Full wwPDB X-ray Structure Validation Report ⓘ

Apr 21, 2026 – 06:53 AM JST

PDB ID	:	9L89 / pdb_00009l89
Title	:	Structure of human TREX1-DNA complex with K160R disease-associated mutation
Authors	:	Zhu, J.; Wang, L.; Lin, C.; Zhou, W.
Deposited on	:	2024-12-27
Resolution	:	2.38 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:

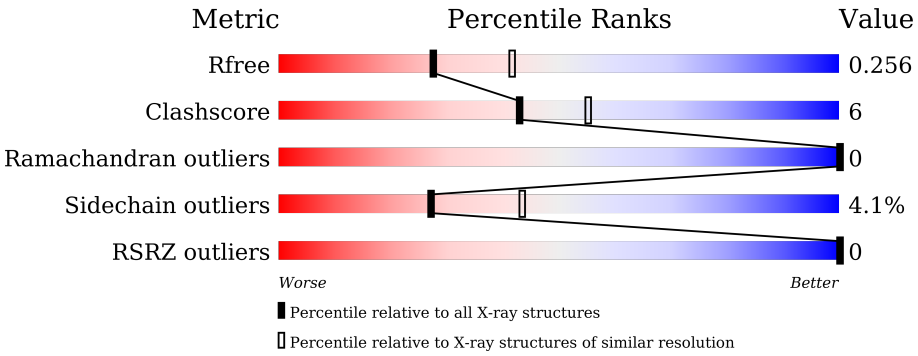
MolProbity	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

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 2.38 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}	180053	7164 (2.40-2.36)
Clashscore	190562	7722 (2.40-2.36)
Ramachandran outliers	187476	7626 (2.40-2.36)
Sidechain outliers	187428	7627 (2.40-2.36)
RSRZ outliers	180081	7170 (2.40-2.36)

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	242	77% 10% • 12%
1	B	242	80% 8% 12%
1	E	242	76% 12% • 12%
1	F	242	76% 12% 12%
2	C	22	36% 36% 27%
2	D	22	27% 45% 27%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	G	22	36% 36% 27%
2	H	22	50% 23% 27%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 8051 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 Three-prime repair exonuclease 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	212	Total	C	N	O	S	0	0	0
			1610	1028	279	291	12			
1	B	213	Total	C	N	O	S	0	1	0
			1623	1038	281	291	13			
1	E	213	Total	C	N	O	S	0	0	0
			1626	1038	284	292	12			
1	F	214	Total	C	N	O	S	0	0	0
			1631	1041	285	293	12			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	160	ARG	LYS	engineered mutation	UNP Q9NSU2
B	160	ARG	LYS	engineered mutation	UNP Q9NSU2
E	160	ARG	LYS	engineered mutation	UNP Q9NSU2
F	160	ARG	LYS	engineered mutation	UNP Q9NSU2

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	16	Total	C	N	O	P	0	0	0
			327	154	59	98	16			
2	D	16	Total	C	N	O	P	0	0	0
			325	154	56	99	16			
2	G	16	Total	C	N	O	P	0	0	0
			327	154	59	98	16			
2	H	16	Total	C	N	O	P	0	0	0
			327	154	59	98	16			

- Molecule 3 is water.

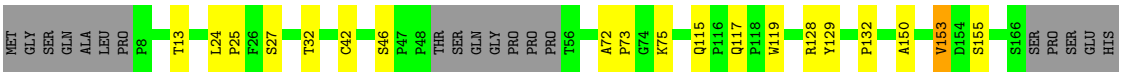
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	28	Total O 28 28	0	0
3	B	64	Total O 64 64	0	0
3	C	9	Total O 9 9	0	0
3	D	11	Total O 11 11	0	0
3	E	65	Total O 65 65	0	0
3	F	58	Total O 58 58	0	0
3	G	8	Total O 8 8	0	0
3	H	12	Total O 12 12	0	0

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.

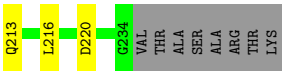
• Molecule 1: Three-prime repair exonuclease 1

Chain A:



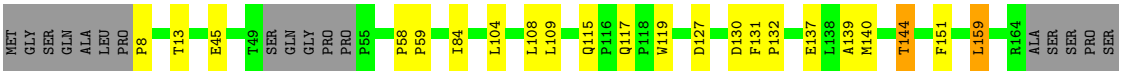
• Molecule 1: Three-prime repair exonuclease 1

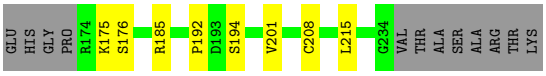
Chain B:



• Molecule 1: Three-prime repair exonuclease 1

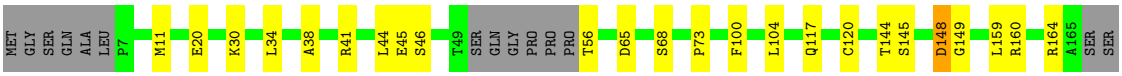
Chain E:

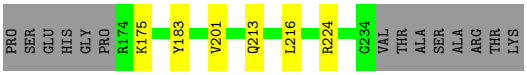




• Molecule 1: Three-prime repair exonuclease 1

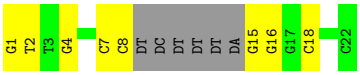
Chain F:





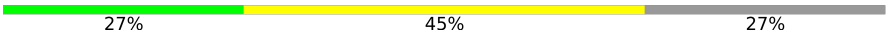
• Molecule 2: DNA (5'-D(P*GP*TP*TP*GP*GP*CP*CP*CP*TP*CP*TP*TP*TP*AP*GP*G P*GP*CP*CP*AP*TP*C)-3')

Chain C:



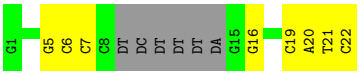
• Molecule 2: DNA (5'-D(P*GP*TP*TP*GP*GP*CP*CP*CP*TP*CP*TP*TP*TP*AP*GP*G P*GP*CP*CP*AP*TP*C)-3')

Chain D:



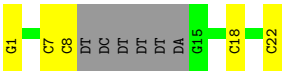
• Molecule 2: DNA (5'-D(P*GP*TP*TP*GP*GP*CP*CP*CP*TP*CP*TP*TP*TP*AP*GP*G P*GP*CP*CP*AP*TP*C)-3')

Chain G:



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

Chain H:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	43.78Å 91.29Å 158.24Å 90.00° 90.86° 90.00°	Depositor
Resolution (Å)	43.77 – 2.38 43.77 – 2.38	Depositor EDS
% Data completeness (in resolution range)	99.0 (43.77-2.38) 94.8 (43.77-2.38)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.79 (at 2.37Å)	Xtriage
Refinement program	PHENIX 1.21.1_5127	Depositor
R, R_{free}	0.220 , 0.255 0.220 , 0.256	Depositor DCC
R_{free} test set	2505 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å^2)	29.4	Xtriage
Anisotropy	0.686	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 28.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.43$, $< L^2 > = 0.25$	Xtriage
Estimated twinning fraction	0.167 for h,-k,-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	8051	wwPDB-VP
Average B, all atoms (Å^2)	43.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.27% 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

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	0.13	0/1651	0.29	0/2254
1	B	0.15	0/1669	0.31	0/2278
1	E	0.13	0/1668	0.32	0/2275
1	F	0.18	0/1673	0.32	0/2283
2	C	0.17	0/364	0.35	0/557
2	D	0.22	0/361	0.47	0/552
2	G	0.26	0/364	0.49	0/557
2	H	0.25	0/364	0.40	0/557
All	All	0.16	0/8114	0.34	0/11313

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1610	0	1611	13	0
1	B	1623	0	1630	13	0
1	E	1626	0	1633	15	0
1	F	1631	0	1637	15	0
2	C	327	0	181	8	0
2	D	325	0	182	13	0
2	G	327	0	181	10	0
2	H	327	0	181	6	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	28	0	0	2	0
3	B	64	0	0	6	0
3	C	9	0	0	0	0
3	D	11	0	0	4	0
3	E	65	0	0	3	2
3	F	58	0	0	7	0
3	G	8	0	0	2	0
3	H	12	0	0	2	2
All	All	8051	0	7236	89	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 6.

All (89) 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:33:GLU:OE2	3:B:301:HOH:O	1.83	0.95
2:D:1:DG:N3	3:D:101:HOH:O	2.14	0.81
2:C:7:DC:H2"	2:C:8:DC:H5'	1.64	0.80
1:F:45:GLU:OE1	3:F:301:HOH:O	2.01	0.78
2:G:6:DC:OP1	3:G:101:HOH:O	2.01	0.78
2:H:1:DG:N2	3:H:102:HOH:O	2.21	0.74
1:B:148:ASP:O	3:B:302:HOH:O	2.04	0.73
1:F:73:PRO:O	3:F:302:HOH:O	2.07	0.72
1:A:32:THR:HG22	1:A:73:PRO:HG3	1.73	0.70
2:G:6:DC:H1'	2:G:7:DC:H4'	1.73	0.70
1:E:130:ASP:OD1	3:E:302:HOH:O	2.09	0.70
1:E:192:PRO:O	3:E:301:HOH:O	2.09	0.69
1:B:164:ARG:NH2	2:D:8:DC:OP2	2.26	0.69
1:B:190:SER:O	3:B:303:HOH:O	2.11	0.68
1:E:176:SER:O	1:E:185:ARG:NH2	2.25	0.68
1:E:127:ASP:OD1	3:E:303:HOH:O	2.11	0.67
2:C:2:DT:H5'	2:C:2:DT:H6	1.60	0.65
2:C:1:DG:H2'	2:C:2:DT:H5"	1.80	0.64
2:G:6:DC:C2	2:G:7:DC:H1'	2.34	0.63
2:C:1:DG:H8	2:C:1:DG:H5"	1.65	0.61
2:G:22:DC:H5"	2:G:22:DC:H6	1.66	0.61
2:H:18:DC:OP1	3:H:101:HOH:O	2.16	0.61
1:B:41:ARG:NH2	1:B:220:ASP:OD2	2.32	0.61
2:D:22:DC:H6	2:D:22:DC:H5"	1.67	0.59
1:F:100:PHE:HA	1:F:104:LEU:HD23	1.84	0.58

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:8:DC:H2"	2:D:9:DT:H5"	1.86	0.57
1:A:231:PRO:HB2	1:A:235:VAL:HG22	1.87	0.57
1:F:11:MET:O	1:F:41:ARG:HD2	2.06	0.56
2:D:22:DC:H5"	2:D:22:DC:C6	2.43	0.54
1:B:106:ASN:ND2	3:B:311:HOH:O	2.42	0.53
1:A:119:TRP:HE3	1:A:150:ALA:HB2	1.74	0.53
1:F:183:TYR:OH	3:F:303:HOH:O	2.16	0.52
1:F:213:GLN:OE1	3:F:304:HOH:O	2.19	0.52
1:F:149:GLY:O	3:F:305:HOH:O	2.19	0.52
1:A:153:VAL:HG23	1:A:223:ALA:HA	1.92	0.52
2:C:15:DG:H2"	2:C:16:DG:C8	2.45	0.52
2:G:6:DC:C1'	2:G:7:DC:H4'	2.38	0.51
2:D:3:DT:O4	2:D:19:DC:N4	2.44	0.50
1:B:63:VAL:O	3:B:304:HOH:O	2.20	0.50
1:E:8:PRO:HG3	1:E:151:PHE:CZ	2.47	0.49
2:G:5:DG:H2"	2:G:6:DC:H2'	1.95	0.49
1:F:160:ARG:O	1:F:164:ARG:HB2	2.12	0.49
1:E:208:CYS:HB3	1:E:215:LEU:HD22	1.95	0.49
1:A:27:SER:O	1:A:75:LYS:NZ	2.45	0.48
2:D:19:DC:H2"	2:D:20:DA:C8	2.48	0.48
1:B:213:GLN:NE2	3:B:315:HOH:O	2.46	0.48
1:A:13:THR:HG21	1:A:115:GLN:HB3	1.94	0.48
1:A:25:PRO:O	3:A:301:HOH:O	2.20	0.48
1:F:38:ALA:HB3	1:F:65:ASP:HB2	1.95	0.48
1:F:145:SER:HA	1:F:148:ASP:OD2	2.14	0.48
2:H:7:DC:H2"	2:H:8:DC:O5'	2.14	0.47
2:G:21:DT:OP2	3:G:102:HOH:O	2.20	0.47
1:E:117:GLN:HG2	1:E:119:TRP:CE2	2.50	0.46
2:G:19:DC:H2"	2:G:20:DA:C8	2.51	0.46
2:D:7:DC:H2"	2:D:8:DC:O4'	2.16	0.46
1:A:24:LEU:HD13	2:C:2:DT:C2	2.51	0.46
2:D:21:DT:OP2	3:D:102:HOH:O	2.21	0.46
1:B:117:GLN:HG3	1:B:119:TRP:CE2	2.52	0.45
1:E:139:ALA:HA	1:E:144:THR:HG23	1.99	0.45
1:F:30:LYS:NZ	3:F:313:HOH:O	2.47	0.45
2:C:2:DT:H5'	2:C:2:DT:C6	2.48	0.45
1:E:13:THR:HG21	1:E:115:GLN:HB3	1.99	0.45
1:A:117:GLN:HG2	1:A:119:TRP:CE2	2.53	0.44
2:C:4:DG:O6	2:C:18:DC:N4	2.51	0.44
1:E:58:PRO:HA	1:E:59:PRO:HD3	1.88	0.44
1:B:160:ARG:NE	2:D:7:DC:H5"	2.33	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:8:DC:OP1	2:D:8:DC:H4'	2.17	0.44
2:G:6:DC:C2'	2:G:7:DC:H4'	2.48	0.43
2:H:1:DG:H2'	2:H:1:DG:N3	2.33	0.43
1:F:144:THR:HG23	3:F:332:HOH:O	2.17	0.43
1:F:20:GLU:OE1	2:H:22:DC:H5'	2.19	0.43
2:D:20:DA:OP2	3:D:104:HOH:O	2.21	0.43
1:E:137:GLU:HA	1:E:140:MET:HE3	2.01	0.42
1:A:132:PRO:HB2	1:A:232:MET:HB2	2.01	0.42
1:B:36:LEU:HD12	1:B:67:LEU:HD23	2.02	0.42
2:G:7:DC:H42	2:G:16:DG:H21	1.68	0.42
1:F:216:LEU:HD23	1:F:216:LEU:HA	1.91	0.42
1:A:153:VAL:HG11	1:A:219:VAL:HG22	2.02	0.42
1:E:159:LEU:HD13	1:E:159:LEU:HA	1.90	0.41
1:E:109:LEU:HD13	1:E:109:LEU:HA	1.95	0.41
1:A:72:ALA:HA	1:A:73:PRO:HD3	1.91	0.41
2:D:4:DG:OP2	3:D:103:HOH:O	2.21	0.41
1:F:11:MET:SD	1:F:120:CYS:HB2	2.60	0.41
1:B:191:PRO:HA	1:B:192:PRO:HD3	1.91	0.41
1:B:216:LEU:HD23	1:B:216:LEU:HA	1.95	0.41
1:E:104:LEU:O	1:E:108:LEU:HG	2.21	0.41
1:E:131:PHE:HB2	1:E:132:PRO:HD3	2.03	0.40
1:A:129:TYR:HB3	3:A:305:HOH:O	2.21	0.40
2:H:7:DC:H2'	2:H:7:DC:OP1	2.21	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)
3:E:336:HOH:O	3:H:108:HOH:O[2_446]	2.05	0.15
3:E:351:HOH:O	3:H:112:HOH:O[2_446]	2.14	0.06

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	206/242 (85%)	201 (98%)	5 (2%)	0	100	100
1	B	208/242 (86%)	206 (99%)	2 (1%)	0	100	100
1	E	207/242 (86%)	202 (98%)	5 (2%)	0	100	100
1	F	208/242 (86%)	204 (98%)	4 (2%)	0	100	100
All	All	829/968 (86%)	813 (98%)	16 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	174/198 (88%)	166 (95%)	8 (5%)	24	38
1	B	176/198 (89%)	173 (98%)	3 (2%)	53	72
1	E	176/198 (89%)	169 (96%)	7 (4%)	28	44
1	F	176/198 (89%)	165 (94%)	11 (6%)	16	26
All	All	702/792 (89%)	673 (96%)	29 (4%)	27	43

All (29) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	42	CYS
1	A	46	SER
1	A	128	ARG
1	A	153	VAL
1	A	155	SER
1	A	190	SER
1	A	201	VAL
1	A	219	VAL
1	B	71	VAL
1	B	83	GLU
1	B	109	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	E	45	GLU
1	E	84	ILE
1	E	144	THR
1	E	159	LEU
1	E	175	LYS
1	E	194	SER
1	E	201	VAL
1	F	34	LEU
1	F	44	LEU
1	F	46	SER
1	F	56	THR
1	F	68	SER
1	F	117	GLN
1	F	148	ASP
1	F	159	LEU
1	F	175	LYS
1	F	201	VAL
1	F	224	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	117	GLN
1	B	189	GLN
1	E	209	GLN
1	F	12	GLN
1	F	209	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)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

There are no oligosaccharides 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	212/242 (87%)	-1.21	0	100	100	22, 41, 67, 88	0
1	B	213/242 (88%)	-1.35	0	100	100	19, 30, 48, 76	1 (0%)
1	E	213/242 (88%)	-1.35	0	100	100	22, 33, 52, 76	0
1	F	214/242 (88%)	-1.35	0	100	100	22, 31, 55, 74	0
2	C	16/22 (72%)	-0.92	0	100	100	36, 83, 123, 123	0
2	D	16/22 (72%)	-1.23	0	100	100	23, 55, 111, 112	0
2	G	16/22 (72%)	-0.61	0	100	100	36, 90, 141, 155	0
2	H	16/22 (72%)	-1.08	0	100	100	30, 56, 111, 114	0
All	All	916/1056 (86%)	-1.29	0	100	100	19, 34, 75, 155	1 (0%)

There are no RSRZ outliers to report.

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no oligosaccharides 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.