Full wwPDB X-ray Structure Validation Report ⓘ

Jan 12, 2022 – 09:26 am GMT

PDB ID	:	7O0G
Title	:	Structure of the foamy viral protease-reverse transcriptase in complex with RNA/DNA hybrid.
Authors	:	Nowak, E.; Nowacka, M.; Nowotny, M.
Deposited on	:	2021-03-26
Resolution	:	3.10 Å(reported)

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

|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.24
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.24

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.10 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.
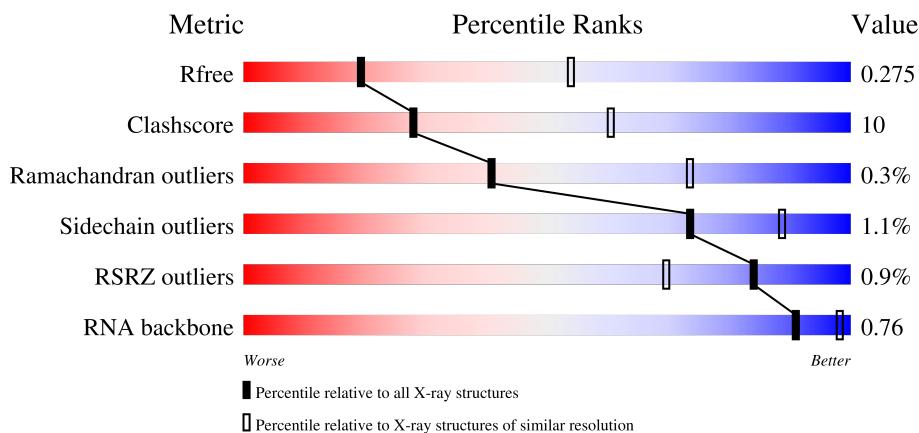
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1094 (3.10-3.10)
Clashscore	141614	1184 (3.10-3.10)
Ramachandran outliers	138981	1141 (3.10-3.10)
Sidechain outliers	138945	1141 (3.10-3.10)
RSRZ outliers	127900	1067 (3.10-3.10)
RNA backbone	3102	1116 (3.40-2.80)

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	752	74% 23% ·
2	E	18	72% 22% 6%
3	F	15	53% 47%
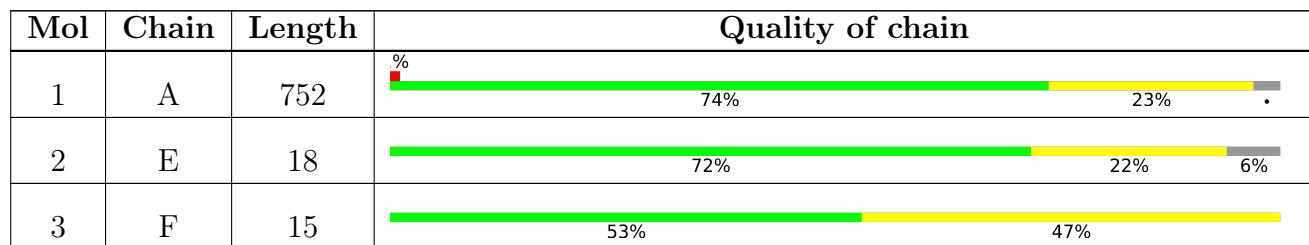
2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 6270 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 Pr125Pol.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	733	Total	C	N	O	S	0	0	0
			5569	3607	921	1033	8			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	586	LEU	UNK	variant	UNP D5JWV1
A	598	ASN	ASP	engineered mutation	UNP D5JWV1
A	668	ASN	ASP	engineered mutation	UNP D5JWV1

- Molecule 2 is a RNA chain called RNA (5'-R(*UP*UP*CP*UP*UP*GP*UP*CP*CP*AP* GP*GP*AP*GP*AP*GP*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	17	Total	C	N	O	P	0	0	0
			361	162	64	119	16			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	F	15	Total	C	N	O	P	0	2	0
			339	163	59	101	16			

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	O	0	0
			1	1		

3 Residue-property plots ⓘ

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: Pr125Pol
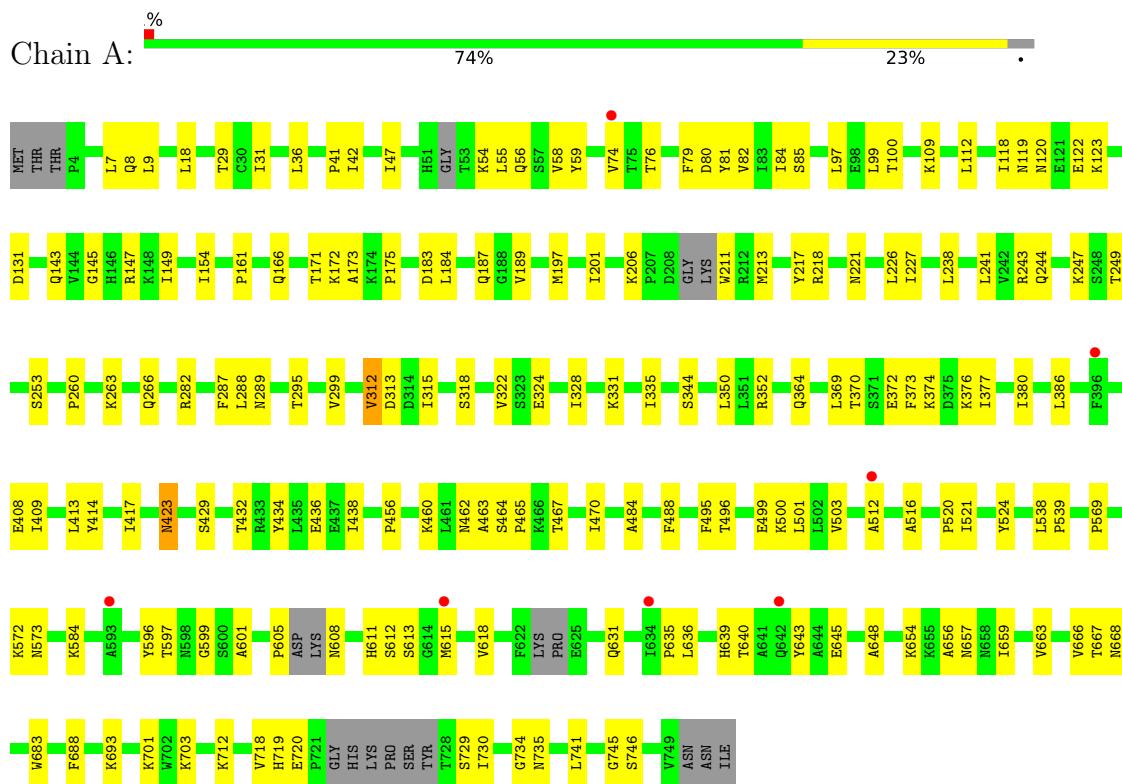
Chain A:

- Molecule 2: RNA (5'-R(*UP*UP*CP*UP*UP*GP*UP*CP*CP*AP*GP*GP*AP*GP*AP*GP *G)-3')
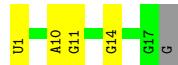
Chain E:

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

Chain F:

C1　C7 T8 G9　C12 A13 A14 G15
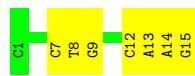
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α, β, γ	105.03Å 105.03Å 253.53Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.52 – 3.10 48.52 – 3.10	Depositor EDS
% Data completeness (in resolution range)	99.7 (48.52-3.10) 93.4 (48.52-3.10)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.21 (at 3.12Å)	Xtriage
Refinement program	PHENIX 1.19_4092, PHENIX 1.19_4092	Depositor
R, R_{free}	0.247 , 0.273 0.245 , 0.275	Depositor DCC
R_{free} test set	1334 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	111.8	Xtriage
Anisotropy	0.396	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< |L| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	6270	wwPDB-VP
Average B, all atoms (Å^2)	126.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.15% 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 (i)

5.1 Standard geometry (i)

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	0.25	0/5710	0.47	0/7816
2	E	0.13	0/403	0.72	0/627
3	F	0.50	0/378	0.91	0/580
All	All	0.27	0/6491	0.53	0/9023

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 (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	5569	0	5294	112	0
2	E	361	0	184	2	0
3	F	339	0	193	5	0
4	A	1	0	0	0	0
All	All	6270	0	5671	118	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 10.

All (118) 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:29:THR:HG22	1:A:85:SER:H	1.21	1.02
1:A:42:ILE:HG23	1:A:58:VAL:HG23	1.65	0.76
1:A:374:LYS:HA	1:A:377:ILE:HD12	1.69	0.75
1:A:206:LYS:H	1:A:211:TRP:HA	1.53	0.74
1:A:312:VAL:HG23	1:A:313:ASP:H	1.56	0.71
1:A:54:LYS:HD2	1:A:55:LEU:H	1.56	0.69
1:A:244:GLN:HB2	1:A:318:SER:HB2	1.76	0.68
1:A:253:SER:HB3	1:A:344:SER:HB2	1.77	0.66
1:A:605:PRO:O	1:A:608:ASN:N	2.29	0.66
1:A:470:ILE:HB	1:A:484:ALA:HB3	1.78	0.65
1:A:656:ALA:HB1	1:A:663:VAL:HG11	1.79	0.64
1:A:147:ARG:HG3	1:A:149:ILE:HG12	1.83	0.60
1:A:423:ASN:O	1:A:423:ASN:ND2	2.30	0.59
1:A:597:THR:HB	1:A:648:ALA:HB1	1.85	0.57
1:A:413:LEU:HD12	1:A:413:LEU:H	1.70	0.57
1:A:640:THR:HG23	1:A:643:TYR:H	1.71	0.56
1:A:166:GLN:OE1	1:A:201:ILE:HG22	2.05	0.55
1:A:370:THR:HG22	1:A:372:GLU:H	1.71	0.55
1:A:184:LEU:HD12	1:A:189:VAL:HG11	1.90	0.54
1:A:145:GLY:O	1:A:282:ARG:NH1	2.34	0.54
3:F:12:DC:H2'	3:F:13:DA:C8	2.42	0.54
1:A:171:THR:O	1:A:175:PRO:HD3	2.09	0.53
3:F:14:DA:H2"	3:F:15:DG:H5"	1.90	0.53
1:A:369:LEU:HD22	1:A:373:PHE:CE2	2.43	0.53
1:A:613:SER:OG	1:A:639:HIS:O	2.26	0.53
1:A:119:ASN:OD1	1:A:122:GLU:HG3	2.09	0.53
3:F:8:DT:H2'	3:F:9:DG:C8	2.44	0.53
1:A:432:THR:O	1:A:436:GLU:HG3	2.09	0.52
1:A:247:LYS:HB3	1:A:318:SER:HB3	1.90	0.52
1:A:118:ILE:HG13	1:A:123:LYS:HG2	1.90	0.52
1:A:373:PHE:CZ	1:A:377:ILE:HD11	2.44	0.52
1:A:243:ARG:NH1	1:A:364:GLN:O	2.43	0.52
1:A:54:LYS:HD2	1:A:55:LEU:N	2.23	0.51
1:A:59:TYR:HB2	1:A:74:VAL:HG23	1.91	0.51
1:A:109:LYS:NZ	1:A:131:ASP:HA	2.27	0.50
1:A:226:LEU:HD12	1:A:288:LEU:HD21	1.94	0.50
1:A:414:TYR:O	1:A:417:ILE:HG22	2.12	0.50
3:F:8:DT:H2'	3:F:9:DG:H8	1.77	0.49
1:A:122:GLU:OE1	1:A:322:VAL:HG11	2.12	0.49
1:A:331:LYS:O	1:A:335:ILE:HG12	2.12	0.49
1:A:42:ILE:CG2	1:A:58:VAL:HG23	2.37	0.49
1:A:109:LYS:HZ2	1:A:131:ASP:HA	1.78	0.49

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:29:THR:HG22	1:A:85:SER:N	2.06	0.48
1:A:462:ASN:ND2	1:A:569:PRO:O	2.38	0.48
1:A:599:GLY:N	1:A:645:GLU:OE2	2.45	0.48
1:A:173:ALA:HB2	1:A:211:TRP:CD2	2.48	0.48
2:E:10:A:H2'	2:E:11:G:C8	2.49	0.48
1:A:495:PHE:HD2	1:A:499:GLU:HB3	1.78	0.48
1:A:295:THR:O	1:A:299:VAL:HG13	2.13	0.48
1:A:729:SER:OG	1:A:730:ILE:N	2.46	0.47
1:A:41:PRO:HG3	1:A:59:TYR:CZ	2.49	0.47
1:A:429:SER:HB3	1:A:584:LYS:HD3	1.96	0.47
1:A:373:PHE:CE2	1:A:377:ILE:HD11	2.50	0.47
1:A:495:PHE:CD2	1:A:499:GLU:HB3	2.50	0.46
1:A:118:ILE:HD12	1:A:122:GLU:HB2	1.96	0.46
1:A:143:GLN:HA	1:A:187:GLN:HE22	1.80	0.46
1:A:47:ILE:HD12	1:A:56:GLN:H	1.79	0.46
1:A:512:ALA:O	1:A:516:ALA:N	2.49	0.46
1:A:161:PRO:HG2	1:A:197:MET:SD	2.56	0.46
1:A:238:LEU:HA	1:A:241:LEU:HD23	1.98	0.46
1:A:172:LYS:HD2	1:A:172:LYS:HA	1.79	0.46
1:A:376:LYS:O	1:A:380:ILE:HG13	2.16	0.45
1:A:635:PRO:HB3	1:A:745:GLY:O	2.16	0.45
1:A:463:ALA:HB3	1:A:501:LEU:HD22	1.99	0.45
1:A:201:ILE:HD11	1:A:213:MET:HG3	1.98	0.45
1:A:7:LEU:HD12	1:A:8:GLN:H	1.82	0.45
1:A:36:LEU:HD13	1:A:59:TYR:CE1	2.52	0.45
3:F:7:DC:H2'	3:F:8:DT:C6	2.52	0.45
1:A:147:ARG:HA	1:A:147:ARG:HD2	1.78	0.45
1:A:408:GLU:HG3	1:A:467:THR:HG21	1.98	0.45
1:A:499:GLU:O	1:A:503:VAL:HG23	2.17	0.44
1:A:631:GLN:HB2	1:A:741:LEU:HD21	1.99	0.44
1:A:701:LYS:HD3	1:A:701:LYS:HA	1.66	0.44
1:A:666:VAL:HA	1:A:718:VAL:O	2.17	0.44
1:A:36:LEU:HD21	1:A:76:THR:HA	1.99	0.44
1:A:79:PHE:HB3	1:A:81:TYR:CE1	2.52	0.44
1:A:112:LEU:HD13	1:A:352:ARG:HH12	1.82	0.44
1:A:370:THR:HG22	1:A:372:GLU:N	2.32	0.44
1:A:688:PHE:CE2	1:A:703:LYS:HG2	2.52	0.44
1:A:456:PRO:HB3	1:A:520:PRO:HG2	2.00	0.44
1:A:667:THR:O	1:A:719:HIS:ND1	2.43	0.43
1:A:31:ILE:HG22	1:A:82:VAL:HG13	2.00	0.43
1:A:227:ILE:N	1:A:289:ASN:OD1	2.41	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:612:SER:O	1:A:746:SER:HB2	2.18	0.43
1:A:683:TRP:HB3	1:A:688:PHE:HD1	1.84	0.43
1:A:460:LYS:HG2	1:A:524:TYR:HB2	2.00	0.43
1:A:324:GLU:O	1:A:328:ILE:HG12	2.19	0.42
1:A:512:ALA:HB1	1:A:521:ILE:HD13	2.01	0.42
1:A:596:TYR:HB2	1:A:618:VAL:HG22	2.01	0.42
1:A:601:ALA:HB1	1:A:611:HIS:ND1	2.34	0.42
1:A:9:LEU:HD21	1:A:84:ILE:HD12	2.01	0.42
1:A:221:ASN:HB3	1:A:288:LEU:HD12	2.02	0.42
1:A:464:SER:HB2	1:A:465:PRO:HD2	2.00	0.42
1:A:538:LEU:N	1:A:539:PRO:HD2	2.34	0.42
1:A:18:LEU:HD23	1:A:18:LEU:HA	1.82	0.42
1:A:218:ARG:NH2	2:E:1:U:O3'	2.52	0.42
1:A:693:LYS:HD3	1:A:693:LYS:HA	1.62	0.42
1:A:154:ILE:HB	1:A:287:PHE:CZ	2.55	0.42
1:A:596:TYR:CD2	1:A:734:GLY:HA3	2.55	0.42
1:A:635:PRO:O	1:A:636:LEU:HD23	2.20	0.42
1:A:409:ILE:O	1:A:438:ILE:HD11	2.19	0.41
1:A:668:ASN:HB3	1:A:735:ASN:OD1	2.20	0.41
1:A:260:PRO:HA	1:A:282:ARG:HG3	2.01	0.41
1:A:217:TYR:O	1:A:221:ASN:ND2	2.53	0.41
1:A:496:THR:HG23	1:A:499:GLU:H	1.85	0.41
1:A:100:THR:OG1	1:A:183:ASP:OD2	2.37	0.41
1:A:112:LEU:HD12	1:A:350:LEU:HG	2.01	0.41
1:A:249:THR:HA	1:A:315:ILE:O	2.21	0.41
1:A:386:LEU:HD23	1:A:386:LEU:HA	1.87	0.41
1:A:668:ASN:HA	1:A:719:HIS:CE1	2.55	0.41
1:A:659:ILE:O	1:A:712:LYS:HE2	2.21	0.41
1:A:97:LEU:HD13	1:A:99:LEU:HD13	2.02	0.40
1:A:434:TYR:O	1:A:438:ILE:HG12	2.21	0.40
1:A:99:LEU:HB3	1:A:100:THR:H	1.76	0.40
1:A:386:LEU:H	1:A:423:ASN:HD22	1.68	0.40
1:A:263:LYS:HA	1:A:266:GLN:HG3	2.04	0.40
1:A:488:PHE:CZ	1:A:500:LYS:HB3	2.57	0.40
1:A:572:LYS:HG3	1:A:573:ASN:OD1	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	721/752 (96%)	703 (98%)	16 (2%)	2 (0%)	41	73

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	312	VAL
1	A	720	GLU

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	566/677 (84%)	560 (99%)	6 (1%)	73	89

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	80	ASP
1	A	120	ASN
1	A	423	ASN
1	A	615	MET
1	A	654	LYS
1	A	657	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	423	ASN

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	E	16/18 (88%)	1 (6%)	0

All (1) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	E	14	G

There are no RNA pucker outliers to report.

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

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

5.5 Carbohydrates ⓘ

There are no monosaccharides 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, 95th 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(Å²)	Q<0.9
1	A	733/752 (97%)	-0.12	7 (0%)	82	67	76, 126, 169, 224	0
2	E	17/18 (94%)	-0.18	0	100	100	104, 122, 173, 183	0
3	F	15/15 (100%)	-0.32	0	100	100	99, 116, 188, 198	0
All	All	765/785 (97%)	-0.12	7 (0%)	84	69	76, 125, 169, 224	0

All (7) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	74	VAL	2.8
1	A	512	ALA	2.7
1	A	593	ALA	2.6
1	A	396	PHE	2.5
1	A	615	MET	2.2
1	A	642	GLN	2.1
1	A	634	ILE	2.0

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 monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.