



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 06:24 AM GMT

PDB ID : 2WWT
Title : Intracellular subtilisin precursor from *B. clausii*
Authors : Vevodova, J.; Gamble, M.; Ariza, A.; Dodson, E.; Jones, D.D.; Wilson, K.S.
Deposited on : 2009-10-27
Resolution : 2.68 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
<http://wwpdb.org/validation/2016/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 : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

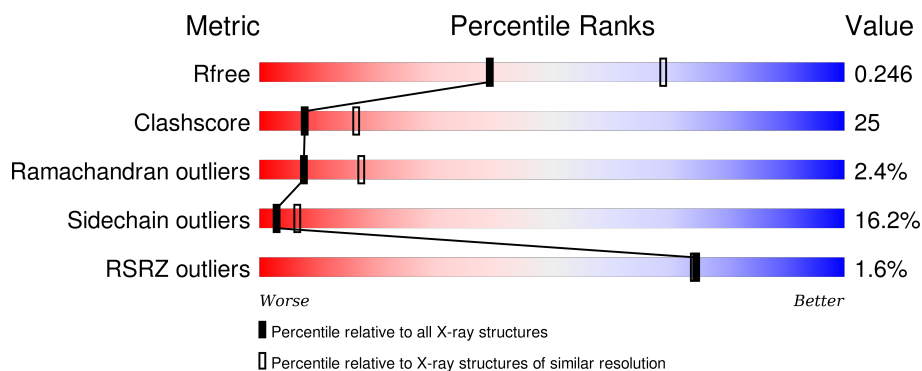
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION


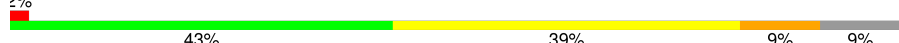



The reported resolution of this entry is 2.68 Å.

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}	91344	2780 (2.70-2.66)
Clashscore	102246	3138 (2.70-2.66)
Ramachandran outliers	100387	3089 (2.70-2.66)
Sidechain outliers	100360	3089 (2.70-2.66)
RSRZ outliers	91569	2789 (2.70-2.66)

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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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	329	
1	B	329	
1	C	329	
1	D	329	
1	E	329	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	329	 A horizontal bar chart showing the quality of chain F. The bar is divided into five segments: a small red segment at the beginning labeled '2%', followed by a large green segment labeled '53%', a yellow segment labeled '34%', a small orange segment labeled '5%', and a small grey segment at the end labeled '8%'.

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 13310 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 INTRACELLULAR SUBTILISIN PROTEASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	299	Total	C	N	O	S	0	0	0
			2200	1387	373	433	7			
1	B	298	Total	C	N	O	S	0	0	0
			2194	1384	372	431	7			
1	C	300	Total	C	N	O	S	0	0	0
			2209	1392	374	436	7			
1	D	306	Total	C	N	O	S	0	0	0
			2258	1421	382	448	7			
1	E	299	Total	C	N	O	S	6	0	0
			2200	1387	373	433	7			
1	F	302	Total	C	N	O	S	0	0	0
			2224	1401	376	440	7			

There are 54 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	322	LEU	-	EXPRESSION TAG	UNP D0AB41
A	323	GLU	-	EXPRESSION TAG	UNP D0AB41
A	324	HIS	-	EXPRESSION TAG	UNP D0AB41
A	325	HIS	-	EXPRESSION TAG	UNP D0AB41
A	326	HIS	-	EXPRESSION TAG	UNP D0AB41
A	327	HIS	-	EXPRESSION TAG	UNP D0AB41
A	328	HIS	-	EXPRESSION TAG	UNP D0AB41
A	329	HIS	-	EXPRESSION TAG	UNP D0AB41
A	250	ALA	SER	ENGINEERED MUTATION	UNP D0AB41
B	322	LEU	-	EXPRESSION TAG	UNP D0AB41
B	323	GLU	-	EXPRESSION TAG	UNP D0AB41
B	324	HIS	-	EXPRESSION TAG	UNP D0AB41
B	325	HIS	-	EXPRESSION TAG	UNP D0AB41
B	326	HIS	-	EXPRESSION TAG	UNP D0AB41
B	327	HIS	-	EXPRESSION TAG	UNP D0AB41
B	328	HIS	-	EXPRESSION TAG	UNP D0AB41
B	329	HIS	-	EXPRESSION TAG	UNP D0AB41

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	250	ALA	SER	ENGINEERED MUTATION	UNP D0AB41
C	322	LEU	-	EXPRESSION TAG	UNP D0AB41
C	323	GLU	-	EXPRESSION TAG	UNP D0AB41
C	324	HIS	-	EXPRESSION TAG	UNP D0AB41
C	325	HIS	-	EXPRESSION TAG	UNP D0AB41
C	326	HIS	-	EXPRESSION TAG	UNP D0AB41
C	327	HIS	-	EXPRESSION TAG	UNP D0AB41
C	328	HIS	-	EXPRESSION TAG	UNP D0AB41
C	329	HIS	-	EXPRESSION TAG	UNP D0AB41
C	250	ALA	SER	ENGINEERED MUTATION	UNP D0AB41
D	322	LEU	-	EXPRESSION TAG	UNP D0AB41
D	323	GLU	-	EXPRESSION TAG	UNP D0AB41
D	324	HIS	-	EXPRESSION TAG	UNP D0AB41
D	325	HIS	-	EXPRESSION TAG	UNP D0AB41
D	326	HIS	-	EXPRESSION TAG	UNP D0AB41
D	327	HIS	-	EXPRESSION TAG	UNP D0AB41
D	328	HIS	-	EXPRESSION TAG	UNP D0AB41
D	329	HIS	-	EXPRESSION TAG	UNP D0AB41
D	250	ALA	SER	ENGINEERED MUTATION	UNP D0AB41
E	322	LEU	-	EXPRESSION TAG	UNP D0AB41
E	323	GLU	-	EXPRESSION TAG	UNP D0AB41
E	324	HIS	-	EXPRESSION TAG	UNP D0AB41
E	325	HIS	-	EXPRESSION TAG	UNP D0AB41
E	326	HIS	-	EXPRESSION TAG	UNP D0AB41
E	327	HIS	-	EXPRESSION TAG	UNP D0AB41
E	328	HIS	-	EXPRESSION TAG	UNP D0AB41
E	329	HIS	-	EXPRESSION TAG	UNP D0AB41
E	250	ALA	SER	ENGINEERED MUTATION	UNP D0AB41
F	322	LEU	-	EXPRESSION TAG	UNP D0AB41
F	323	GLU	-	EXPRESSION TAG	UNP D0AB41
F	324	HIS	-	EXPRESSION TAG	UNP D0AB41
F	325	HIS	-	EXPRESSION TAG	UNP D0AB41
F	326	HIS	-	EXPRESSION TAG	UNP D0AB41
F	327	HIS	-	EXPRESSION TAG	UNP D0AB41
F	328	HIS	-	EXPRESSION TAG	UNP D0AB41
F	329	HIS	-	EXPRESSION TAG	UNP D0AB41
F	250	ALA	SER	ENGINEERED MUTATION	UNP D0AB41

- Molecule 2 is SODIUM ION (three-letter code: NA) (formula: Na).

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
-----	-------	----------	-------	---------	---------

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	1	Total Na 1 1	0	0
2	D	1	Total Na 1 1	0	0
2	C	1	Total Na 1 1	0	0
2	F	1	Total Na 1 1	0	0
2	E	1	Total Na 1 1	0	0

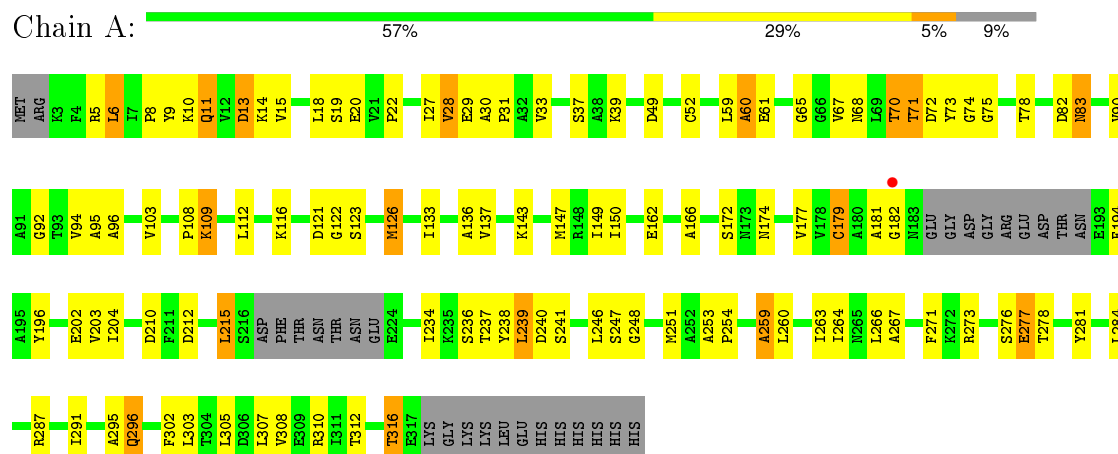
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	4	Total O 4 4	0	0
3	B	3	Total O 3 3	0	0
3	C	4	Total O 4 4	0	0
3	D	3	Total O 3 3	0	0
3	E	3	Total O 3 3	0	0
3	F	3	Total O 3 3	0	0

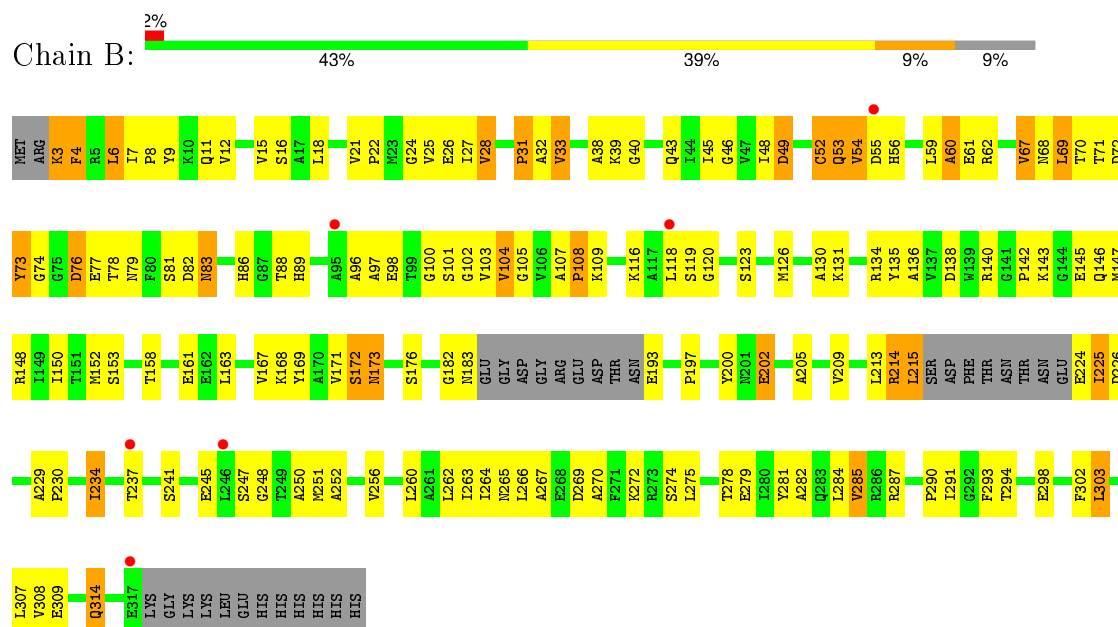
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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: INTRACELLULAR SUBTILISIN PROTEASE

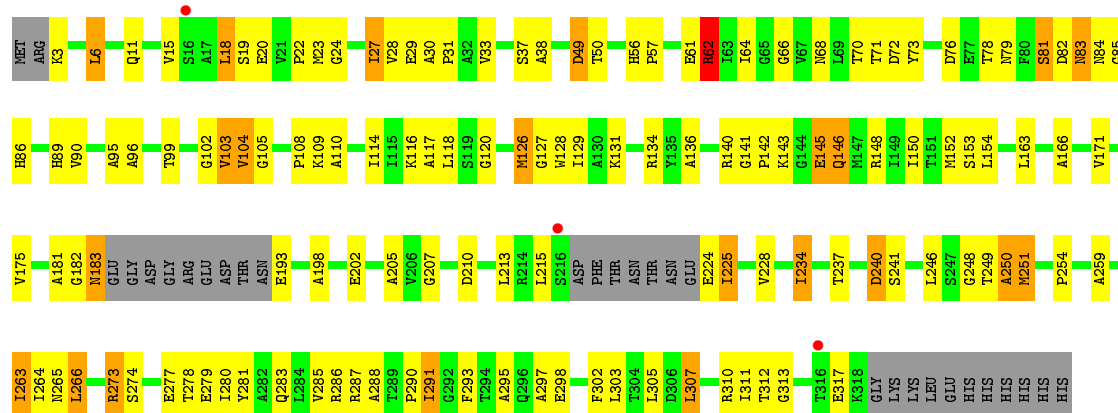


• Molecule 1: INTRACELLULAR SUBTILISIN PROTEASE

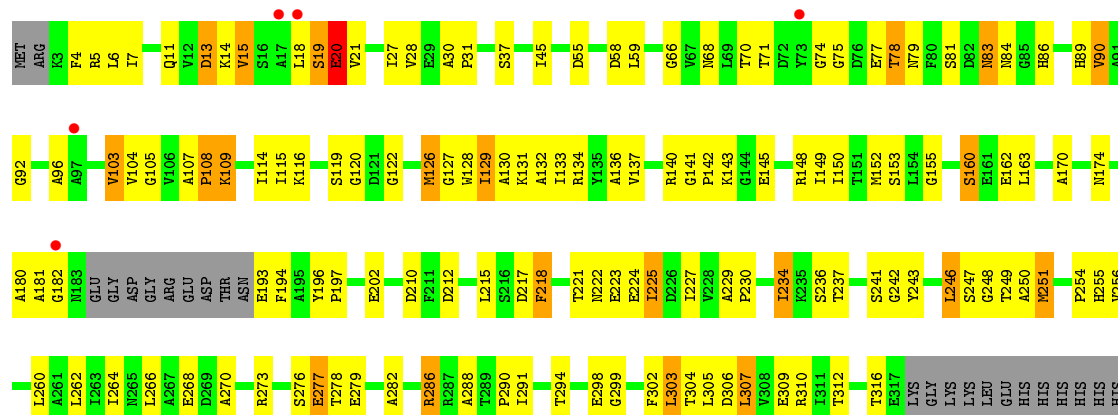


• Molecule 1: INTRACELLULAR SUBTILISIN PROTEASE

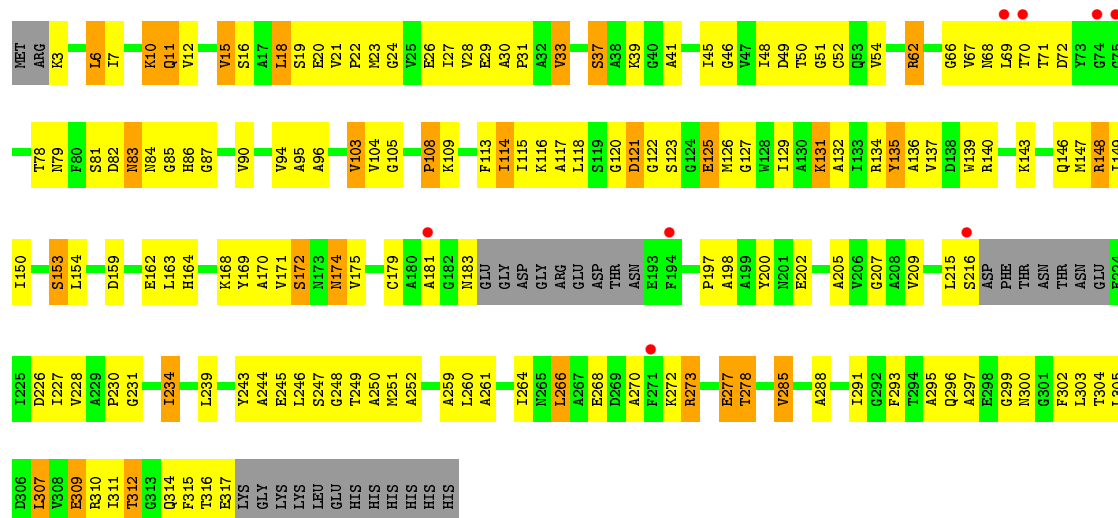




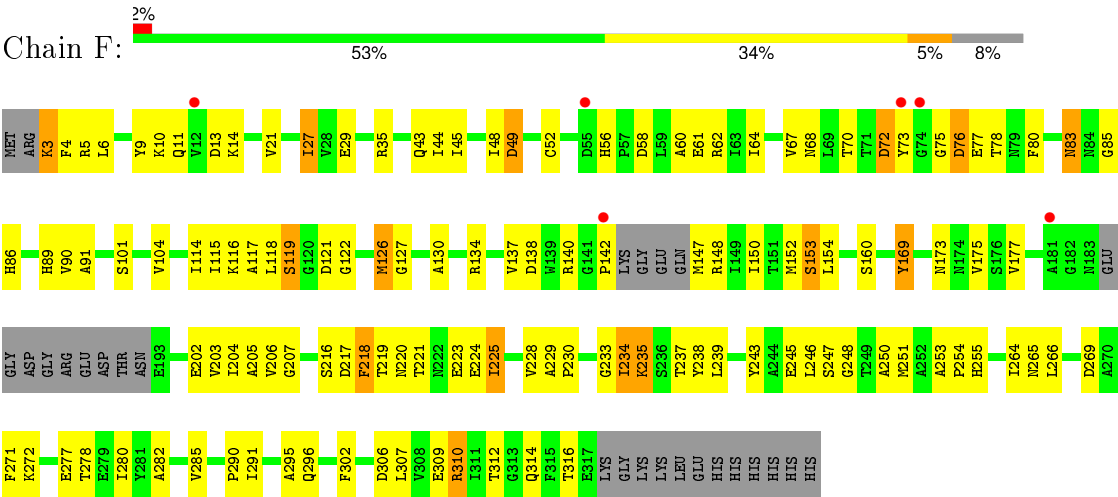
• Molecule 1: INTRACELLULAR SUBTILISIN PROTEASE



• Molecule 1: INTRACELLULAR SUBTILISIN PROTEASE



• Molecule 1: INTRACELLULAR SUBTILISIN PROTEASE



4 Data and refinement statistics

Property	Value	Source
Space group	P 31	Depositor
Cell constants a, b, c, α , β , γ	121.64Å 121.64Å 106.20Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	105.57 – 2.68 105.34 – 2.68	Depositor EDS
% Data completeness (in resolution range)	98.8 (105.57-2.68) 99.0 (105.34-2.68)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.06 (at 2.69Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.166 , 0.252 0.166 , 0.246	Depositor DCC
R_{free} test set	2485 reflections (5.36%)	DCC
Wilson B-factor (Å ²)	41.8	Xtriage
Anisotropy	0.368	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 14.9	EDS
Estimated twinning fraction	0.720 for H, K, L 0.280 for -H, H+K, -L 0.094 for -h,-k,l 0.097 for h,-h-k,-l 0.245 for -k,-h,-l	Xtriage
Reported twinning fraction	0.720 for H, K, L 0.280 for -H, H+K, -L	Depositor
L-test for twinning ²	$\langle L \rangle = 0.44$, $\langle L^2 \rangle = 0.26$	Xtriage
Outliers	0 of 48934 reflections	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	13310	wwPDB-VP
Average B, all atoms (Å ²)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.23% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NA

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	0.95	2/2237 (0.1%)	0.98	5/3039 (0.2%)
1	B	0.75	0/2231	0.88	1/3031 (0.0%)
1	C	0.93	1/2246 (0.0%)	0.99	2/3051 (0.1%)
1	D	0.83	0/2297	0.90	0/3123
1	E	0.80	2/2237 (0.1%)	0.87	1/3039 (0.0%)
1	F	0.74	0/2261	0.84	0/3073
All	All	0.84	5/13509 (0.0%)	0.91	9/18356 (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	2
1	C	0	1
All	All	0	3

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	140	ARG	CB-CG	-6.47	1.35	1.52
1	A	179	CYS	CB-SG	-6.05	1.72	1.82
1	A	259	ALA	CA-CB	-5.67	1.40	1.52
1	C	193	GLU	CB-CG	5.53	1.62	1.52
1	E	309	GLU	CB-CG	5.05	1.61	1.52

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	62	ARG	NE-CZ-NH1	-6.70	116.95	120.30
1	C	266	LEU	CB-CG-CD1	-5.96	100.86	111.00
1	E	140	ARG	CA-CB-CG	5.90	126.38	113.40
1	A	196	TYR	C-N-CD	-5.85	107.72	120.60
1	A	112	LEU	CA-CB-CG	5.76	128.55	115.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	13	ASP	Peptide
1	A	181	ALA	Peptide
1	C	313	GLY	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	2200	0	2166	78	0
1	B	2194	0	2160	146	0
1	C	2209	0	2172	101	0
1	D	2258	0	2211	133	0
1	E	2200	0	2166	120	0
1	F	2224	0	2175	113	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
2	E	1	0	0	0	0
2	F	1	0	0	0	0
3	A	4	0	0	0	0
3	B	3	0	0	0	0
3	C	4	0	0	1	0
3	D	3	0	0	0	0
3	E	3	0	0	0	0
3	F	3	0	0	0	0
All	All	13310	0	13050	664	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 25.

The worst 5 of 664 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:234:ILE:HD11	1:A:251:MET:CB	1.55	1.36
1:A:234:ILE:CD1	1:A:251:MET:HB3	1.55	1.34
1:F:234:ILE:HD11	1:F:251:MET:HB3	1.20	1.14
1:C:285:VAL:HG21	1:D:282:ALA:HB1	1.31	1.08
1:C:136:ALA:HB3	1:C:150:ILE:HD11	1.31	1.08

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	293/329 (89%)	266 (91%)	23 (8%)	4 (1%)	14	32
1	B	292/329 (89%)	253 (87%)	30 (10%)	9 (3%)	5	11
1	C	294/329 (89%)	261 (89%)	25 (8%)	8 (3%)	6	14
1	D	302/329 (92%)	259 (86%)	35 (12%)	8 (3%)	7	15
1	E	293/329 (89%)	261 (89%)	22 (8%)	10 (3%)	5	10
1	F	296/329 (90%)	263 (89%)	30 (10%)	3 (1%)	19	42
All	All	1770/1974 (90%)	1563 (88%)	165 (9%)	42 (2%)	7	17

5 of 42 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	182	GLY
1	A	239	LEU
1	B	32	ALA
1	B	60	ALA
1	B	182	GLY

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	224/252 (89%)	197 (88%)	27 (12%)	6	13
1	B	223/252 (88%)	184 (82%)	39 (18%)	2	5
1	C	225/252 (89%)	185 (82%)	40 (18%)	2	5
1	D	231/252 (92%)	197 (85%)	34 (15%)	4	9
1	E	224/252 (89%)	180 (80%)	44 (20%)	1	4
1	F	228/252 (90%)	193 (85%)	35 (15%)	3	7
All	All	1355/1512 (90%)	1136 (84%)	219 (16%)	3	6

5 of 219 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	266	LEU
1	D	140	ARG
1	F	153	SER
1	C	277	GLU
1	D	55	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 26 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	83	ASN
1	D	314	GLN
1	F	255	HIS
1	D	164	HIS
1	D	201	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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 5 ligands modelled in this entry, 5 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	299/329 (90%)	-0.14	1 (0%) 94 95	19, 29, 44, 59	0
1	B	298/329 (90%)	0.20	6 (2%) 68 68	22, 42, 67, 89	0
1	C	300/329 (91%)	-0.08	3 (1%) 84 84	19, 28, 44, 77	0
1	D	306/329 (93%)	0.10	5 (1%) 74 74	20, 37, 56, 69	0
1	E	299/329 (90%)	0.24	8 (2%) 58 57	25, 42, 63, 70	1 (0%)
1	F	302/329 (91%)	0.28	6 (1%) 68 68	28, 48, 74, 98	0
All	All	1804/1974 (91%)	0.10	29 (1%) 74 74	19, 37, 65, 98	1 (0%)

The worst 5 of 29 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	18	LEU	3.4
1	F	73	TYR	3.3
1	D	17	ALA	3.3
1	E	69	LEU	3.3
1	E	74	GLY	3.2

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 carbohydrates in this entry.

6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å ²)	Q<0.9
2	NA	E	401	1/1	0.82	0.16	-1.37	33,33,33,33	0
2	NA	C	401	1/1	0.98	0.12	-1.58	22,22,22,22	0
2	NA	F	401	1/1	0.98	0.10	-2.26	44,44,44,44	0
2	NA	D	401	1/1	0.92	0.07	-2.27	37,37,37,37	0
2	NA	B	401	1/1	0.88	0.08	-2.29	37,37,37,37	0

6.5 Other polymers [i](#)

There are no such residues in this entry.