7%)
2	F	0.40	0/564	0.78	0/864
3	E	0.51	1/2124 (0.0%)	0.72	1/2871 (0.0%)
All	All	0.77	3/8117 (0.0%)	0.96	20/11279 (0.2%)

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	4
3	E	0	1
All	All	0	5

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	298	DG	O3'-P	6.35	1.68	1.61
1	A	227	GLU	CD-OE1	5.14	1.31	1.25
3	E	5	PRO	N-CD	5.07	1.54	1.47

All (20) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	298	DG	O5'-P-OP1	-22.97	83.14	110.70
2	D	298	DG	O5'-P-OP1	-12.03	94.87	105.70

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	57	ARG	NE-CZ-NH1	10.18	125.39	120.30
1	B	57	ARG	NE-CZ-NH2	-10.00	115.30	120.30
2	C	298	DG	O5'-P-OP2	8.90	121.38	110.70
1	A	57	ARG	NE-CZ-NH2	-7.98	116.31	120.30
2	D	313	DG	O5'-P-OP1	-7.95	98.54	105.70
1	A	67	GLN	C-N-CD	-6.94	105.34	120.60
1	B	78	ARG	CG-CD-NE	-6.71	97.72	111.80
2	D	298	DG	O5'-P-OP2	6.68	118.72	110.70
2	D	313	DG	O5'-P-OP2	6.40	118.39	110.70
1	A	150	ARG	NE-CZ-NH1	6.34	123.47	120.30
2	D	310	DC	C1'-O4'-C4'	-5.70	104.40	110.10
3	E	4	GLY	C-N-CD	5.67	140.31	128.40
1	B	58	LEU	CA-CB-CG	5.61	128.20	115.30
2	D	318	DG	O5'-P-OP1	-5.39	100.85	105.70
1	B	67	GLN	C-N-CD	5.17	139.26	128.40
1	A	78	ARG	NE-CZ-NH2	-5.10	117.75	120.30
1	A	179	ARG	NE-CZ-NH1	5.03	122.81	120.30
1	B	78	ARG	NE-CZ-NH2	-5.00	117.80	120.30

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	105	PRO	Peptide
1	A	200	ALA	Peptide
1	A	243	GLY	Peptide
1	A	67	GLN	Peptide
3	E	67	GLN	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	2099	0	2076	26	0
1	B	2089	0	2061	74	0
2	C	527	0	298	12	0
2	D	527	0	298	4	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	F	527	0	298	6	0
3	E	2069	0	2042	56	0
4	A	113	0	0	6	1
4	B	49	0	0	1	0
4	C	6	0	0	0	0
4	D	11	0	0	1	0
4	E	3	0	0	0	0
4	F	1	0	0	0	0
All	All	8021	0	7073	168	1

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 11.

All (168) 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:B:3:GLU:OE1	1:B:175:GLY:CA	1.73	1.36
1:B:3:GLU:OE1	1:B:175:GLY:C	1.63	1.33
1:A:19:ARG:NH1	1:B:201:LEU:HD13	1.53	1.23
1:A:19:ARG:HH12	1:B:201:LEU:CD1	1.61	1.14
3:E:262:CYS:O	3:E:265:MET:HB2	1.46	1.13
3:E:6:GLU:OE1	3:E:6:GLU:N	1.89	1.06
1:B:3:GLU:OE1	1:B:175:GLY:HA3	1.53	1.05
1:B:263:TYR:OH	2:C:297:CTG:OP1	1.75	1.04
1:B:242:LYS:H	1:B:252:ASP:CG	1.62	1.03
3:E:194:ALA:O	3:E:198:LEU:CD2	2.07	1.01
1:B:76:VAL:HG23	1:B:127:GLY:HA2	1.45	0.96
3:E:195:ARG:O	3:E:199:GLU:HG2	1.65	0.95
1:A:19:ARG:HH12	1:B:201:LEU:HD13	0.76	0.91
3:E:263:TYR:OH	2:F:297:CTG:OP1	1.90	0.90
1:B:155:LYS:H	1:B:155:LYS:HD3	1.35	0.89
3:E:194:ALA:O	3:E:198:LEU:HD22	1.73	0.88
1:B:242:LYS:HA	1:B:252:ASP:CB	2.09	0.82
1:B:3:GLU:CD	1:B:175:GLY:HA3	1.99	0.81
3:E:262:CYS:O	3:E:265:MET:CB	2.28	0.81
1:B:3:GLU:CG	1:B:175:GLY:HA3	2.12	0.80
3:E:194:ALA:O	3:E:198:LEU:HD23	1.79	0.79
1:B:155:LYS:H	1:B:155:LYS:CD	1.97	0.78
3:E:95:ARG:O	3:E:96:HIS:HB2	1.83	0.77
2:F:317:DG:H2"	2:F:318:DG:OP2	1.84	0.77
1:B:242:LYS:N	1:B:252:ASP:CG	2.39	0.76

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:69:GLN:OE1	1:B:70:GLN:N	2.19	0.76
1:B:200:ALA:O	1:B:201:LEU:HB2	1.86	0.75
3:E:198:LEU:HD22	3:E:198:LEU:H	1.53	0.74
1:B:3:GLU:OE1	1:B:176:ASN:N	2.19	0.74
3:E:50:SER:OG	3:E:51:ALA:N	2.21	0.73
3:E:152:LEU:HD21	3:E:198:LEU:HB3	1.70	0.72
2:D:313:DG:N3	4:D:401:HOH:O	2.23	0.71
3:E:265:MET:O	3:E:268:MET:HB2	1.90	0.71
1:B:78:ARG:NH1	2:C:300:DC:OP1	2.24	0.71
3:E:195:ARG:O	3:E:199:GLU:CG	2.41	0.68
1:B:76:VAL:CG2	1:B:127:GLY:HA2	2.24	0.67
1:A:199:GLU:O	1:A:200:ALA:HB3	1.95	0.67
1:B:74:ALA:O	1:B:126:GLY:HA2	1.95	0.66
1:B:3:GLU:OE2	1:B:177:TYR:CD2	2.48	0.66
1:B:242:LYS:HA	1:B:252:ASP:HB2	1.78	0.65
1:B:260:LEU:HD22	1:B:263:TYR:HB3	1.79	0.65
1:A:199:GLU:O	1:A:200:ALA:CB	2.45	0.64
3:E:69:GLN:HG2	3:E:69:GLN:O	1.99	0.62
1:B:3:GLU:CD	1:B:175:GLY:CA	2.60	0.62
1:B:264:GLY:HA2	1:B:280:TRP:CH2	2.35	0.61
1:A:154:ASP:OD1	1:A:155:LYS:N	2.33	0.61
3:E:196:SER:HA	3:E:199:GLU:HG3	1.83	0.61
1:A:97:ALA:O	1:A:100:ARG:NH1	2.34	0.61
2:D:300:DC:H1'	2:D:301:DT:H5'	1.82	0.61
1:B:257:ARG:HA	1:B:260:LEU:HD12	1.81	0.60
3:E:98:HIS:NE2	3:E:119:ARG:HD3	2.16	0.60
1:B:81:MET:HE1	2:C:297:CTG:H5'	1.83	0.60
1:B:160:PRO:HB3	1:B:191:PHE:HA	1.86	0.58
3:E:198:LEU:O	3:E:200:ALA:N	2.35	0.58
1:A:184:TYR:O	1:A:187:LYS:HE2	2.04	0.57
3:E:263:TYR:CD1	3:E:280:TRP:CD1	2.93	0.57
2:C:308:DG:H2"	2:C:309:DA:H5'	1.87	0.57
1:B:3:GLU:HG2	1:B:175:GLY:N	2.21	0.56
3:E:153:ALA:O	3:E:154:ASP:O	2.24	0.56
3:E:242:ARG:HG2	3:E:252:ASP:HB3	1.86	0.56
3:E:269:SER:HB3	3:E:281:PHE:CE1	2.40	0.56
2:C:306:DT:H2'	2:C:307:DA:C8	2.40	0.56
3:E:51:ALA:O	3:E:52:ARG:HB2	2.05	0.56
1:B:3:GLU:OE1	1:B:175:GLY:N	2.36	0.56
3:E:95:ARG:O	3:E:96:HIS:CB	2.48	0.56
2:F:293:DT:C2'	2:F:294:DC:O5'	2.55	0.55

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:308:DG:H2"	2:D:309:DA:H5'	1.87	0.55
2:D:317:DG:H2"	2:D:318:DG:OP2	2.07	0.55
3:E:99:LEU:HD22	3:E:123:TRP:CE2	2.42	0.54
3:E:263:TYR:HD1	3:E:280:TRP:CD1	2.25	0.54
1:B:161:ILE:HD13	1:B:194:ALA:HA	1.90	0.53
1:A:147:ASN:HD22	1:A:150:ARG:NH1	2.07	0.53
3:E:82:SER:OG	3:E:116:ASP:OD1	2.23	0.52
1:B:67:GLN:HA	1:B:68:PRO:C	2.30	0.52
2:C:307:DA:H2"	2:C:308:DG:O5'	2.09	0.52
1:B:191:PHE:O	1:B:288:LEU:HD12	2.10	0.51
3:E:54:LYS:HD2	3:E:168:GLN:OE1	2.09	0.51
1:B:232:VAL:HB	1:B:233:PRO:CD	2.40	0.51
3:E:95:ARG:HG2	3:E:96:HIS:CD2	2.46	0.51
1:B:155:LYS:N	1:B:155:LYS:HD3	2.15	0.50
1:B:159:ARG:HD2	1:B:274:ARG:HH22	1.76	0.50
3:E:98:HIS:CD2	3:E:119:ARG:HG2	2.46	0.50
1:B:242:LYS:N	1:B:252:ASP:OD2	2.43	0.50
1:A:241:GLY:O	1:A:244:TYR:HB2	2.12	0.49
3:E:152:LEU:CD2	3:E:198:LEU:HB3	2.40	0.49
1:A:198:LEU:O	1:A:201:LEU:HB2	2.13	0.49
1:A:184:TYR:O	1:A:187:LYS:CE	2.60	0.49
1:A:67:GLN:C	1:A:68:PRO:O	2.50	0.49
2:F:293:DT:H2"	2:F:294:DC:O5'	2.12	0.49
1:B:222:ASN:N	1:B:222:ASN:OD1	2.45	0.49
3:E:95:ARG:HB3	3:E:96:HIS:HD2	1.77	0.49
1:B:105:PRO:HB2	1:B:106:PRO:HD3	1.93	0.49
2:C:306:DT:H2'	2:C:307:DA:N7	2.27	0.49
1:A:130:GLN:NE2	1:A:133:ARG:HE	2.10	0.49
2:C:295:DC:H2"	2:C:296:DA:C8	2.48	0.49
1:B:97:ALA:O	1:B:100:ARG:NH1	2.44	0.49
1:A:63:LEU:HB3	1:A:64:PRO:CD	2.42	0.48
1:B:95:ARG:O	1:B:96:HIS:HB2	2.13	0.48
1:B:67:GLN:HB3	1:B:68:PRO:HA	1.94	0.48
1:B:242:LYS:HA	1:B:252:ASP:CG	2.34	0.48
3:E:154:ASP:OD1	3:E:155:LYS:N	2.46	0.48
1:B:88:VAL:HB	1:B:89:PRO:HD2	1.95	0.47
1:A:153:ALA:CB	4:A:406:HOH:O	2.61	0.47
1:A:195:ARG:HD3	4:A:462:HOH:O	2.13	0.47
3:E:198:LEU:H	3:E:198:LEU:CD2	2.26	0.47
1:B:21:LEU:N	1:B:21:LEU:HD23	2.29	0.47
1:B:3:GLU:CG	1:B:175:GLY:CA	2.88	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:159:ARG:CB	1:B:160:PRO:CD	2.92	0.47
1:B:181:GLU:O	1:B:185:ARG:HG3	2.15	0.47
1:B:78:ARG:NH2	1:B:124:ASP:OD2	2.47	0.47
1:B:198:LEU:HD21	1:B:225:LEU:HD23	1.96	0.47
1:B:155:LYS:HE3	1:B:155:LYS:HB2	1.68	0.47
3:E:97:ALA:O	3:E:100:ARG:NH1	2.48	0.46
1:B:265:MET:HB3	1:B:266:PRO:HD2	1.97	0.46
3:E:15:ASN:ND2	3:E:49:ALA:O	2.49	0.46
2:F:294:DC:H2"	2:F:295:DC:C6	2.51	0.46
1:B:2:PRO:N	2:C:297:CTG:O2	2.49	0.46
1:A:143:GLN:HG3	4:A:438:HOH:O	2.15	0.46
1:B:244:TYR:HB3	4:B:434:HOH:O	2.16	0.45
1:B:67:GLN:HA	1:B:68:PRO:O	2.16	0.45
1:A:69:GLN:HB3	4:A:480:HOH:O	2.15	0.45
1:B:256:PHE:CE2	1:B:260:LEU:HD11	2.52	0.45
3:E:37:GLU:O	3:E:39:PRO:HD3	2.17	0.45
3:E:67:GLN:HB3	3:E:68:PRO:CD	2.46	0.44
3:E:5:PRO:HD3	3:E:178:LEU:HD11	1.99	0.44
1:B:186:LEU:O	1:B:188:ILE:N	2.51	0.44
1:A:19:ARG:HG3	1:B:150:ARG:HG3	1.98	0.43
1:A:88:VAL:HG21	1:A:94:PRO:HD3	2.00	0.43
1:B:253:PHE:O	1:B:254:ALA:C	2.57	0.43
3:E:5:PRO:HB2	3:E:6:GLU:OE1	2.19	0.43
1:B:242:LYS:CA	1:B:252:ASP:CG	2.86	0.43
3:E:34:ARG:CZ	3:E:34:ARG:HB2	2.48	0.43
3:E:98:HIS:NE2	3:E:119:ARG:CD	2.79	0.43
2:F:292:DG:H2'	2:F:293:DT:C6	2.54	0.43
1:B:199:GLU:O	1:B:200:ALA:HB2	2.18	0.43
3:E:163:GLU:OE1	3:E:274:ARG:NH1	2.52	0.43
3:E:198:LEU:HD22	3:E:198:LEU:N	2.27	0.43
1:B:198:LEU:CD2	1:B:225:LEU:HD23	2.49	0.42
2:C:291:DC:H2'	2:C:292:DG:C8	2.54	0.42
3:E:252:ASP:C	3:E:254:ALA:N	2.73	0.42
1:A:147:ASN:HD22	1:A:150:ARG:CZ	2.32	0.42
1:A:2:PRO:CA	4:A:450:HOH:O	2.67	0.42
3:E:262:CYS:O	3:E:265:MET:CG	2.67	0.42
3:E:272:GLN:HA	3:E:272:GLN:HE21	1.84	0.42
1:B:201:LEU:HA	1:B:201:LEU:HD22	1.93	0.42
2:C:306:DT:C2'	2:C:307:DA:C8	3.03	0.42
3:E:281:PHE:N	3:E:281:PHE:CD1	2.88	0.42
1:B:20:ALA:C	1:B:21:LEU:HD23	2.40	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:89:PRO:HG2	1:B:92:GLU:HB2	2.02	0.41
1:A:155:LYS:HB2	1:A:155:LYS:HE2	1.90	0.41
3:E:252:ASP:C	3:E:254:ALA:H	2.24	0.41
1:B:178:LEU:HD23	1:B:178:LEU:HA	1.88	0.41
1:B:184:TYR:CG	1:B:184:TYR:O	2.73	0.41
3:E:262:CYS:O	3:E:265:MET:HG2	2.20	0.41
3:E:98:HIS:CD2	3:E:119:ARG:HD3	2.54	0.41
3:E:259:TRP:CZ3	3:E:260:LEU:HD23	2.55	0.41
3:E:52:ARG:HD2	3:E:129:TRP:CE2	2.55	0.41
1:B:175:GLY:HA3	2:C:298:DG:OP1	2.20	0.41
1:B:62:PRO:HG2	1:B:70:GLN:HB2	2.02	0.41
1:B:95:ARG:O	1:B:96:HIS:CB	2.66	0.41
1:A:130:GLN:HA	1:A:131:PRO:HD3	1.85	0.41
3:E:23:PHE:HA	3:E:103:THR:HA	2.03	0.41
1:B:241:GLY:O	1:B:242:LYS:HB2	2.22	0.40
3:E:9:LEU:HD23	3:E:9:LEU:HA	1.87	0.40
1:B:130:GLN:NE2	1:B:133:ARG:HE	2.19	0.40
1:B:23:PHE:O	1:B:44:ALA:HA	2.22	0.40
3:E:147:ASN:HD22	3:E:147:ASN:C	2.25	0.40
1:A:25:GLY:HA2	4:A:459:HOH:O	2.21	0.40

All (1) 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)
4:A:457:HOH:O	4:A:496:HOH:O[3_554]	2.12	0.08

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	258/390 (66%)	247 (96%)	8 (3%)	3 (1%)	16 35

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	B	258/390 (66%)	240 (93%)	13 (5%)	5 (2%)	10 22
3	E	253/400 (63%)	219 (87%)	24 (10%)	10 (4%)	4 7
All	All	769/1180 (65%)	706 (92%)	45 (6%)	18 (2%)	8 17

All (18) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	68	PRO
1	A	106	PRO
1	B	200	ALA
3	E	52	ARG
3	E	68	PRO
3	E	154	ASP
1	B	127	GLY
1	B	242	LYS
3	E	199	GLU
1	B	105	PRO
3	E	36	PRO
3	E	167	ASP
1	A	200	ALA
3	E	44	ALA
1	B	287	PRO
3	E	106	PRO
3	E	266	PRO
3	E	285	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	219/328 (67%)	211 (96%)	8 (4%)	41 69
1	B	216/328 (66%)	196 (91%)	20 (9%)	11 23
3	E	215/334 (64%)	195 (91%)	20 (9%)	11 23
All	All	650/990 (66%)	602 (93%)	48 (7%)	17 35

All (48) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	7	LEU
1	A	34	ARG
1	A	70	GLN
1	A	75	LEU
1	A	81	MET
1	A	100	ARG
1	A	155	LYS
1	A	196	SER
1	B	7	LEU
1	B	21	LEU
1	B	42	SER
1	B	63	LEU
1	B	69	GLN
1	B	70	GLN
1	B	100	ARG
1	B	105	PRO
1	B	106	PRO
1	B	125	LEU
1	B	139	GLN
1	B	144	PHE
1	B	150	ARG
1	B	155	LYS
1	B	159	ARG
1	B	177	TYR
1	B	199	GLU
1	B	201	LEU
1	B	257	ARG
1	B	288	LEU
3	E	34	ARG
3	E	54	LYS
3	E	68	PRO
3	E	70	GLN
3	E	95	ARG
3	E	123	TRP
3	E	124	ASP
3	E	128	LYS
3	E	139	GLN
3	E	144	PHE
3	E	147	ASN
3	E	155	LYS
3	E	199	GLU
3	E	227	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	E	252	ASP
3	E	257	ARG
3	E	265	MET
3	E	269	SER
3	E	272	GLN
3	E	290	PRO

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (10) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	70	GLN
1	A	130	GLN
1	A	142	GLN
1	A	147	ASN
1	B	70	GLN
1	B	130	GLN
1	B	222	ASN
1	B	238	GLN
1	B	272	GLN
3	E	272	GLN

5.3.3 RNA

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains

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	CTG	C	297	2	16,23,24	1.30	2 (12%)	18,35,38	1.44	1 (5%)
2	CTG	D	297	2	16,23,24	1.43	3 (18%)	18,35,38	1.37	2 (11%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
2	CTG	F	297	2	16,23,24	0.86	1 (6%)	18,35,38	1.76	3 (16%)

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	CTG	C	297	2	-	0/7/45/46	0/2/2/2
2	CTG	D	297	2	-	0/7/45/46	0/2/2/2
2	CTG	F	297	2	-	0/7/45/46	0/2/2/2

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	297	CTG	C2-N3	-2.74	1.32	1.38
2	C	297	CTG	C2-N3	-2.30	1.33	1.38
2	C	297	CTG	C4-N3	-2.13	1.34	1.37
2	D	297	CTG	O4'-C1'	2.07	1.47	1.42
2	F	297	CTG	O6-C6	2.11	1.44	1.40
2	D	297	CTG	O5-C5	2.42	1.47	1.43

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	F	297	CTG	C2'-C1'-N1	-5.64	107.92	115.64
2	C	297	CTG	C2'-C1'-N1	-4.64	109.28	115.64
2	F	297	CTG	O5-C5-C5M	-2.66	103.76	109.09
2	D	297	CTG	C3'-C2'-C1'	-2.28	96.86	102.40
2	F	297	CTG	C4-N3-C2	-2.16	123.08	126.83
2	D	297	CTG	N3-C2-N1	2.74	119.50	116.82

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	297	CTG	3	0
2	F	297	CTG	1	0

5.5 Carbohydrates

There are no carbohydrates in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

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	264/390 (67%)	0.07	1 (0%) 93 94	2, 11, 42, 92	0
1	B	263/390 (67%)	0.36	5 (1%) 70 69	4, 21, 56, 108	0
2	C	25/26 (96%)	0.41	2 (8%) 15 12	12, 43, 56, 63	0
2	D	25/26 (96%)	0.10	1 (4%) 42 40	18, 42, 71, 91	0
2	F	25/26 (96%)	0.71	3 (12%) 6 4	44, 73, 86, 93	0
3	E	260/400 (65%)	1.33	67 (25%) 1 0	40, 74, 101, 169	0
All	All	862/1258 (68%)	0.57	79 (9%) 11 9	2, 31, 89, 169	0

All (79) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	E	75	LEU	9.6
3	E	47	ILE	7.9
3	E	44	ALA	6.9
3	E	106	PRO	6.6
3	E	22	VAL	6.5
3	E	107	GLY	6.5
3	E	64	PRO	6.3
3	E	67	GLN	5.8
3	E	65	GLY	5.8
3	E	60	LEU	5.6
3	E	46	ARG	5.6
3	E	63	LEU	5.6
3	E	58	LEU	5.3
3	E	66	ALA	5.0
3	E	41	GLU	4.9
3	E	43	SER	4.8
3	E	23	PHE	4.7
3	E	241	GLY	4.6
1	B	242	LYS	4.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
3	E	21	LEU	4.4
3	E	105	PRO	4.3
3	E	129	TRP	4.2
3	E	61	SER	4.1
3	E	125	LEU	4.1
3	E	159	ARG	4.0
3	E	109	ARG	4.0
3	E	55	GLU	3.8
3	E	133	ARG	3.7
3	E	243	GLY	3.8
3	E	108	PRO	3.7
1	B	2	PRO	3.7
3	E	2	GLY	3.7
3	E	48	SER	3.7
3	E	244	TYR	3.7
3	E	77	PHE	3.7
3	E	40	PHE	3.5
2	D	306	DT	3.5
3	E	200	ALA	3.4
3	E	26	CYS	3.3
3	E	78	ARG	3.1
3	E	12	GLN	3.0
2	F	306	DT	3.0
3	E	110	LEU	3.0
3	E	69	GLN	3.0
3	E	126	GLY	2.9
2	C	318	DG	2.9
3	E	92	GLU	2.9
2	F	318	DG	2.8
3	E	38	VAL	2.8
3	E	56	LEU	2.7
3	E	242	ARG	2.7
3	E	42	SER	2.7
3	E	25	GLY	2.6
3	E	34	ARG	2.6
3	E	155	LYS	2.6
1	A	253	PHE	2.6
3	E	165	LEU	2.5
3	E	76	VAL	2.5
2	C	317	DG	2.5
3	E	19	ARG	2.5
3	E	287	PRO	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	280	TRP	2.4
3	E	177	TYR	2.4
3	E	271	LEU	2.3
3	E	277	ARG	2.3
3	E	45	TYR	2.2
3	E	170	PHE	2.2
3	E	132	GLY	2.2
3	E	290	PRO	2.2
3	E	20	ALA	2.2
3	E	91	GLU	2.1
3	E	157	PHE	2.1
1	B	241	GLY	2.1
1	B	200	ALA	2.0
3	E	104	ALA	2.0
3	E	53	GLY	2.0
2	F	291	DC	2.0
3	E	174	ILE	2.0
3	E	138	LEU	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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å^2)	Q<0.9
2	CTG	C	297	22/23	0.92	0.27	-	26,38,49,54	0
2	CTG	D	297	22/23	0.92	0.24	-	18,26,32,40	0
2	CTG	F	297	22/23	0.90	0.18	-	47,60,70,71	0

6.3 Carbohydrates (i)

There are no carbohydrates 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.