wwPDB X-ray Structure Validation Summary Report (i)

Feb 21, 2024 – 10:01 PM EST

PDB ID	:	4R71
Title	:	Structure of the Qbeta holoenzyme complex in the P1211 crystal form
Authors	:	Gytz, H.; Seweryn, P.; Kutlubaeva, Z.; Chetverin, A.B.; Brodersen, D.E.; Knudsen, C.R.
Deposited on	:	2014-08-26
Resolution	:	3.21 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 (i) 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 (1)) were used in the production of this report:

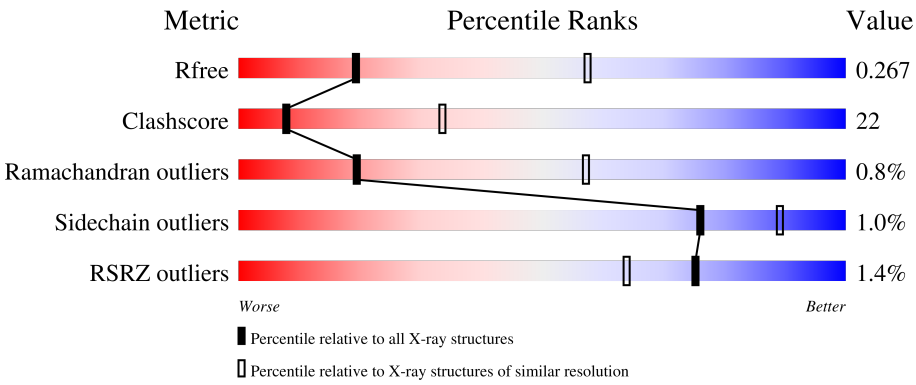
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 3.21 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.



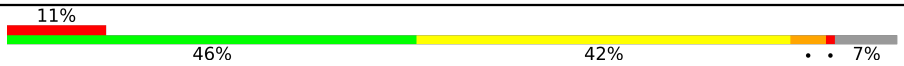
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1335 (3.24-3.20)
Clashscore	141614	1460 (3.24-3.20)
Ramachandran outliers	138981	1437 (3.24-3.20)
Sidechain outliers	138945	1436 (3.24-3.20)
RSRZ outliers	127900	1291 (3.24-3.20)

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	694	56% 36% • 6%
1	C	694	53% 40% 6%
2	B	595	.% 67% 26% 7%
2	D	595	.% 65% 27% 7%
3	E	171	6% 45% 42% 5% • 8%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	171	

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 21108 atoms, of which 1 is hydrogen 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 Elongation factor Ts, Elongation factor Tu.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	650	4968	3136	851	957	24	0	0	0
1	C	650	4967	3136	850	957	24	0	0	0

There are 34 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1000	HIS	-	linker	UNP P0CE48
A	1395	GLY	-	linker	UNP P0CE48
A	1396	ALA	-	linker	UNP P0CE48
A	1397	SER	-	linker	UNP P0CE48
A	1398	GLY	-	linker	UNP P0CE48
A	1399	ALA	-	linker	UNP P0CE48
A	1400	ALA	-	linker	UNP P0CE48
A	1401	GLY	-	linker	UNP P0CE48
A	1402	GLY	-	linker	UNP P0CE48
A	1403	GLY	-	linker	UNP P0CE48
A	1404	GLY	-	linker	UNP P0CE48
A	1405	GLU	-	linker	UNP P0CE48
A	1406	ASN	-	linker	UNP P0CE48
A	1407	LEU	-	linker	UNP P0CE48
A	1408	TYR	-	linker	UNP P0CE48
A	1409	PHE	-	linker	UNP P0CE48
A	1410	GLN	-	linker	UNP P0CE48
C	999	HIS	-	linker	UNP P0CE48
C	1394	GLY	-	linker	UNP P0CE48
C	1395	ALA	-	linker	UNP P0CE48
C	1396	SER	-	linker	UNP P0CE48
C	1397	GLY	-	linker	UNP P0CE48
C	1398	ALA	-	linker	UNP P0CE48
C	1399	ALA	-	linker	UNP P0CE48
C	1400	GLY	-	linker	UNP P0CE48

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	1401	GLY	-	linker	UNP P0CE48
C	1402	GLY	-	linker	UNP P0CE48
C	1403	GLY	-	linker	UNP P0CE48
C	1404	GLU	-	linker	UNP P0CE48
C	1405	ASN	-	linker	UNP P0CE48
C	1406	LEU	-	linker	UNP P0CE48
C	1407	TYR	-	linker	UNP P0CE48
C	1408	PHE	-	linker	UNP P0CE48
C	1409	GLN	-	linker	UNP P0CE48

- Molecule 2 is a protein called RNA-directed RNA polymerase beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	556	4375	2769	765	820	21	0	0	0
			Total	C	N	O	S			
2	D	553	4355	2758	761	815	21	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-5	SER	-	linker	UNP P14647
B	-4	GLY	-	linker	UNP P14647
B	-3	GLY	-	linker	UNP P14647
B	-2	GLY	-	linker	UNP P14647
B	-1	GLY	-	linker	UNP P14647
B	0	SER	-	linker	UNP P14647
D	-5	SER	-	linker	UNP P14647
D	-4	GLY	-	linker	UNP P14647
D	-3	GLY	-	linker	UNP P14647
D	-2	GLY	-	linker	UNP P14647
D	-1	GLY	-	linker	UNP P14647
D	0	SER	-	linker	UNP P14647

- Molecule 3 is a protein called 30S ribosomal protein S1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	H	N	O			
3	E	157	1215	766	1	209	239	0	0	0
			Total	C	N	O				
3	F	159	1228	774	211	243		0	0	0

There are 2 discrepancies between the modelled and reference sequences:

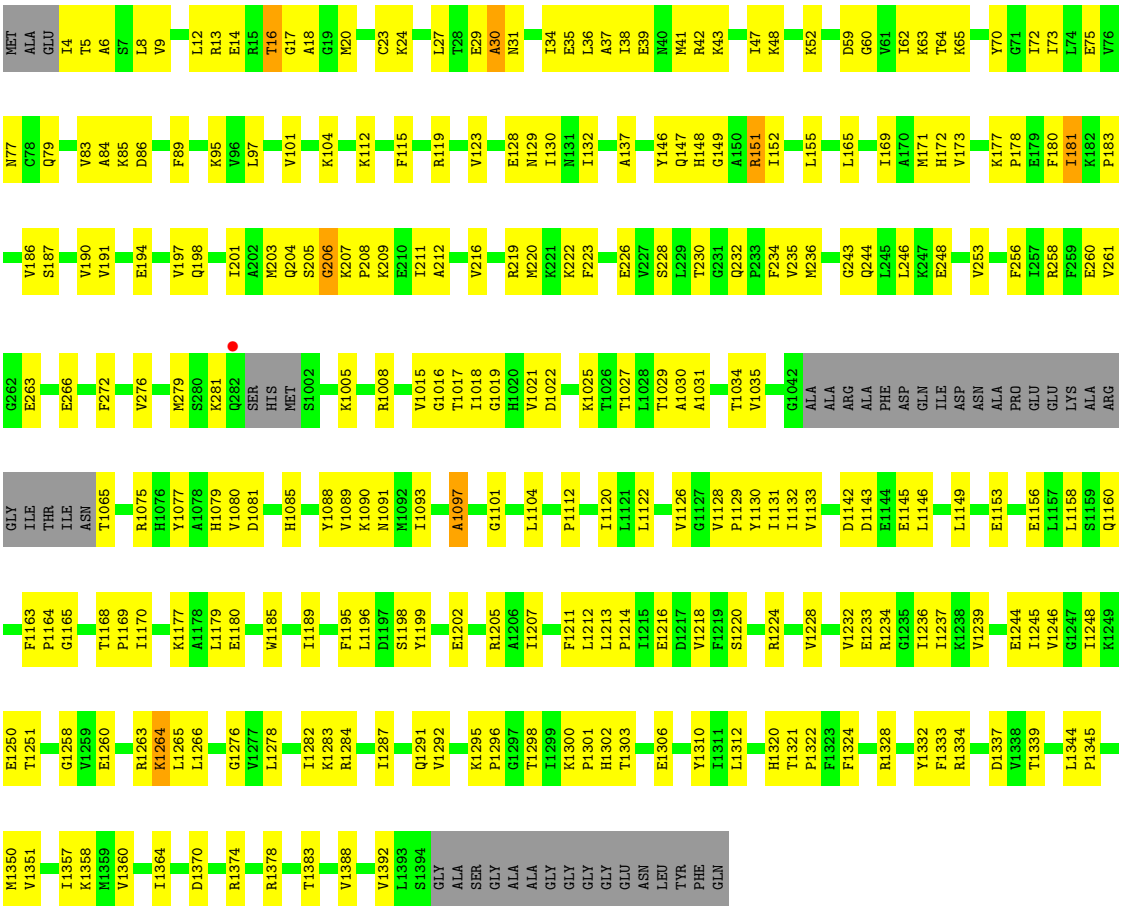
Chain	Residue	Modelled	Actual	Comment	Reference
E	1	GLY	-	expression tag	UNP P0AG67
F	1	GLY	-	expression tag	UNP P0AG67

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.

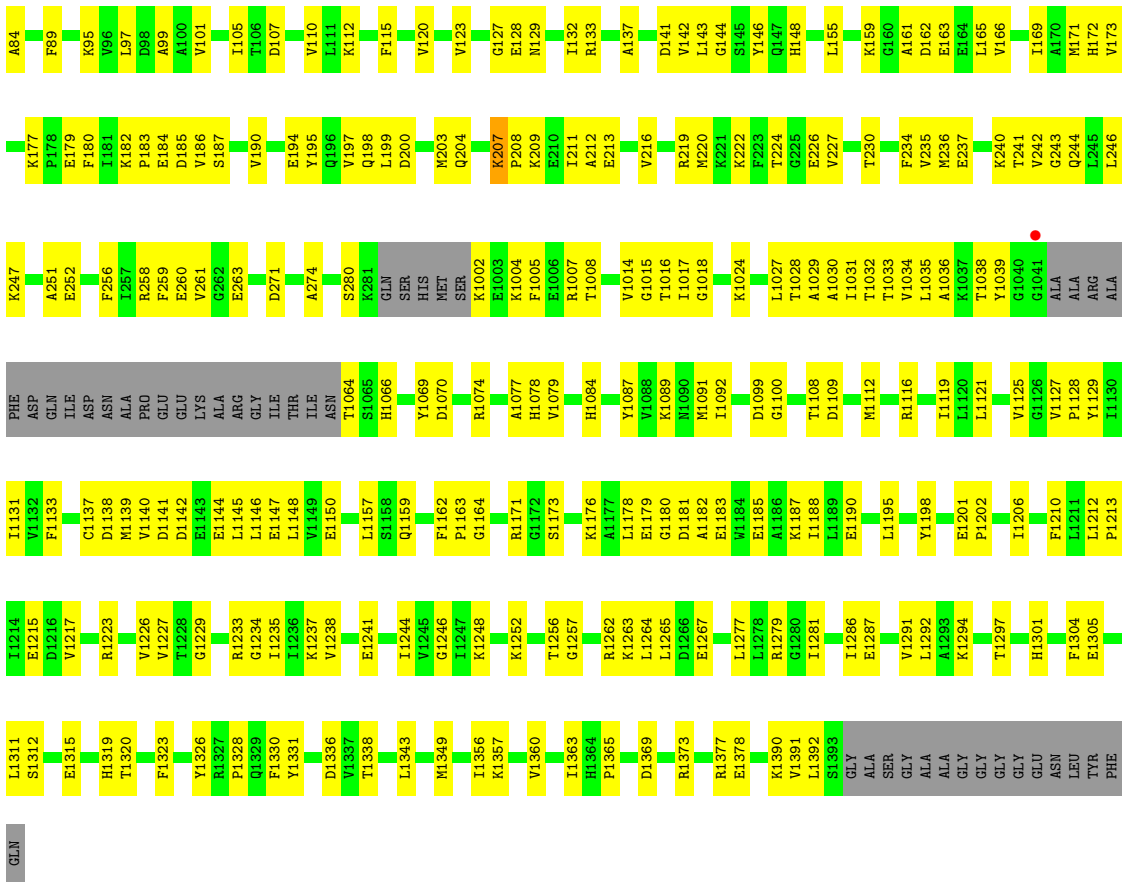
• Molecule 1: Elongation factor Ts, Elongation factor Tu

Chain A: 56% 36% • 6%



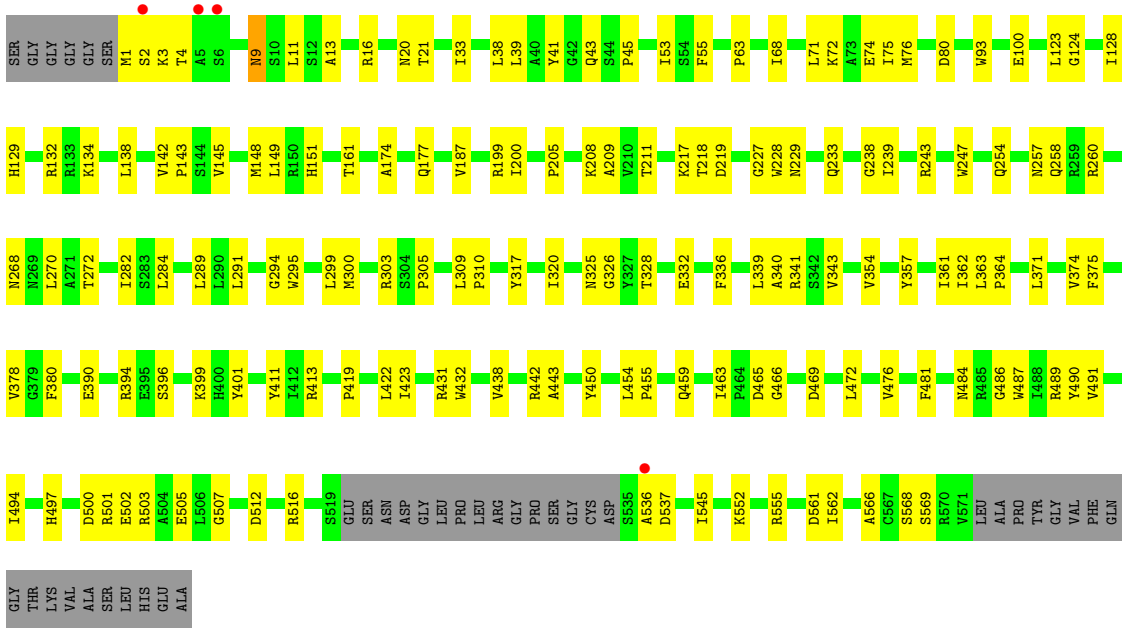
• Molecule 1: Elongation factor Ts, Elongation factor Tu

Chain C: 53% 40% 6%



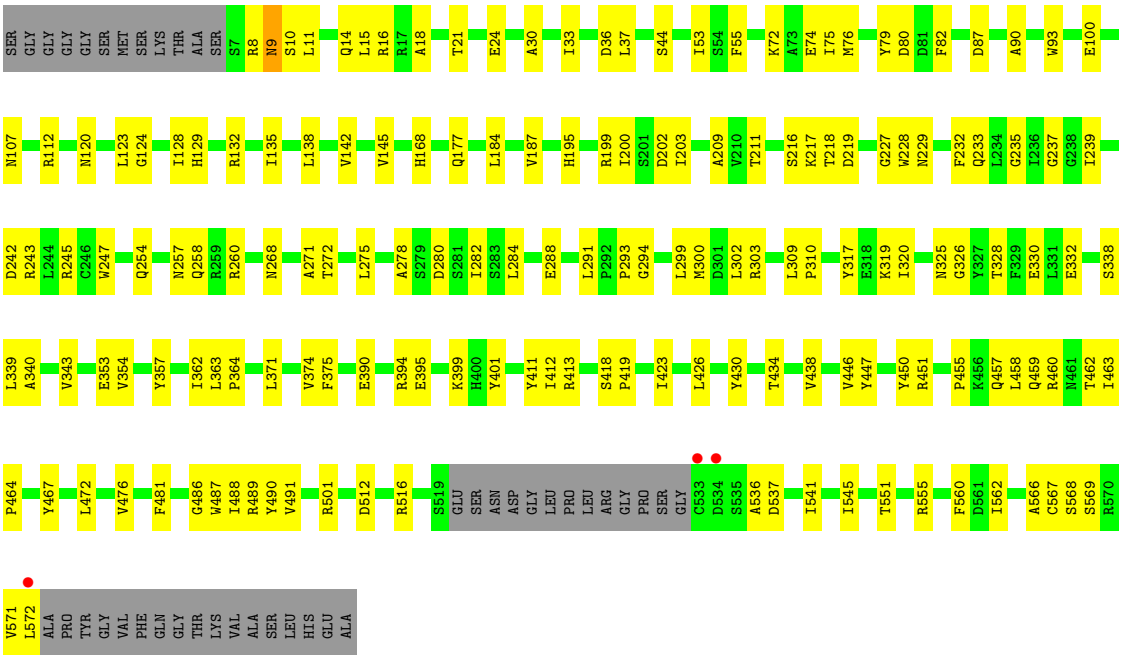
- Molecule 2: RNA-directed RNA polymerase beta chain

Chain B:



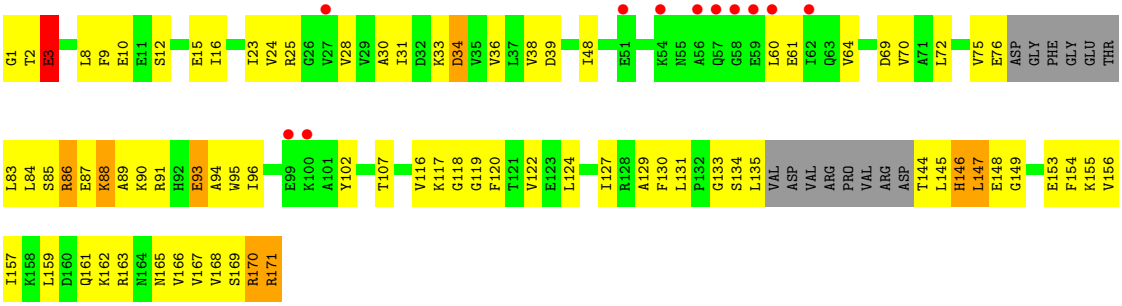
- Molecule 2: RNA-directed RNA polymerase beta chain

Chain D:



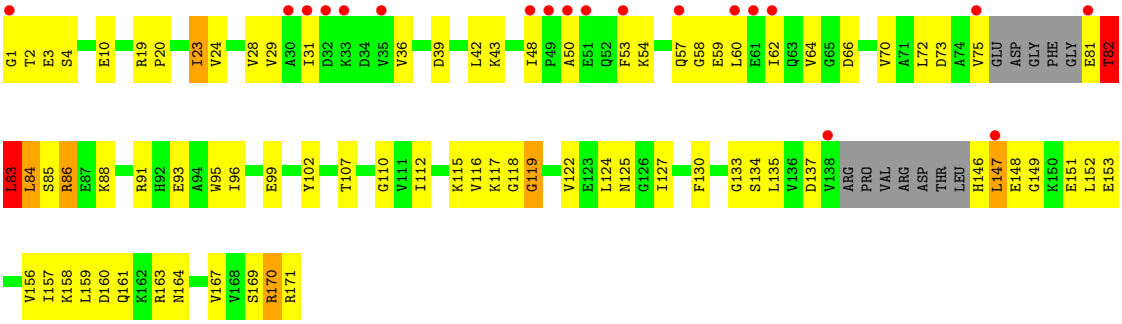
- Molecule 3: 30S ribosomal protein S1

Chain E:



- Molecule 3: 30S ribosomal protein S1

Chain F:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	99.73Å 115.45Å 178.52Å 90.00° 96.14° 90.00°	Depositor
Resolution (Å)	59.83 – 3.21 67.32 – 3.21	Depositor EDS
% Data completeness (in resolution range)	98.6 (59.83-3.21) 98.6 (67.32-3.21)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.13 (at 3.19Å)	Xtriage
Refinement program	PHENIX (phenix.refine: dev_1702)	Depositor
R, R_{free}	0.213 , 0.268 0.213 , 0.267	Depositor DCC
R_{free} test set	3261 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	73.4	Xtriage
Anisotropy	0.116	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 45.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	21108	wwPDB-VP
Average B, all atoms (Å^2)	78.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.97% 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	0.26	0/5042	0.44	0/6798
1	C	0.25	0/5041	0.43	0/6797
2	B	0.26	0/4471	0.42	0/6060
2	D	0.26	0/4451	0.41	0/6035
3	E	0.26	0/1225	0.45	0/1647
3	F	0.25	0/1239	0.45	0/1667
All	All	0.26	0/21469	0.43	0/29004

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	4968	0	5036	242	0
1	C	4967	0	5034	267	0
2	B	4375	0	4335	131	0
2	D	4355	0	4305	137	0
3	E	1214	1	1242	97	0
3	F	1228	0	1253	91	0
All	All	21107	1	21205	912	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 22.

The worst 5 of 912 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:1065:THR:HG22	2:B:536:ALA:HB1	1.23	1.15
1:C:208:PRO:HG2	1:C:211:ILE:HB	1.31	1.11
2:D:463:ILE:HG22	2:D:491:VAL:HG11	1.37	1.02
1:C:1147:GLU:HG3	1:C:1171:ARG:HH22	1.25	1.01
3:F:148:GLU:HB3	3:F:149:GLY:HA3	1.41	1.00

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	644/694 (93%)	609 (95%)	28 (4%)	7 (1%)	14 50
1	C	644/694 (93%)	614 (95%)	27 (4%)	3 (0%)	29 66
2	B	552/595 (93%)	524 (95%)	27 (5%)	1 (0%)	47 79
2	D	549/595 (92%)	521 (95%)	28 (5%)	0	100 100
3	E	151/171 (88%)	130 (86%)	15 (10%)	6 (4%)	3 20
3	F	153/171 (90%)	130 (85%)	18 (12%)	5 (3%)	4 24
All	All	2693/2920 (92%)	2528 (94%)	143 (5%)	22 (1%)	19 57

5 of 22 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	17	GLY
1	A	30	ALA
3	E	34	ASP
3	F	82	THR
3	F	83	LEU

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	527/555 (95%)	524 (99%)	3 (1%)	86	93
1	C	526/555 (95%)	525 (100%)	1 (0%)	93	97
2	B	478/505 (95%)	476 (100%)	2 (0%)	91	95
2	D	475/505 (94%)	474 (100%)	1 (0%)	93	97
3	E	131/143 (92%)	123 (94%)	8 (6%)	18	53
3	F	133/143 (93%)	125 (94%)	8 (6%)	19	53
All	All	2270/2406 (94%)	2247 (99%)	23 (1%)	76	89

5 of 23 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	E	171	ARG
3	F	81	GLU
3	F	23	ILE
3	F	82	THR
2	D	9	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	1319	HIS
2	D	14	GLN
2	D	457	GLN
1	C	55	ASN
1	A	1320	HIS

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

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 (i)

6.1 Protein, DNA and RNA chains (i)

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(Å^2)	Q<0.9
1	A	650/694 (93%)	-0.19	1 (0%)	95	94	30, 64, 128, 183	0
1	C	650/694 (93%)	-0.11	1 (0%)	95	94	35, 74, 126, 188	0
2	B	556/595 (93%)	-0.10	4 (0%)	87	82	34, 64, 104, 240	0
2	D	553/595 (92%)	-0.08	3 (0%)	91	86	40, 72, 113, 204	0
3	E	157/171 (91%)	0.36	11 (7%)	16	10	41, 104, 166, 195	0
3	F	159/171 (92%)	0.46	19 (11%)	4	3	64, 112, 188, 268	0
All	All	2725/2920 (93%)	-0.06	39 (1%)	75	64	30, 71, 138, 268	0

The worst 5 of 39 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	F	1	GLY	9.6
3	F	57	GLN	5.4
2	D	533	CYS	4.9
3	F	50	ALA	4.5
2	D	534	ASP	4.3

6.2 Non-standard residues in protein, DNA, RNA chains (i)

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

6.3 Carbohydrates (i)

There are no monosaccharides 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.