wwPDB X-ray Structure Validation Summary Report (i)

Mar 4, 2024 – 09:58 PM EST

PDB ID	:	1X0U
Title	:	Crystal Structure of the carboxyl transferase subunit of putative PCC of Sulfolobus tokodaii
Authors	:	Kakuta, Y.; Sueda, S.; Kondo, H.
Deposited on	:	2005-03-29
Resolution	:	2.20 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:

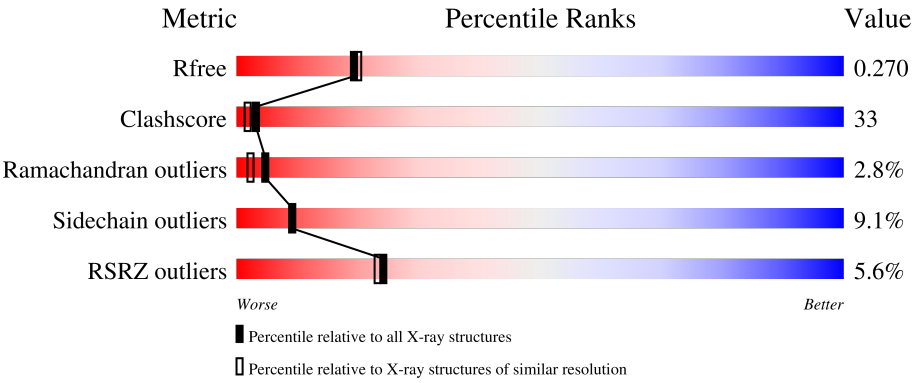
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 2.20 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.



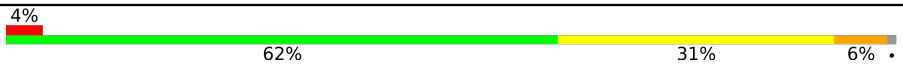
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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	522	5% 57% 35% 6% ··
1	B	522	6% 61% 30% 8% ·
1	C	522	8% 55% 35% 9% ·
1	D	522	4% 58% 33% 8% ··
1	E	522	6% 59% 34% 5% ··

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	522	

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 25335 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 hypothetical methylmalonyl-CoA decarboxylase alpha subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	518	Total	C	N	O	S	0	0	0
			4002	2560	683	746	13			
1	B	518	Total	C	N	O	S	0	0	0
			4002	2560	683	746	13			
1	C	518	Total	C	N	O	S	0	0	0
			4002	2560	683	746	13			
1	D	518	Total	C	N	O	S	0	0	0
			4002	2560	683	746	13			
1	E	518	Total	C	N	O	S	0	0	0
			4002	2560	683	746	13			
1	F	518	Total	C	N	O	S	0	0	0
			4002	2560	683	746	13			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	2	ALA	SER	conflict	UNP Q974R9
B	2	ALA	SER	conflict	UNP Q974R9
C	2	ALA	SER	conflict	UNP Q974R9
D	2	ALA	SER	conflict	UNP Q974R9
E	2	ALA	SER	conflict	UNP Q974R9
F	2	ALA	SER	conflict	UNP Q974R9

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	216	Total	O	0	0
			216	216		
2	B	214	Total	O	0	0
			214	214		
2	C	205	Total	O	0	0
			205	205		

Continued on next page...

Continued from previous page...

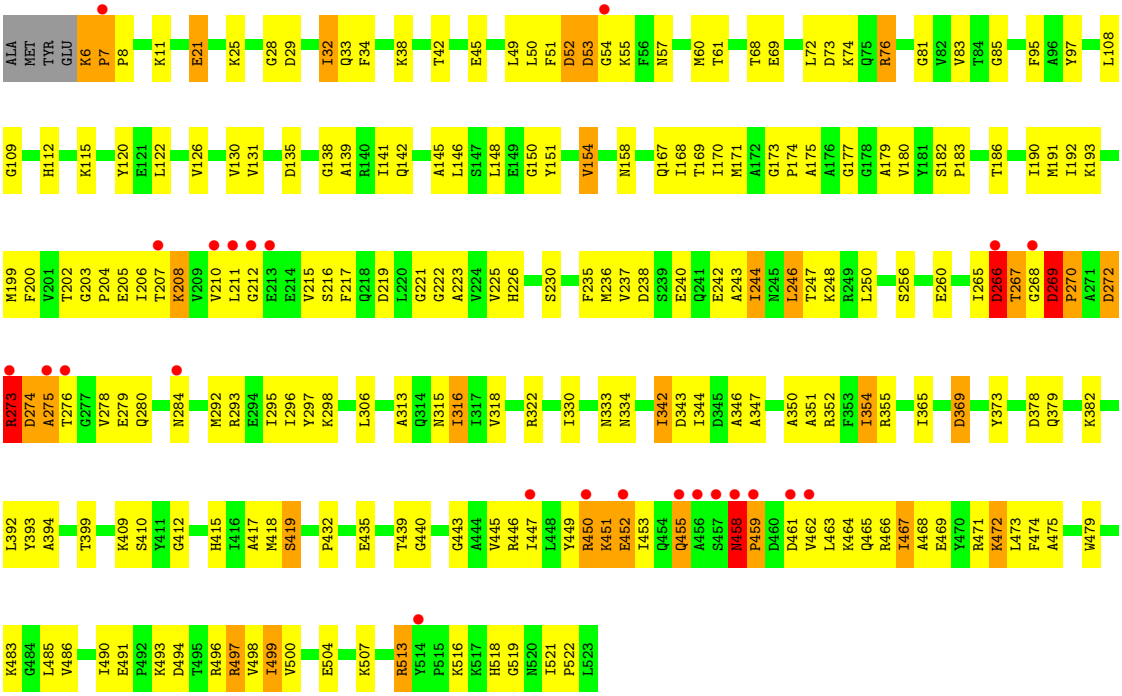
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	D	235	Total O 235 235	0	0
2	E	230	Total O 230 230	0	0
2	F	223	Total O 223 223	0	0

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.

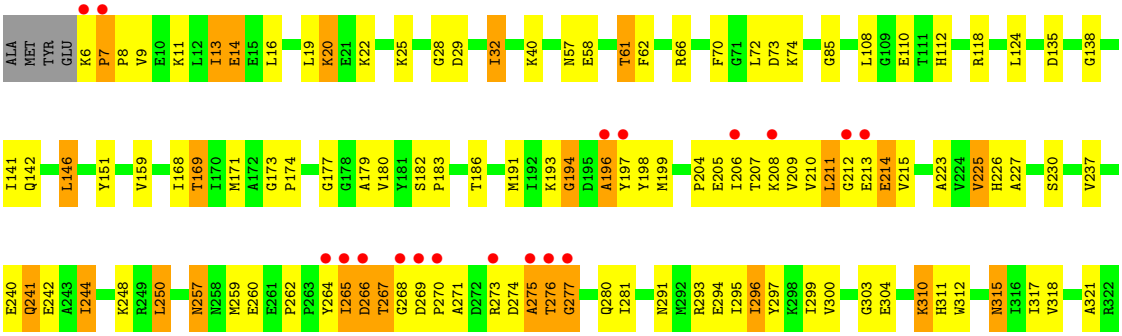
• Molecule 1: hypothetical methylmalonyl-CoA decarboxylase alpha subunit

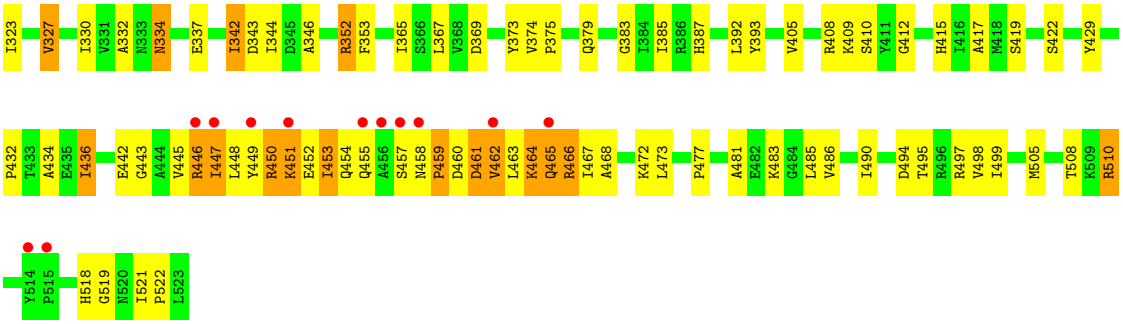
Chain A:



• Molecule 1: hypothetical methylmalonyl-CoA decarboxylase alpha subunit

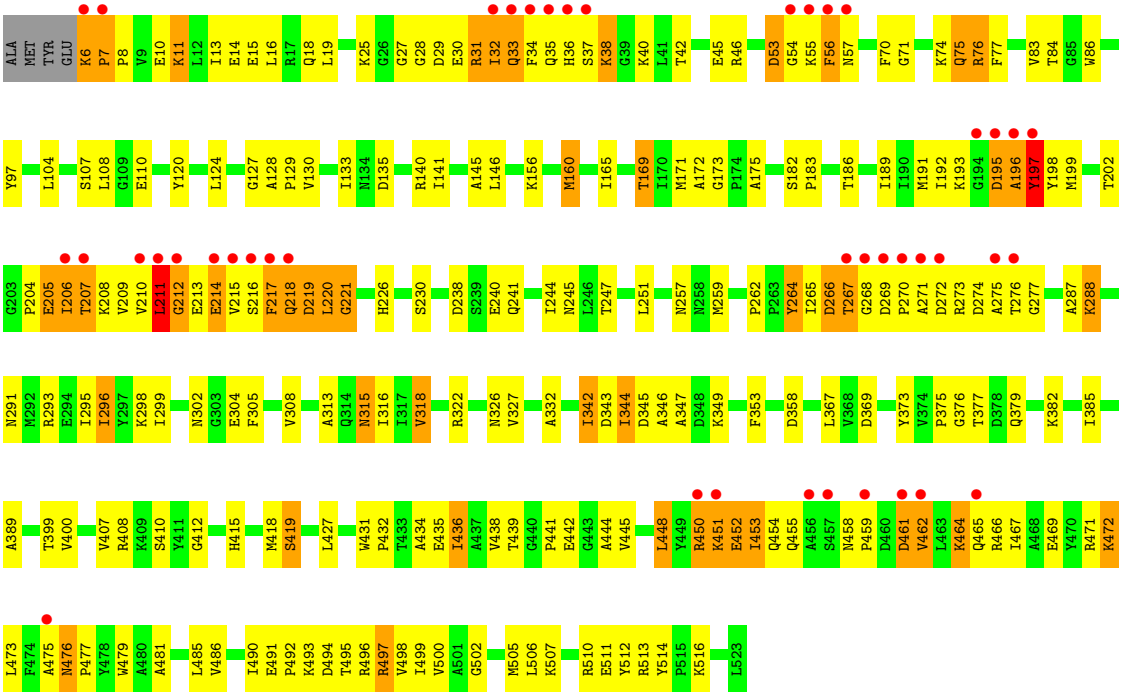
Chain B:





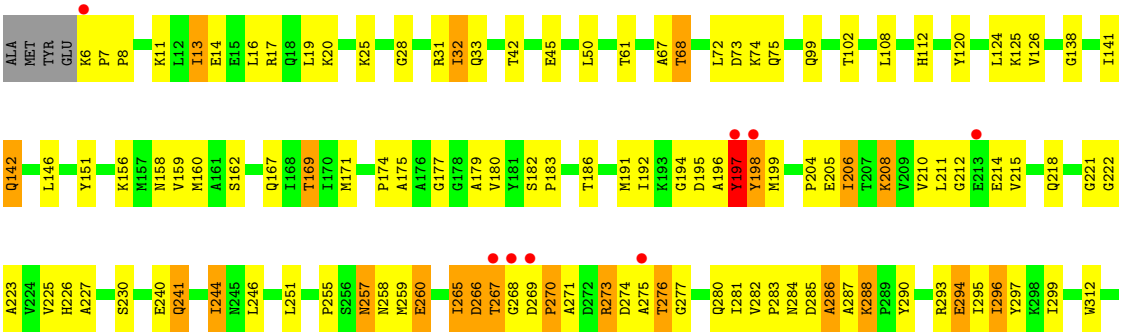
• Molecule 1: hypothetical methylmalonyl-CoA decarboxylase alpha subunit

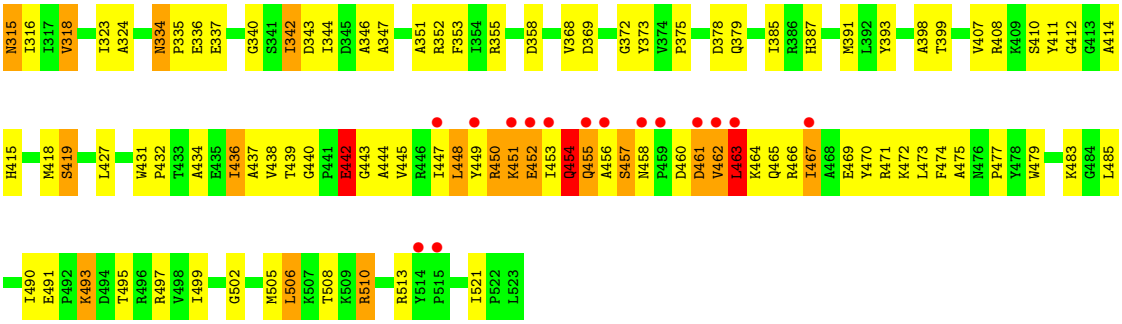
Chain C:



• Molecule 1: hypothetical methylmalonyl-CoA decarboxylase alpha subunit

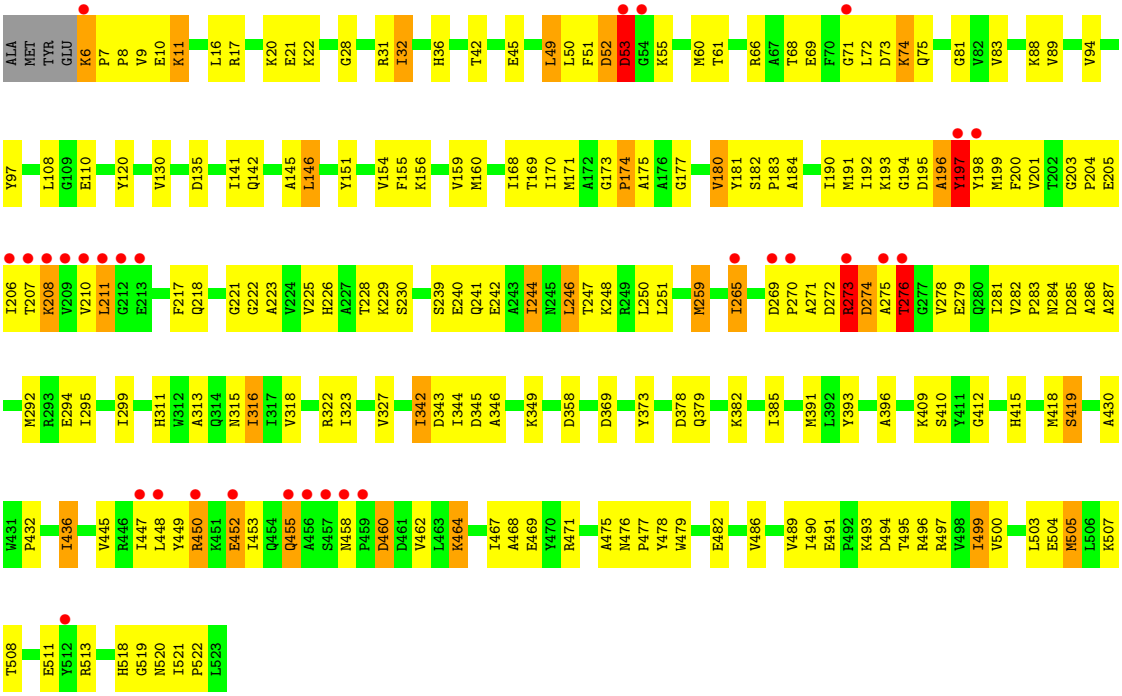
Chain D:





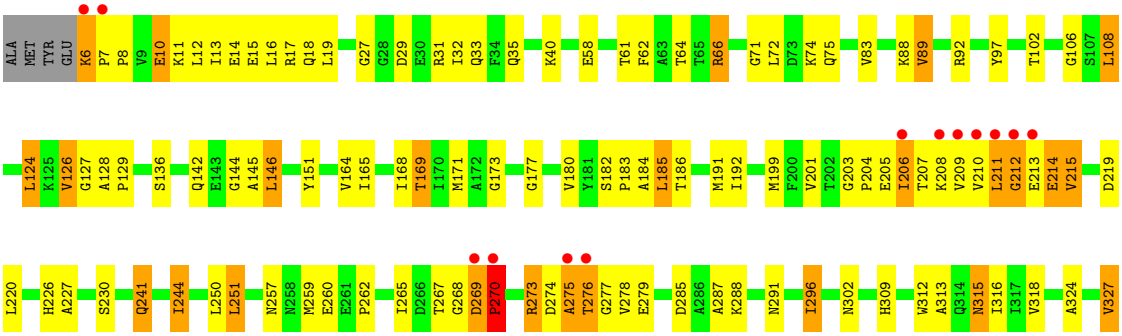
• Molecule 1: hypothetical methylmalonyl-CoA decarboxylase alpha subunit

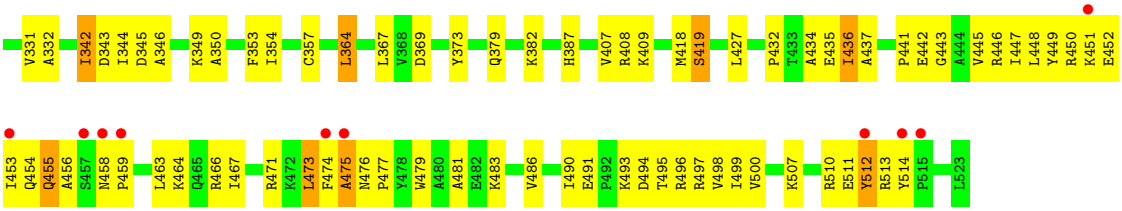
Chain E:



• Molecule 1: hypothetical methylmalonyl-CoA decarboxylase alpha subunit

Chain F:




V331
A332
I342
D343
I344
D345
A346
K349
A350
F353
I354
C357
L364
L367
V368
D369
Y373
Q379
K382
H387
V407
R408
K409
M418
S419
L427
P432
T433
A434
E435
T436
A437
P441
E442
G443
A444
V445
R446
I447
L448
Y449
R450
K451
E452
I453
Q454
Q455
A456
S457
N458
P459
L463
K464
Q465
R466
I467
R471
K472
L473
F474
A475
N476
P477
Y478
W479
A480
A481
E482
K483
V486
I490
E491
P492
K493
D494
T495
R496
R497
V498
I499
V500
K507
R510
E511
Y512
R513
Y514
P515
L523

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	99.61Å 181.47Å 112.70Å 90.00° 116.21° 90.00°	Depositor
Resolution (Å)	48.03 – 2.20 48.03 – 2.20	Depositor EDS
% Data completeness (in resolution range)	92.4 (48.03-2.20) 92.4 (48.03-2.20)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	0.10	Depositor
$< I/\sigma(I) >$ [1]	3.61 (at 2.20Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.228 , 0.271 0.227 , 0.270	Depositor DCC
R_{free} test set	9045 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	16.7	Xtriage
Anisotropy	1.205	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 16.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.479 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	25335	wwPDB-VP
Average B, all atoms (Å^2)	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.73% 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	0.39	0/4086	0.67	2/5531 (0.0%)
1	B	0.37	0/4086	0.64	0/5531
1	C	0.38	0/4086	0.66	4/5531 (0.1%)
1	D	0.37	0/4086	0.65	2/5531 (0.0%)
1	E	0.38	0/4086	0.66	0/5531
1	F	0.36	0/4086	0.63	0/5531
All	All	0.38	0/24516	0.65	8/33186 (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

There are no bond length outliers.

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	463	LEU	CA-CB-CG	6.71	130.73	115.30
1	C	53	ASP	N-CA-C	6.57	128.74	111.00
1	C	197	TYR	N-CA-C	5.92	126.97	111.00
1	D	454	GLN	N-CA-C	-5.37	96.51	111.00
1	A	269	ASP	CB-CG-OD2	5.24	123.01	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	269	ASP	Peptide

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	4002	0	4019	302	0
1	B	4002	0	4019	274	0
1	C	4002	0	4019	339	0
1	D	4002	0	4019	286	0
1	E	4002	0	4019	268	0
1	F	4002	0	4019	246	0
2	A	216	0	0	5	0
2	B	214	0	0	4	0
2	C	205	0	0	7	0
2	D	235	0	0	9	0
2	E	230	0	0	9	0
2	F	223	0	0	7	0
All	All	25335	0	24114	1609	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 33.

The worst 5 of 1609 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:273:ARG:HB3	1:E:497:ARG:HG3	1.21	1.18
1:A:275:ALA:HA	1:A:493:LYS:HD2	1.25	1.16
1:B:466:ARG:NH2	1:B:468:ALA:H	1.46	1.12
1:F:273:ARG:HD3	1:F:493:LYS:HD2	1.32	1.11
1:B:466:ARG:HG2	1:B:467:ILE:H	1.13	1.09

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	516/522 (99%)	471 (91%)	31 (6%)	14 (3%)	5 2
1	B	516/522 (99%)	471 (91%)	30 (6%)	15 (3%)	4 2
1	C	516/522 (99%)	461 (89%)	37 (7%)	18 (4%)	3 1
1	D	516/522 (99%)	474 (92%)	25 (5%)	17 (3%)	4 2
1	E	516/522 (99%)	468 (91%)	37 (7%)	11 (2%)	7 4
1	F	516/522 (99%)	479 (93%)	25 (5%)	12 (2%)	6 3
All	All	3096/3132 (99%)	2824 (91%)	185 (6%)	87 (3%)	5 2

5 of 87 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	7	PRO
1	A	52	ASP
1	A	267	THR
1	A	270	PRO
1	A	273	ARG

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	413/416 (99%)	376 (91%)	37 (9%)	9 9
1	B	413/416 (99%)	372 (90%)	41 (10%)	8 7
1	C	413/416 (99%)	372 (90%)	41 (10%)	8 7
1	D	413/416 (99%)	376 (91%)	37 (9%)	9 9
1	E	413/416 (99%)	377 (91%)	36 (9%)	10 10
1	F	413/416 (99%)	380 (92%)	33 (8%)	12 12
All	All	2478/2496 (99%)	2253 (91%)	225 (9%)	9 9

5 of 225 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	462	VAL
1	F	455	GLN
1	D	436	ILE
1	F	436	ILE
1	F	146	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 75 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	454	GLN
1	F	291	ASN
1	E	18	GLN
1	E	415	HIS
1	B	415	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

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	518/522 (99%)	-0.24	24 (4%)	32	31	8, 17, 57, 83	0
1	B	518/522 (99%)	-0.22	30 (5%)	23	22	7, 19, 61, 82	0
1	C	518/522 (99%)	-0.03	43 (8%)	11	10	9, 19, 68, 81	0
1	D	518/522 (99%)	-0.22	23 (4%)	34	32	10, 19, 60, 83	0
1	E	518/522 (99%)	-0.18	30 (5%)	23	22	9, 17, 57, 81	0
1	F	518/522 (99%)	-0.20	23 (4%)	34	32	9, 19, 61, 82	0
All	All	3108/3132 (99%)	-0.18	173 (5%)	24	23	7, 19, 61, 83	0

The worst 5 of 173 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	276	THR	11.0
1	C	54	GLY	10.8
1	C	215	VAL	9.5
1	E	456	ALA	9.1
1	D	461	ASP	7.7

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.