WORLDWIDE PDB

PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report (i)

Oct 22, 2024 – 09:59 AM EDT

PDB ID	:	3UB0
Title	:	Crystal structure of the nonstructural protein 7 and 8 complex of Feline Coronavirus
Authors	:	Xiao, Y.; Hilgenfeld, R.; Ma, Q.
Deposited on	:	2011-10-22
Resolution	:	2.60 Å(reported)

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

|		
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

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 2.60 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	Percentile Ranks	Value
Rfree		0.227
Clashscore		5
Ramachandran outliers		0.9%
Sidechain outliers		9.8%
RSRZ outliers		4.7%

Worse — *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution
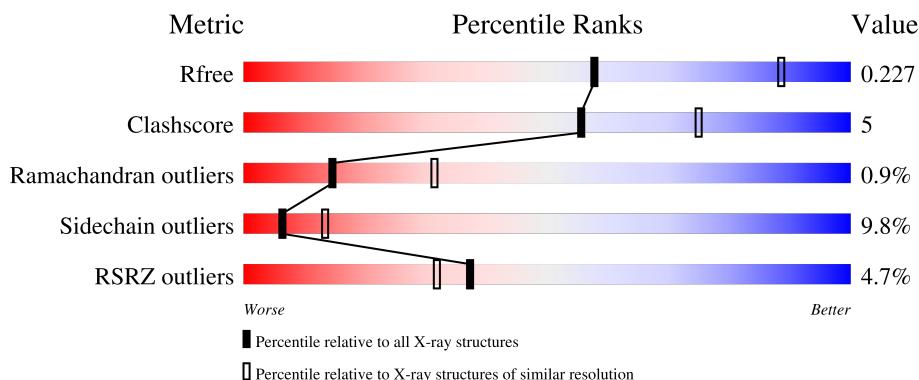
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	3775 (2.60-2.60)
Clashscore	180529	4181 (2.60-2.60)
Ramachandran outliers	177936	4129 (2.60-2.60)
Sidechain outliers	177891	4129 (2.60-2.60)
RSRZ outliers	164620	3775 (2.60-2.60)

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	199	4% / 78% ... 16% · 5%
1	D	199	10% / 78% ... 18% · ·
2	B	87	% / 82% ... 9% · 6%
2	C	87	% / 77% ... 15% · 6%
2	E	87	% / 80% ... 11% · 7%
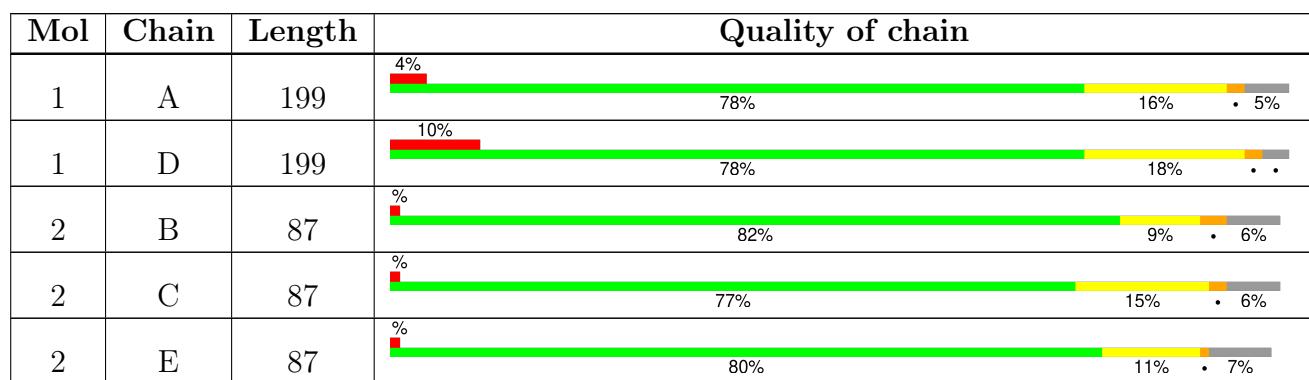
Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	87	

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 5575 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 Non-structural protein 6, nsp6,.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	190	Total	C	N	O	S	Se	0	0	0
			1459	921	254	277	2	5			
1	D	194	Total	C	N	O	S	Se	0	0	0
			1488	938	261	282	2	5			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-3	GLY	-	expression tag	UNP Q98VG9
A	-2	PRO	-	expression tag	UNP Q98VG9
A	-1	LEU	-	expression tag	UNP Q98VG9
A	0	GLY	-	expression tag	UNP Q98VG9
D	-3	GLY	-	expression tag	UNP Q98VG9
D	-2	PRO	-	expression tag	UNP Q98VG9
D	-1	LEU	-	expression tag	UNP Q98VG9
D	0	GLY	-	expression tag	UNP Q98VG9

- Molecule 2 is a protein called Non-structural protein 7, nsp7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	82	Total	C	N	O	S	0	0	0
			651	415	103	127	6			
2	C	82	Total	C	N	O	S	0	0	0
			644	409	102	127	6			
2	E	81	Total	C	N	O	S	0	0	0
			642	410	101	125	6			
2	F	78	Total	C	N	O	S	0	0	0
			614	391	97	120	6			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-3	GLY	-	expression tag	UNP Q98VG9
B	-2	PRO	-	expression tag	UNP Q98VG9
B	-1	LEU	-	expression tag	UNP Q98VG9
B	0	GLY	-	expression tag	UNP Q98VG9
C	-3	GLY	-	expression tag	UNP Q98VG9
C	-2	PRO	-	expression tag	UNP Q98VG9
C	-1	LEU	-	expression tag	UNP Q98VG9
C	0	GLY	-	expression tag	UNP Q98VG9
E	-3	GLY	-	expression tag	UNP Q98VG9
E	-2	PRO	-	expression tag	UNP Q98VG9
E	-1	LEU	-	expression tag	UNP Q98VG9
E	0	GLY	-	expression tag	UNP Q98VG9
F	-3	GLY	-	expression tag	UNP Q98VG9
F	-2	PRO	-	expression tag	UNP Q98VG9
F	-1	LEU	-	expression tag	UNP Q98VG9
F	0	GLY	-	expression tag	UNP Q98VG9

- Molecule 3 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	44	Total 44	O 44	0	0
3	B	6	Total 6	O 6	0	0
3	C	10	Total 10	O 10	0	0
3	D	9	Total 9	O 9	0	0
3	E	4	Total 4	O 4	0	0
3	F	4	Total 4	O 4	0	0

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: Non-structural protein 6, nsp6,
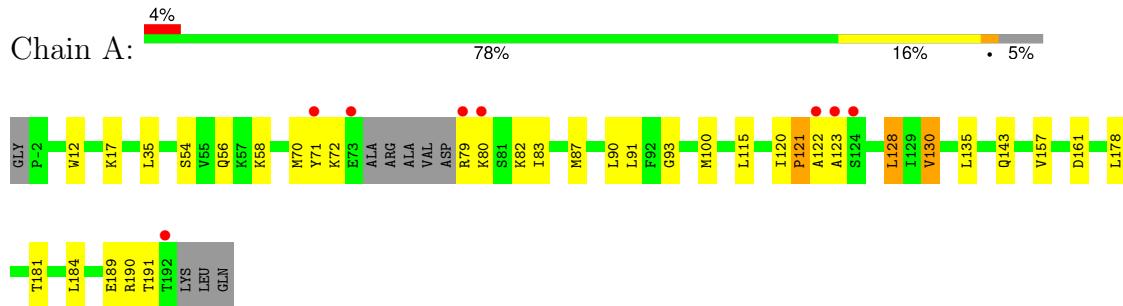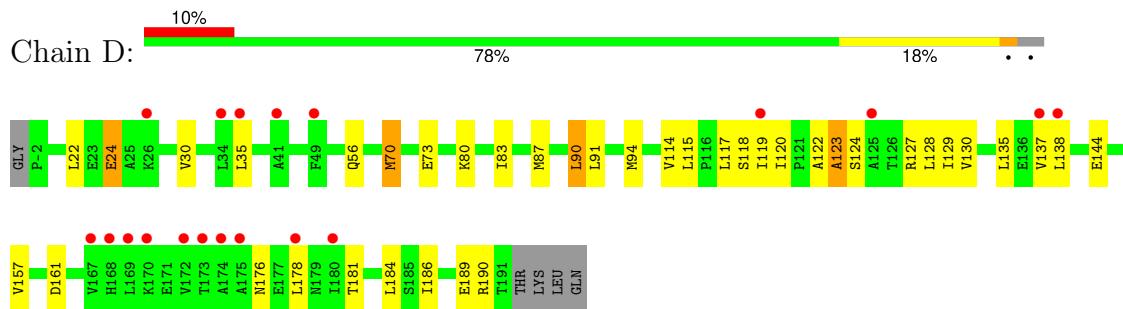
Chain A:

- Molecule 1: Non-structural protein 6, nsp6,

Chain D:

- Molecule 2: Non-structural protein 7, nsp7

Chain B:

- Molecule 2: Non-structural protein 7, nsp7

Chain C:

- Molecule 2: Non-structural protein 7, nsp7

Chain E:

• Molecule 2: Non-structural protein 7, nsp7

Chain F:
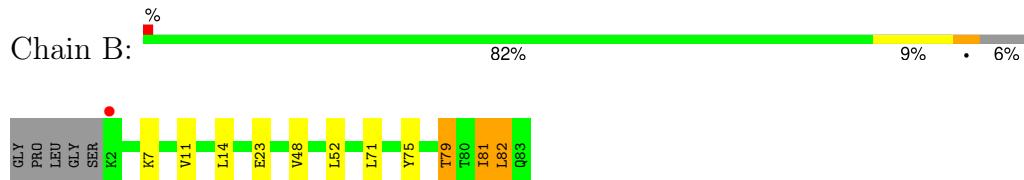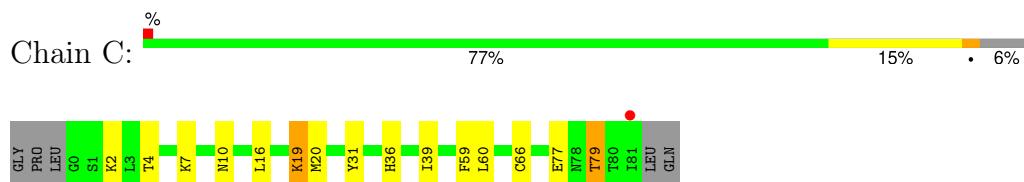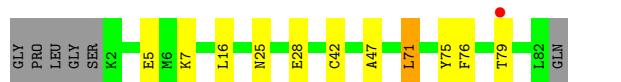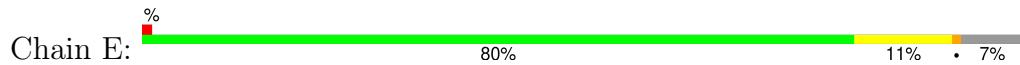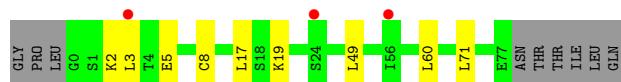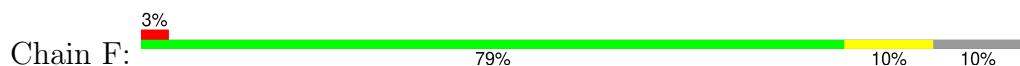
4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	121.45Å 160.30Å 102.66Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	33.45 – 2.60 33.45 – 2.60	Depositor EDS
% Data completeness (in resolution range)	100.0 (33.45-2.60) 99.9 (33.45-2.60)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	0.06	Depositor
$< I/\sigma(I) >$ [1]	2.84 (at 2.61Å)	Xtriage
Refinement program	BUSTER 2.11.1	Depositor
R, R_{free}	0.205 , 0.227 0.208 , 0.227	Depositor DCC
R_{free} test set	1471 reflections (4.72%)	wwPDB-VP
Wilson B-factor (Å^2)	61.6	Xtriage
Anisotropy	0.358	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 76.1	EDS
L-test for twinning[2]	$< |L| > = 0.48, < L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	5575	wwPDB-VP
Average B, all atoms (Å^2)	86.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.56% 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	0.52	0/1474	0.78	0/1984
1	D	0.45	0/1504	0.74	1/2026 (0.0%)
2	B	0.52	0/661	0.76	0/895
2	C	0.53	0/654	0.69	0/885
2	E	0.47	0/652	0.70	0/883
2	F	0.48	0/624	0.68	0/843
All	All	0.49	0/5569	0.74	1/7516 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	123	ALA	C-N-CA	6.92	138.99	121.70

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	1459	0	1517	18	0
1	D	1488	0	1547	25	0
2	B	651	0	650	5	0
2	C	644	0	642	10	0
2	E	642	0	642	6	0
2	F	614	0	611	4	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	44	0	0	0	0
3	B	6	0	0	0	0
3	C	10	0	0	0	0
3	D	9	0	0	0	0
3	E	4	0	0	0	0
3	F	4	0	0	0	0
All	All	5575	0	5609	51	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 5.

All (51) 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:12:TRP:CZ3	1:D:70:MSE:HE3	2.00	0.96
1:A:12:TRP:HZ3	1:D:70:MSE:HE3	1.30	0.94
1:D:87:MSE:HE2	2:E:76:PHE:HB3	1.70	0.74
1:D:83:ILE:HG22	1:D:87:MSE:HE3	1.70	0.73
2:C:10:ASN:HD21	2:C:36:HIS:HA	1.53	0.73
2:B:48:VAL:HG21	2:B:81:ILE:HG21	1.72	0.71
1:A:80:LYS:HB2	1:A:83:ILE:H	1.63	0.64
1:D:87:MSE:HE2	2:E:76:PHE:CB	2.29	0.62
1:D:114:VAL:HG21	1:D:137:VAL:HG22	1.82	0.62
1:A:79:ARG:HE	1:A:80:LYS:HG3	1.62	0.62
1:D:114:VAL:HG21	1:D:137:VAL:CG2	2.29	0.62
2:B:75:TYR:O	2:B:79:THR:HG23	2.04	0.58
1:D:83:ILE:CG2	1:D:87:MSE:HE3	2.35	0.56
1:A:70:MSE:HB2	2:B:11:VAL:HG21	1.88	0.56
1:A:128:LEU:HD13	1:A:130:VAL:HG22	1.91	0.52
1:A:12:TRP:HZ3	1:D:70:MSE:CE	2.13	0.52
1:A:121:PRO:HG3	2:C:31:TYR:HB2	1.92	0.52
1:A:93:GLY:HA2	2:C:79:THR:HG21	1.93	0.51
2:E:75:TYR:O	2:E:79:THR:HG22	2.11	0.51
1:D:87:MSE:HE1	2:F:19:LYS:HE3	1.92	0.50
1:A:12:TRP:CZ3	1:D:70:MSE:CE	2.87	0.50
1:D:94:MSE:HE1	2:F:8:CYS:HB2	1.94	0.49
2:B:75:TYR:O	2:B:79:THR:CG2	2.60	0.49
1:A:121:PRO:HB3	2:C:31:TYR:CD1	2.49	0.48
1:D:157:VAL:HG21	1:D:189:GLU:HG3	1.95	0.48
1:D:120:ILE:HG13	1:D:127:ARG:HB3	1.96	0.47
1:D:127:ARG:HH21	1:D:157:VAL:HG11	1.79	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:138:LEU:HD13	1:D:186:ILE:HD11	1.97	0.47
2:C:20:MET:HE1	2:C:59:PHE:CE1	2.49	0.46
1:D:118:SER:HB2	1:D:129:ILE:HD13	1.96	0.46
1:A:121:PRO:HB2	1:A:122:ALA:H	1.57	0.45
1:A:157:VAL:HG21	1:A:189:GLU:HG3	1.98	0.45
1:A:79:ARG:N	2:C:19:LYS:HZ2	2.15	0.44
1:A:79:ARG:HA	1:D:24:GLU:OE2	2.18	0.44
2:C:7:LYS:HE2	2:C:39:ILE:O	2.17	0.44
1:D:94:MSE:HE1	2:F:8:CYS:CB	2.47	0.43
1:D:70:MSE:HA	1:D:73:GLU:HB2	2.00	0.43
1:D:114:VAL:HG21	1:D:137:VAL:HG21	1.99	0.43
2:E:42:CYS:SG	2:E:47:ALA:HB3	2.57	0.43
1:A:54:SER:O	1:A:58:LYS:HD3	2.18	0.43
1:D:123:ALA:HB3	1:D:124:SER:HB3	2.00	0.43
2:E:25:ASN:OD1	2:E:28:GLU:HG3	2.18	0.43
2:C:20:MET:HE1	2:C:59:PHE:HE1	1.84	0.42
1:A:100:MSE:HE3	2:C:2:LYS:HE3	2.01	0.42
1:A:87:MSE:HE1	2:C:19:LYS:HG3	2.01	0.42
1:D:80:LYS:HA	1:D:83:ILE:HD12	2.01	0.42
2:B:82:LEU:HD12	2:B:82:LEU:HA	1.90	0.41
2:F:2:LYS:HB2	2:F:5:GLU:HB2	2.03	0.41
1:D:114:VAL:CG2	1:D:137:VAL:HG21	2.51	0.41
2:E:16:LEU:CD1	2:E:71:LEU:HG	2.51	0.40
1:D:90:LEU:O	1:D:94:MSE:HG3	2.22	0.40

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	186/199 (94%)	176 (95%)	6 (3%)	4 (2%)	5	10

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	D	192/199 (96%)	183 (95%)	7 (4%)	2 (1%)	13	29
2	B	80/87 (92%)	80 (100%)	0	0	100	100
2	C	80/87 (92%)	77 (96%)	3 (4%)	0	100	100
2	E	79/87 (91%)	78 (99%)	1 (1%)	0	100	100
2	F	76/87 (87%)	76 (100%)	0	0	100	100
All	All	693/746 (93%)	670 (97%)	17 (2%)	6 (1%)	14	31

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	121	PRO
1	A	123	ALA
1	D	144	GLU
1	A	71	TYR
1	A	191	THR
1	D	122	ALA

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	159/160 (99%)	141 (89%)	18 (11%)	4	9
1	D	161/160 (101%)	141 (88%)	20 (12%)	4	7
2	B	77/80 (96%)	69 (90%)	8 (10%)	5	11
2	C	76/80 (95%)	69 (91%)	7 (9%)	7	15
2	E	76/80 (95%)	73 (96%)	3 (4%)	27	53
2	F	72/80 (90%)	67 (93%)	5 (7%)	13	28
All	All	621/640 (97%)	560 (90%)	61 (10%)	6	13

All (61) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	17	LYS
1	A	35	LEU
1	A	56	GLN
1	A	72	LYS
1	A	82	LYS
1	A	90	LEU
1	A	91	LEU
1	A	115	LEU
1	A	120	ILE
1	A	128	LEU
1	A	130	VAL
1	A	135	LEU
1	A	143	GLN
1	A	161	ASP
1	A	178	LEU
1	A	181	THR
1	A	184	LEU
1	A	190	ARG
2	B	7	LYS
2	B	14	LEU
2	B	23	GLU
2	B	52	LEU
2	B	71	LEU
2	B	79	THR
2	B	81	ILE
2	B	82	LEU
2	C	4	THR
2	C	16	LEU
2	C	19	LYS
2	C	60	LEU
2	C	66	CYS
2	C	77	GLU
2	C	79	THR
1	D	22	LEU
1	D	24	GLU
1	D	30	VAL
1	D	35	LEU
1	D	56	GLN
1	D	70	MSE
1	D	90	LEU
1	D	91	LEU
1	D	115	LEU
1	D	117	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	119	ILE
1	D	128	LEU
1	D	130	VAL
1	D	135	LEU
1	D	161	ASP
1	D	176	ASN
1	D	178	LEU
1	D	181	THR
1	D	184	LEU
1	D	190	ARG
2	E	5	GLU
2	E	7	LYS
2	E	71	LEU
2	F	3	LEU
2	F	17	LEU
2	F	49	LEU
2	F	60	LEU
2	F	71	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (11) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	28	ASN
1	A	104	ASN
1	A	148	HIS
1	A	163	ASN
2	C	10	ASN
2	C	63	HIS
1	D	33	GLN
1	D	37	GLN
1	D	104	ASN
1	D	163	ASN
2	E	40	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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 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, 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	185/199 (92%)	-0.05	8 (4%)	40	34	44, 62, 95, 129	0
1	D	189/199 (94%)	0.77	19 (10%)	14	12	58, 112, 218, 238	0
2	B	82/87 (94%)	-0.24	1 (1%)	76	72	48, 64, 84, 102	0
2	C	82/87 (94%)	0.01	1 (1%)	76	72	45, 62, 95, 122	0
2	E	81/87 (93%)	0.21	1 (1%)	76	72	56, 88, 129, 150	0
2	F	78/87 (89%)	0.36	3 (3%)	44	38	56, 82, 134, 161	0
All	All	697/746 (93%)	0.23	33 (4%)	37	32	44, 76, 159, 238	0

All (33) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	174	ALA	5.3
1	A	192	THR	4.8
1	D	175	ALA	4.2
1	D	169	LEU	3.8
1	D	41	ALA	3.7
1	A	123	ALA	3.5
1	D	35	LEU	3.4
1	D	178	LEU	3.4
1	D	167	VAL	3.2
1	A	71	TYR	3.0
2	C	81	ILE	3.0
1	D	125	ALA	3.0
1	A	124	SER	3.0
1	A	122	ALA	2.9
1	D	119	ILE	2.9
1	D	172	VAL	2.9
1	D	170	LYS	2.8
1	D	49	PHE	2.7
1	D	26	LYS	2.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	F	24	SER	2.5
1	D	173	THR	2.5
2	B	2	LYS	2.4
1	D	180	ILE	2.3
1	A	73	GLU	2.3
2	F	3	LEU	2.2
1	A	79	ARG	2.2
1	A	80	LYS	2.1
1	D	34	LEU	2.1
2	F	56	ILE	2.1
1	D	138	LEU	2.1
1	D	137	VAL	2.1
2	E	79	THR	2.0
1	D	168	HIS	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 ⓘ

There are no such residues in this entry.