



wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 15, 2021 – 10:14 PM JST

PDB ID : 7DKH
Title : Crystal structure of the Ctr9/Paf1/Cdc73/Rtf1 quaternary complex
Authors : Chen, F.L.; Liu, B.B.; Guo, L.; Li, D.F.; Zhou, H.; Long, J.F.
Deposited on : 2020-11-24
Resolution : 2.90 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 ⓘ symbol.

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

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.23.2
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) : 2.23.2

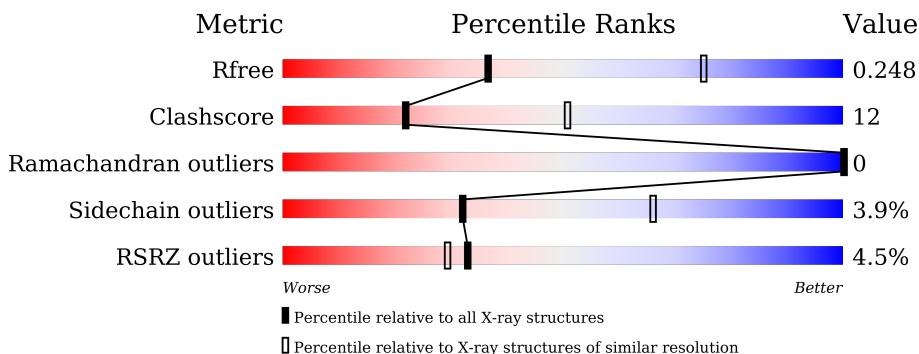
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.90 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.







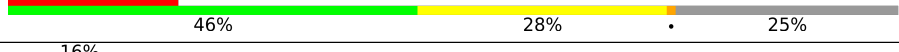
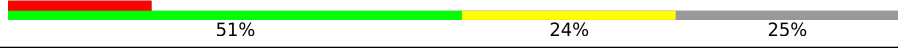
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)

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 $\leq 5\%$. 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	972	 3% 67% 25% • 6%
1	E	972	 6% 68% 25% • 6%
1	I	972	 3% 68% 25% • 6%
2	B	103	 % 67% 29% •
2	F	103	 % 70% 24% • 5%
2	J	103	 % 76% 20% •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	C	57	 58% 32% 9%
3	G	57	 56% 33% 11%
3	K	57	 60% 28% 11%
4	D	68	 9% 62% 13% 25%
4	H	68	 19% 46% 28% 25%
4	L	68	 16% 51% 24% 25%

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 27198 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 RNA polymerase-associated protein CTR9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	N	O	S				Se
1	A	913	Total 7448	C 4751	N 1259	O 1416	S 7	Se 15	0	0	0
1	E	913	Total 7448	C 4751	N 1259	O 1416	S 7	Se 15	0	0	0
1	I	914	Total 7457	C 4757	N 1261	O 1417	S 7	Se 15	0	0	0

- Molecule 2 is a protein called RNA polymerase II-associated protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	Se			
2	B	99	Total 797	C 512	N 129	O 154	Se 2	0	0	0
2	F	98	Total 789	C 506	N 128	O 153	Se 2	0	0	0
2	J	99	Total 797	C 512	N 129	O 154	Se 2	0	0	0

- Molecule 3 is a protein called Cell division control protein 73.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	Se			
3	C	52	Total 403	C 254	N 70	O 78	Se 1	0	0	0
3	G	51	Total 395	C 250	N 69	O 75	Se 1	0	0	0
3	K	51	Total 395	C 250	N 69	O 75	Se 1	0	0	0

- Molecule 4 is a protein called RNA polymerase-associated protein RTF1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	Se			
4	D	51	Total 423	C 267	N 76	O 79	Se 1	0	0	0

Continued on next page...

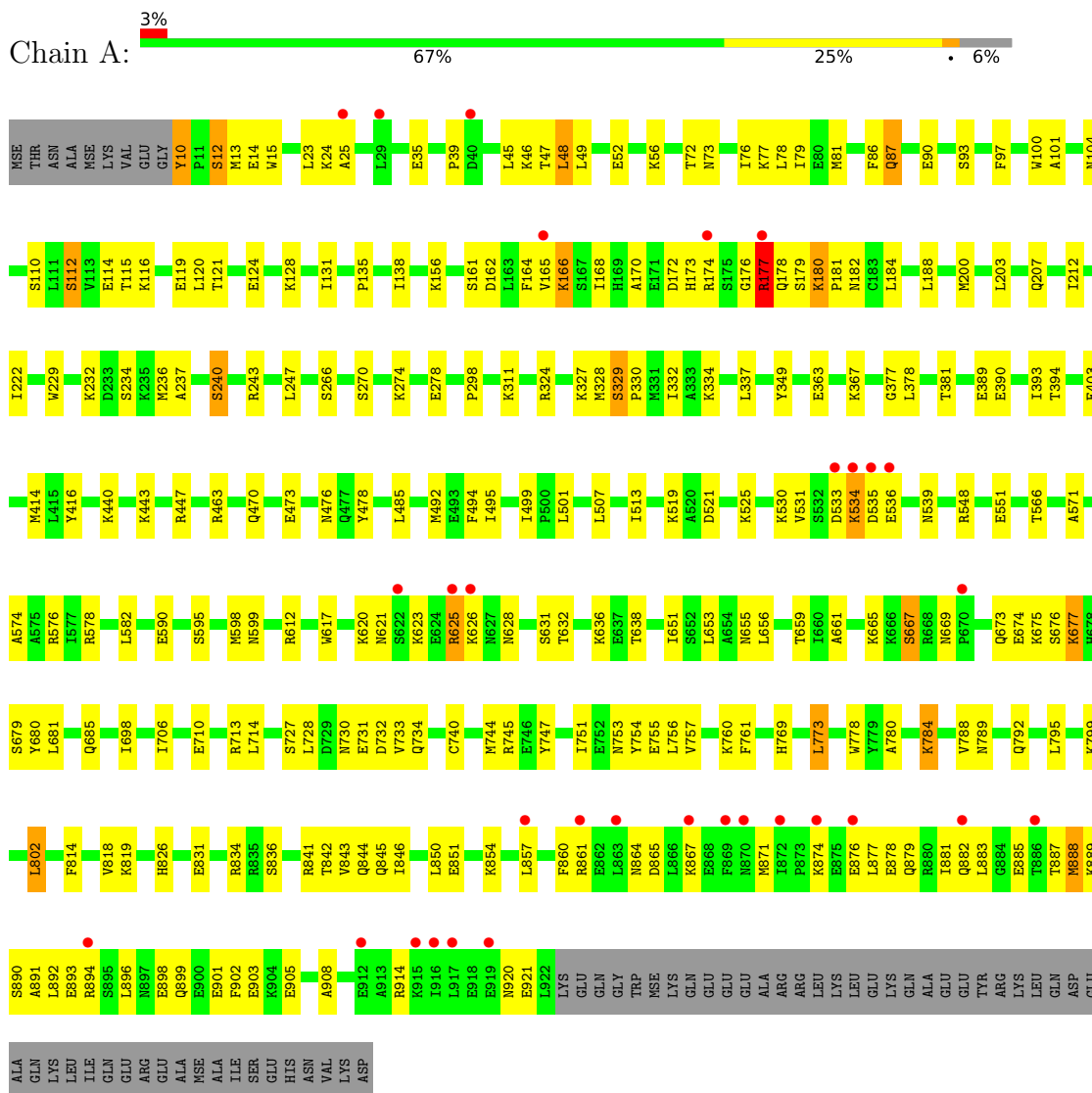
Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	H	51	Total	C	N	O	Se	0	0	0
			423	267	76	79	1			
4	L	51	Total	C	N	O	Se	0	0	0
			423	267	76	79	1			

3 Residue-property plots [i](#)

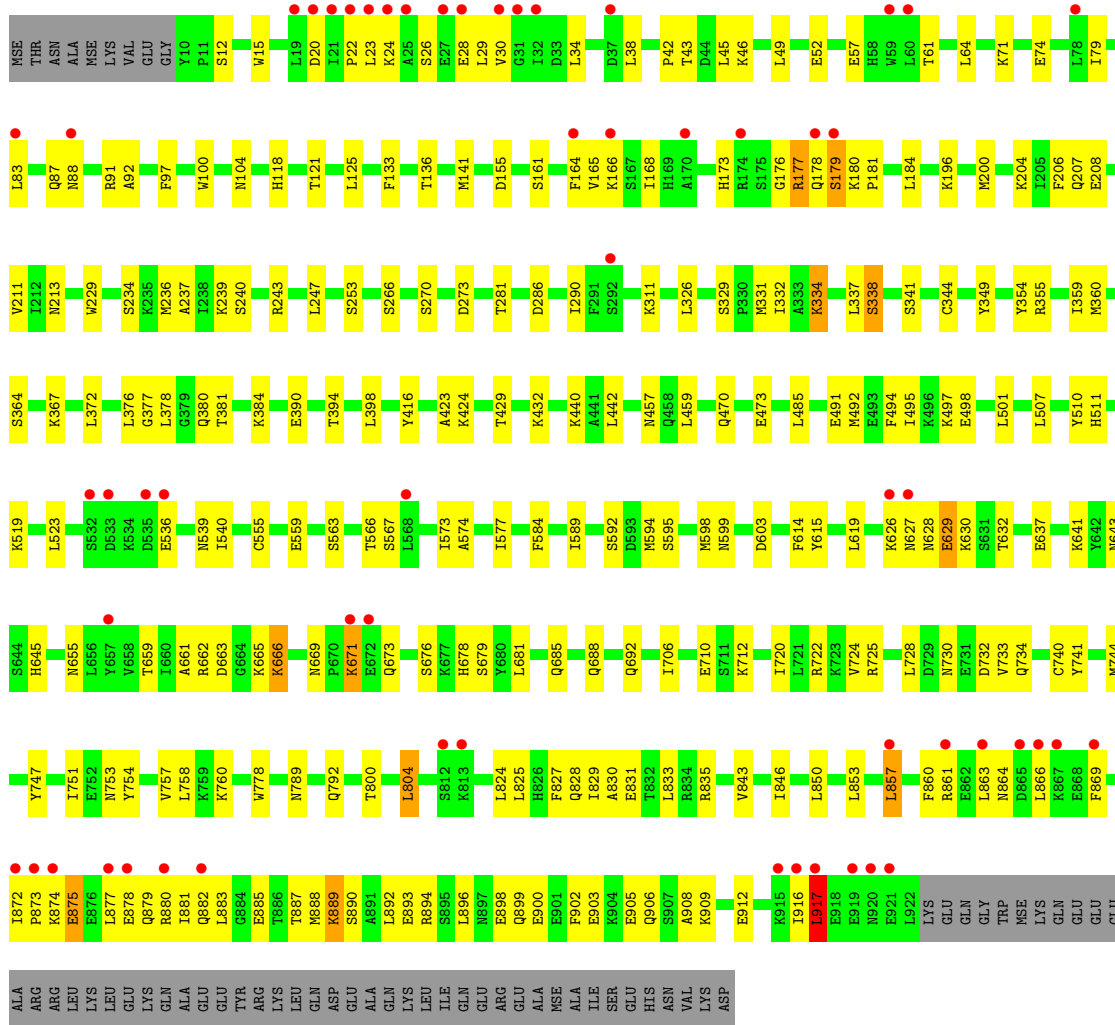
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: RNA polymerase-associated protein CTR9

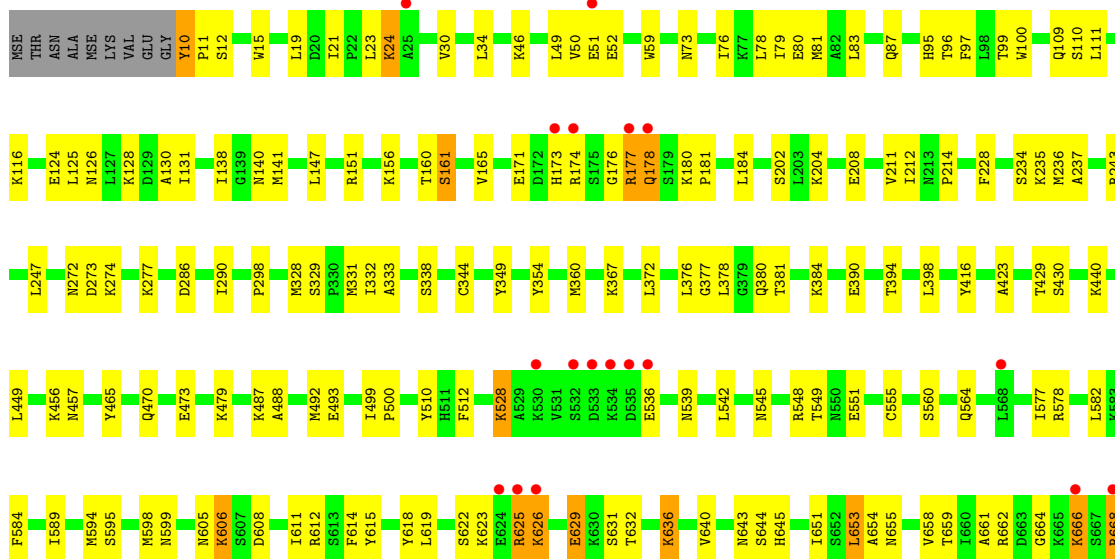


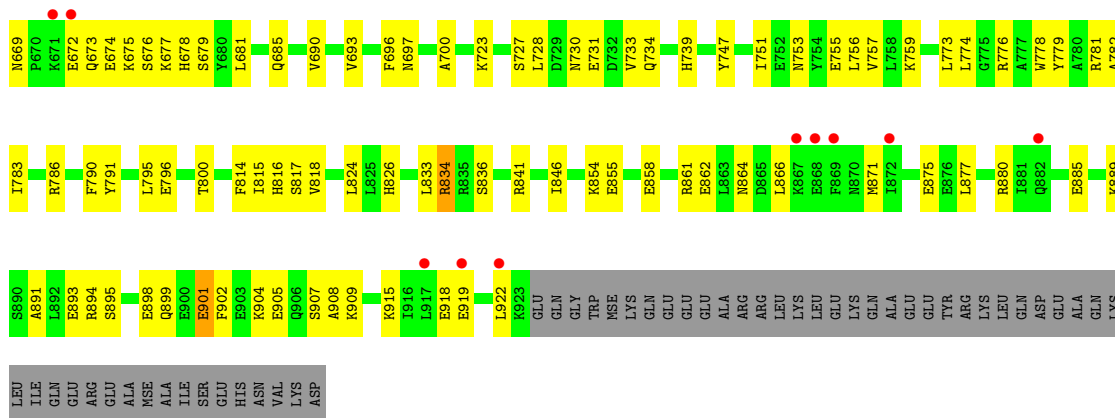
- Molecule 1: RNA polymerase-associated protein CTR9





● Molecule 1: RNA polymerase-associated protein CTR9





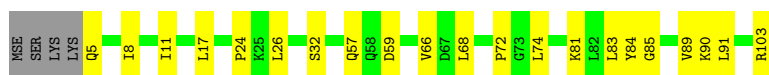
● Molecule 2: RNA polymerase II-associated protein 1



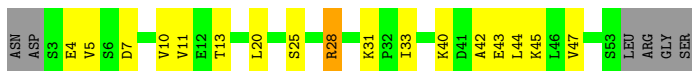
● Molecule 2: RNA polymerase II-associated protein 1



● Molecule 2: RNA polymerase II-associated protein 1

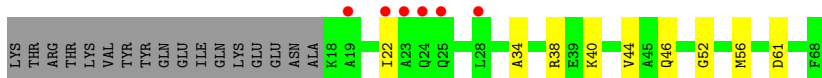


Chain K:  60% 28% 11%



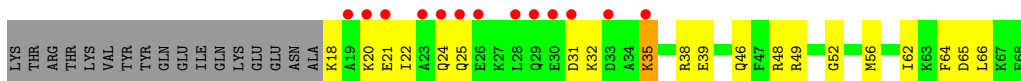
• Molecule 4: RNA polymerase-associated protein RTF1

Chain D:  9% 62% 13% 25%



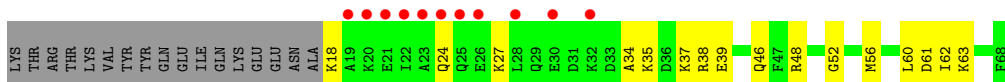
• Molecule 4: RNA polymerase-associated protein RTF1

Chain H:  19% 46% 28% 25%



• Molecule 4: RNA polymerase-associated protein RTF1

Chain L:  16% 51% 24% 25%



4 Data and refinement statistics i

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α , β , γ	190.19Å 190.19Å 216.42Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	48.96 – 2.90 48.96 – 2.90	Depositor EDS
% Data completeness (in resolution range)	99.8 (48.96-2.90) 99.8 (48.96-2.90)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.41 (at 2.91Å)	Xtrriage
Refinement program	PHENIX 1.14 3260	Depositor
R, R_{free}	0.193 , 0.248 0.193 , 0.248	Depositor DCC
R_{free} test set	5119 reflections (5.11%)	wwPDB-VP
Wilson B-factor (Å ²)	73.5	Xtrriage
Anisotropy	0.332	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 68.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.000 for -h,-k,l	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	27198	wwPDB-VP
Average B, all atoms (Å ²)	90.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 22.66 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 5.4490e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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	0.31	0/7569	0.54	1/10175 (0.0%)
1	E	0.28	0/7569	0.54	2/10175 (0.0%)
1	I	0.29	0/7578	0.52	1/10186 (0.0%)
2	B	0.28	0/811	0.57	0/1100
2	F	0.27	0/802	0.59	0/1086
2	J	0.29	0/811	0.60	0/1100
3	C	0.27	0/407	0.58	1/546 (0.2%)
3	G	0.26	0/399	0.53	0/535
3	K	0.29	0/399	0.59	0/535
4	D	0.28	0/424	0.52	0/556
4	H	0.28	0/424	0.62	0/556
4	L	0.27	0/424	0.59	0/556
All	All	0.29	0/27617	0.54	5/37106 (0.0%)

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	1
1	E	0	3
2	F	0	1
All	All	0	5

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	I	125	LEU	CA-CB-CG	6.46	130.15	115.30
1	E	917	LEU	CA-CB-CG	6.26	129.69	115.30
1	A	177	ARG	CG-CD-NE	5.47	123.30	111.80

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	125	LEU	CA-CB-CG	5.29	127.47	115.30
3	C	44	LEU	CA-CB-CG	5.08	126.98	115.30

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	87	GLN	Peptide
1	E	626	LYS	Peptide
1	E	629	GLU	Peptide
1	E	87	GLN	Peptide
2	F	85	GLY	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	7448	0	7513	205	8
1	E	7448	0	7513	197	5
1	I	7457	0	7526	175	5
2	B	797	0	815	30	0
2	F	789	0	803	20	1
2	J	797	0	815	28	2
3	C	403	0	414	15	0
3	G	395	0	410	11	0
3	K	395	0	410	17	0
4	D	423	0	444	6	0
4	H	423	0	444	25	0
4	L	423	0	444	12	0
All	All	27198	0	27551	646	11

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 12.

The worst 5 of 646 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:E:827:PHE:CD2	1:E:880:ARG:HD3	1.86	1.10
1:E:273:ASP:CG	4:H:56:MSE:HE2	1.77	1.04
1:A:176:GLY:O	1:A:177:ARG:HD2	1.59	1.01
1:E:23:LEU:HD23	1:E:24:LYS:H	1.29	0.93
4:H:18:LYS:HD2	4:H:22:ILE:HG22	1.48	0.93

The worst 5 of 11 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)
1:A:179:SER:OG	1:E:177:ARG:NH1[6_655]	1.54	0.66
1:A:179:SER:OG	1:E:177:ARG:CZ[6_655]	1.83	0.37
1:A:179:SER:OG	1:E:177:ARG:NH2[6_655]	1.88	0.32
1:A:172:ASP:OD2	1:E:177:ARG:NH2[6_655]	1.98	0.22
1:A:179:SER:CB	1:E:177:ARG:NH1[6_655]	2.00	0.20

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	911/972 (94%)	874 (96%)	37 (4%)	0	100	100
1	E	911/972 (94%)	867 (95%)	44 (5%)	0	100	100
1	I	912/972 (94%)	865 (95%)	47 (5%)	0	100	100
2	B	97/103 (94%)	88 (91%)	9 (9%)	0	100	100
2	F	94/103 (91%)	84 (89%)	10 (11%)	0	100	100
2	J	97/103 (94%)	86 (89%)	11 (11%)	0	100	100
3	C	50/57 (88%)	47 (94%)	3 (6%)	0	100	100
3	G	49/57 (86%)	46 (94%)	3 (6%)	0	100	100
3	K	49/57 (86%)	47 (96%)	2 (4%)	0	100	100
4	D	49/68 (72%)	47 (96%)	2 (4%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
4	H	49/68 (72%)	45 (92%)	4 (8%)	0	100	100
4	L	49/68 (72%)	49 (100%)	0	0	100	100
All	All	3317/3600 (92%)	3145 (95%)	172 (5%)	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	821/853 (96%)	787 (96%)	34 (4%)	30	64
1	E	821/853 (96%)	792 (96%)	29 (4%)	36	70
1	I	822/853 (96%)	783 (95%)	39 (5%)	26	59
2	B	94/95 (99%)	94 (100%)	0	100	100
2	F	93/95 (98%)	88 (95%)	5 (5%)	22	54
2	J	94/95 (99%)	94 (100%)	0	100	100
3	C	46/49 (94%)	45 (98%)	1 (2%)	52	81
3	G	45/49 (92%)	42 (93%)	3 (7%)	16	43
3	K	45/49 (92%)	44 (98%)	1 (2%)	52	81
4	D	44/59 (75%)	43 (98%)	1 (2%)	50	80
4	H	44/59 (75%)	41 (93%)	3 (7%)	16	42
4	L	44/59 (75%)	42 (96%)	2 (4%)	27	61
All	All	3013/3168 (95%)	2895 (96%)	118 (4%)	32	66

5 of 118 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	760	LYS
1	I	834	ARG
3	G	6	SER
1	I	759	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	I	626	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 19 such sidechains are listed below:

Mol	Chain	Res	Type
1	I	526	GLN
1	I	864	ASN
1	I	879	GLN
1	I	697	ASN
1	E	628	ASN

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	898/972 (92%)	0.20	31 (3%) 44 38	36, 78, 142, 193	0
1	E	898/972 (92%)	0.32	57 (6%) 20 16	45, 87, 168, 237	0
1	I	899/972 (92%)	0.15	28 (3%) 49 44	40, 82, 140, 194	0
2	B	97/103 (94%)	0.12	1 (1%) 82 82	43, 73, 112, 151	0
2	F	96/103 (93%)	0.24	1 (1%) 82 82	50, 80, 132, 163	0
2	J	97/103 (94%)	0.09	0 100 100	44, 73, 113, 125	0
3	C	51/57 (89%)	-0.07	0 100 100	50, 79, 128, 135	0
3	G	50/57 (87%)	0.01	0 100 100	55, 90, 141, 154	0
3	K	50/57 (87%)	-0.10	0 100 100	70, 99, 133, 141	0
4	D	50/68 (73%)	0.64	6 (12%) 4 3	53, 94, 151, 168	0
4	H	50/68 (73%)	0.99	13 (26%) 0 0	59, 106, 169, 201	0
4	L	50/68 (73%)	1.12	11 (22%) 0 0	55, 111, 174, 186	0
All	All	3286/3600 (91%)	0.24	148 (4%) 33 29	36, 83, 153, 237	0

The worst 5 of 148 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	L	28	LEU	7.8
1	E	20	ASP	6.7
1	E	869	PHE	6.2
1	E	21	ILE	6.2
1	E	170	ALA	6.0

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.