Full wwPDB X-ray Structure Validation Report ⓘ

May 13, 2020 – 12:13 am BST

PDB ID	:	5NSK
Title	:	apo Structure of Leucyl aminopeptidase from Trypanosoma brucei
Authors	:	Timm, J.; Wilson, K.S.
Deposited on	:	2017-04-26
Resolution	:	2.60 Å(reported)

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

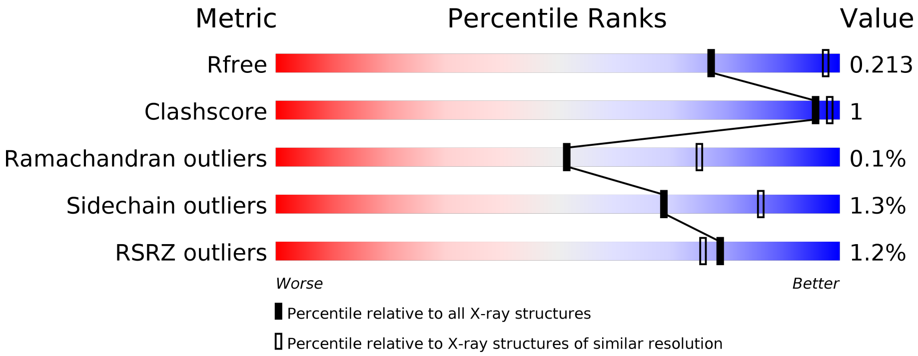
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

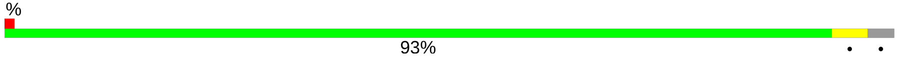
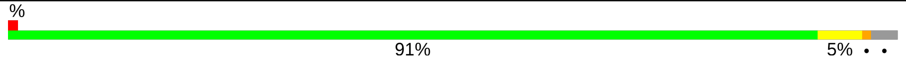
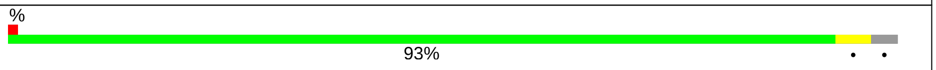
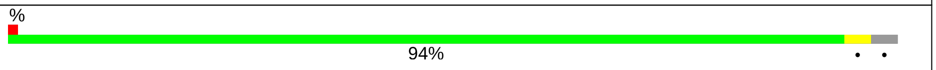
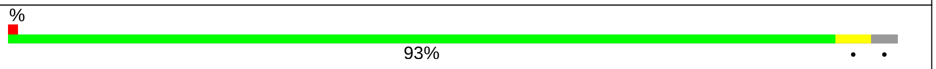
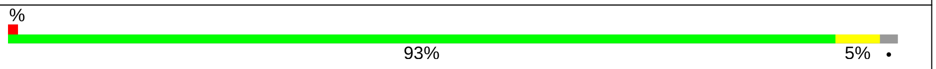
The reported resolution of this entry is 2.60 Å.

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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (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 on 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	521	
1	B	521	
1	C	521	
1	D	521	
1	E	521	
1	F	521	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 22652 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 Aminopeptidase, putative.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	507	Total	C	N	O	S	0	0	0
			3683	2329	627	707	20			
1	B	506	Total	C	N	O	S	0	0	0
			3683	2327	628	709	19			
1	C	506	Total	C	N	O	S	0	0	0
			3700	2335	632	714	19			
1	D	506	Total	C	N	O	S	0	0	0
			3675	2320	627	709	19			
1	E	506	Total	C	N	O	S	0	0	0
			3704	2342	633	709	20			
1	F	508	Total	C	N	O	S	0	0	0
			3707	2342	632	713	20			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	33	ALA	THR	conflict	UNP Q385B0
A	140	THR	ALA	conflict	UNP Q385B0
B	33	ALA	THR	conflict	UNP Q385B0
B	140	THR	ALA	conflict	UNP Q385B0
C	33	ALA	THR	conflict	UNP Q385B0
C	140	THR	ALA	conflict	UNP Q385B0
D	33	ALA	THR	conflict	UNP Q385B0
D	140	THR	ALA	conflict	UNP Q385B0
E	33	ALA	THR	conflict	UNP Q385B0
E	140	THR	ALA	conflict	UNP Q385B0
F	33	ALA	THR	conflict	UNP Q385B0
F	140	THR	ALA	conflict	UNP Q385B0

- Molecule 2 is water.

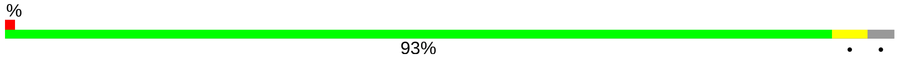
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	83	Total O 83 83	0	0
2	B	62	Total O 62 62	0	0
2	C	119	Total O 119 119	0	0
2	D	71	Total O 71 71	0	0
2	E	87	Total O 87 87	0	0
2	F	78	Total O 78 78	0	0

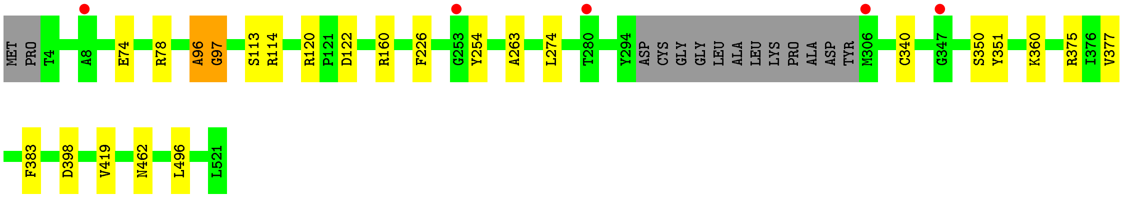
3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA 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: Aminopeptidase, putative

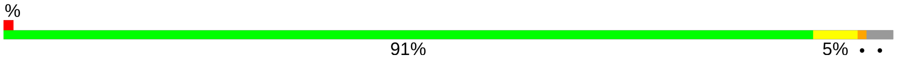
Chain A:

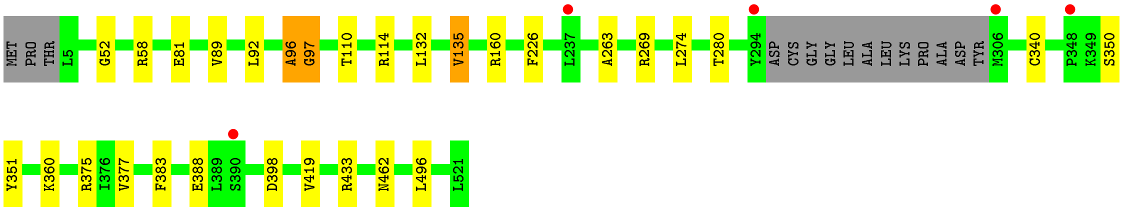




- Molecule 1: Aminopeptidase, putative

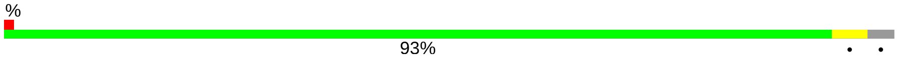
Chain B:

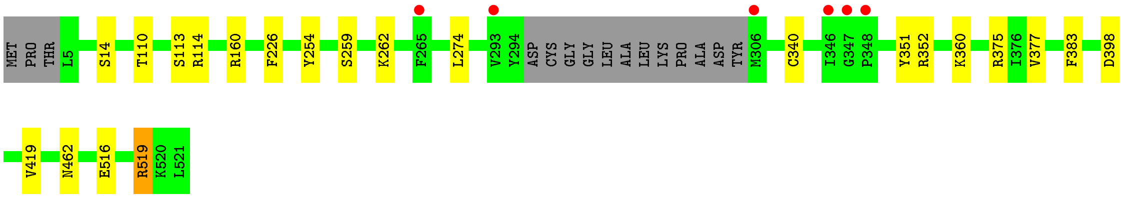




- Molecule 1: Aminopeptidase, putative

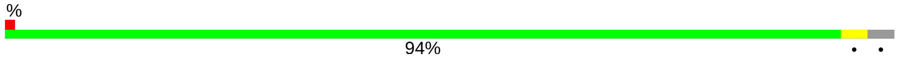
Chain C:

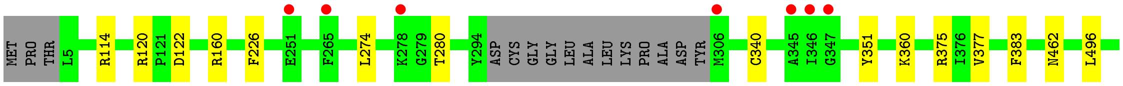




- Molecule 1: Aminopeptidase, putative

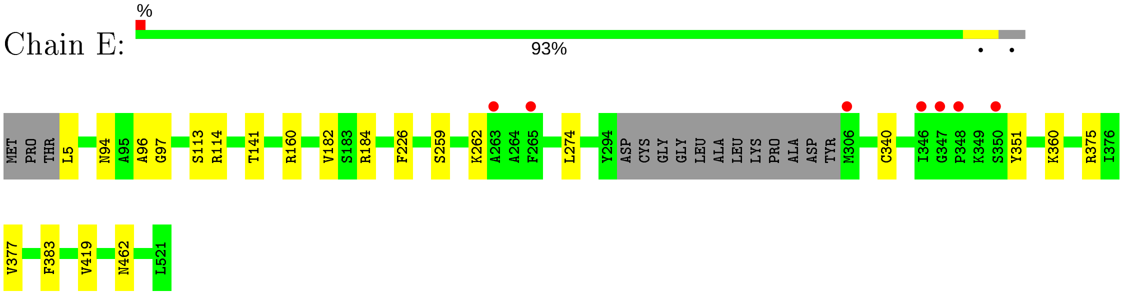
Chain D:





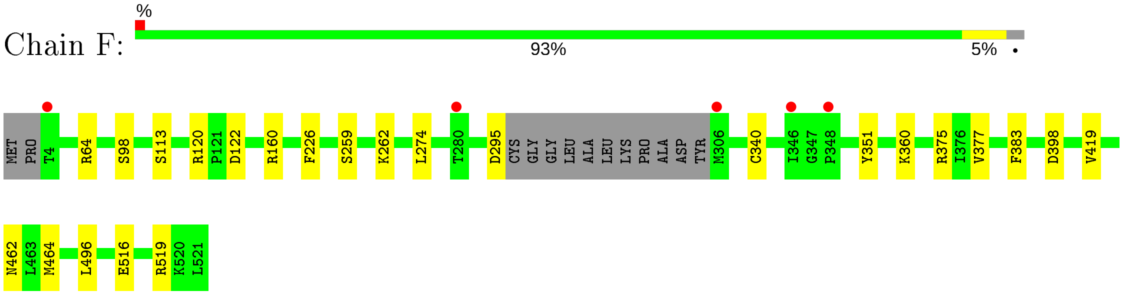
- Molecule 1: Aminopeptidase, putative

Chain E:



- Molecule 1: Aminopeptidase, putative

Chain F:



4 Data and refinement statistics ⓘ

Property	**Value**	**Source**
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	162.33Å 162.45Å 176.99Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	72.61 – 2.60 72.61 – 2.60	Depositor EDS
% Data completeness (in resolution range)	100.0 (72.61-2.60) 100.0 (72.61-2.60)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.11 (at 2.62Å)	Xtriage
Refinement program	REFMAC 5.8.0135	Depositor
R, R_{free}	0.191 , 0.211 0.196 , 0.213	Depositor DCC
R_{free} test set	7257 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	38.1	Xtriage
Anisotropy	0.030	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 26.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.013 for k,h,-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	22652	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 1.96% 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	0.81	1/3757 (0.0%)	0.79	1/5134 (0.0%)
1	B	0.84	2/3757 (0.1%)	0.83	7/5134 (0.1%)
1	C	0.84	2/3774 (0.1%)	0.81	5/5155 (0.1%)
1	D	0.79	0/3748	0.80	3/5122 (0.1%)
1	E	0.85	1/3778 (0.0%)	0.82	2/5159 (0.0%)
1	F	0.84	1/3782 (0.0%)	0.82	5/5169 (0.1%)
All	All	0.83	7/22596 (0.0%)	0.81	23/30873 (0.1%)

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	B	0	1
1	E	0	1
All	All	0	3

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	113	SER	CB-OG	5.97	1.50	1.42
1	E	113	SER	CB-OG	5.71	1.49	1.42
1	B	388	GLU	CD-OE2	-5.57	1.19	1.25
1	A	113	SER	CB-OG	5.44	1.49	1.42
1	C	14	SER	CB-OG	-5.35	1.35	1.42
1	F	113	SER	CB-OG	5.34	1.49	1.42
1	B	81	GLU	CD-OE1	-5.11	1.20	1.25

All (23) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	519	ARG	NE-CZ-NH2	8.25	124.43	120.30
1	C	519	ARG	NE-CZ-NH1	-7.25	116.68	120.30
1	D	160	ARG	NE-CZ-NH1	6.91	123.75	120.30
1	F	160	ARG	NE-CZ-NH1	6.44	123.52	120.30
1	B	269	ARG	CG-CD-NE	-6.33	98.50	111.80
1	C	160	ARG	NE-CZ-NH1	6.04	123.32	120.30
1	F	64	ARG	NE-CZ-NH2	-5.87	117.37	120.30
1	A	160	ARG	NE-CZ-NH1	5.83	123.22	120.30
1	B	160	ARG	NE-CZ-NH1	5.79	123.19	120.30
1	C	14	SER	CB-CA-C	5.74	121.01	110.10
1	B	433	ARG	NE-CZ-NH2	-5.73	117.43	120.30
1	B	58	ARG	NE-CZ-NH2	-5.57	117.52	120.30
1	F	64	ARG	NE-CZ-NH1	5.47	123.04	120.30
1	F	160	ARG	CB-CA-C	-5.46	99.48	110.40
1	B	58	ARG	CG-CD-NE	-5.46	100.33	111.80
1	B	160	ARG	CB-CA-C	-5.42	99.56	110.40
1	B	433	ARG	NE-CZ-NH1	5.29	122.95	120.30
1	D	160	ARG	CB-CA-C	-5.22	99.96	110.40
1	D	280	THR	N-CA-CB	5.21	120.20	110.30
1	E	160	ARG	NE-CZ-NH1	5.21	122.91	120.30
1	F	98	SER	N-CA-CB	-5.19	102.72	110.50
1	C	160	ARG	CB-CA-C	-5.12	100.17	110.40
1	E	160	ARG	CB-CA-C	-5.04	100.32	110.40

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	96	ALA	Peptide
1	B	96	ALA	Peptide
1	E	97	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	3683	0	3559	11	0
1	B	3683	0	3563	11	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	3700	0	3586	9	0
1	D	3675	0	3539	5	0
1	E	3704	0	3611	9	0
1	F	3707	0	3583	9	0
2	A	83	0	0	0	0
2	B	62	0	0	1	0
2	C	119	0	0	0	0
2	D	71	0	0	0	0
2	E	87	0	0	1	0
2	F	78	0	0	0	0
All	All	22652	0	21441	48	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 1.

All (48) 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:F:516:GLU:OE1	1:F:519:ARG:NH1	2.26	0.68
1:B:114:ARG:NH2	1:C:419:VAL:O	2.28	0.66
1:A:340:CYS:HB3	1:A:377:VAL:HG13	1.80	0.64
1:B:419:VAL:O	1:C:114:ARG:NH2	2.30	0.64
1:C:352:ARG:HG2	1:F:295:ASP:HA	1.82	0.61
1:F:340:CYS:HB3	1:F:377:VAL:HG13	1.82	0.61
1:E:340:CYS:HB3	1:E:377:VAL:HG13	1.83	0.61
1:B:340:CYS:HB3	1:B:377:VAL:HG13	1.81	0.60
1:A:114:ARG:NH2	1:E:419:VAL:O	2.33	0.60
1:D:340:CYS:HB3	1:D:377:VAL:HG13	1.84	0.60
1:A:419:VAL:O	1:E:114:ARG:NH2	2.35	0.59
1:D:114:ARG:NH2	1:F:419:VAL:O	2.35	0.59
1:B:89:VAL:HG11	1:B:92:LEU:HD22	1.84	0.59
1:C:340:CYS:HB3	1:C:377:VAL:HG13	1.85	0.57
1:C:516:GLU:OE1	1:C:519:ARG:NH2	2.36	0.57
1:C:360:LYS:HG3	1:C:383:PHE:CD2	2.45	0.51
1:B:360:LYS:HG3	1:B:383:PHE:CD2	2.46	0.51
1:A:360:LYS:HG3	1:A:383:PHE:CD2	2.48	0.49
1:E:360:LYS:HG3	1:E:383:PHE:CD2	2.48	0.48
1:B:132:LEU:O	1:B:135:VAL:HG22	2.14	0.48
1:D:360:LYS:HG3	1:D:383:PHE:CD2	2.48	0.47
1:B:96:ALA:O	1:B:97:GLY:O	2.31	0.47
1:B:96:ALA:O	1:B:97:GLY:C	2.52	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:F:360:LYS:HG3	1:F:383:PHE:CD2	2.50	0.47
1:A:120:ARG:NH1	1:A:122:ASP:OD2	2.48	0.46
1:A:96:ALA:O	1:A:97:GLY:O	2.35	0.46
1:A:96:ALA:O	1:A:97:GLY:C	2.55	0.44
1:E:141:THR:OG1	1:E:184:ARG:NH1	2.52	0.43
1:D:120:ARG:NH1	1:D:122:ASP:OD2	2.51	0.43
1:A:74:GLU:HG3	1:A:78:ARG:HD2	2.01	0.43
1:E:182:VAL:HG11	2:E:679:HOH:O	2.19	0.43
1:B:375:ARG:HD2	1:B:462:ASN:O	2.18	0.43
1:F:120:ARG:NH1	1:F:122:ASP:OD2	2.51	0.42
1:F:375:ARG:HD2	1:F:462:ASN:O	2.19	0.42
1:E:375:ARG:HD2	1:E:462:ASN:O	2.19	0.41
1:C:254:TYR:CE1	1:C:360:LYS:HG2	2.55	0.41
1:B:52:GLY:HA2	2:B:635:HOH:O	2.20	0.41
1:D:375:ARG:HD2	1:D:462:ASN:O	2.21	0.41
1:F:259:SER:HA	1:F:262:LYS:HE2	2.03	0.41
1:C:259:SER:HA	1:C:262:LYS:HE2	2.03	0.41
1:F:464:MET:HG2	1:F:464:MET:H	1.69	0.41
1:A:254:TYR:CE1	1:A:360:LYS:HG2	2.56	0.41
1:E:259:SER:HA	1:E:262:LYS:HE2	2.02	0.41
1:C:375:ARG:HD2	1:C:462:ASN:O	2.22	0.40
1:E:94:ASN:O	1:E:96:ALA:O	2.40	0.40
1:A:263:ALA:CB	1:A:350:SER:HB3	2.52	0.40
1:A:375:ARG:HD2	1:A:462:ASN:O	2.21	0.40
1:B:263:ALA:CB	1:B:350:SER:HB3	2.52	0.40

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	503/521 (96%)	492 (98%)	10 (2%)	1 (0%)	47	71

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	502/521 (96%)	488 (97%)	12 (2%)	2 (0%)	34	57
1	C	502/521 (96%)	491 (98%)	11 (2%)	0	100	100
1	D	502/521 (96%)	490 (98%)	12 (2%)	0	100	100
1	E	502/521 (96%)	491 (98%)	11 (2%)	0	100	100
1	F	504/521 (97%)	493 (98%)	11 (2%)	0	100	100
All	All	3015/3126 (96%)	2945 (98%)	67 (2%)	3 (0%)	51	75

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	97	GLY
1	A	97	GLY
1	B	280	THR

5.3.2 Protein sidechains ⓘ

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	374/411 (91%)	369 (99%)	5 (1%)	69	86
1	B	376/411 (92%)	369 (98%)	7 (2%)	57	79
1	C	380/411 (92%)	375 (99%)	5 (1%)	69	86
1	D	373/411 (91%)	369 (99%)	4 (1%)	73	88
1	E	381/411 (93%)	377 (99%)	4 (1%)	76	90
1	F	379/411 (92%)	374 (99%)	5 (1%)	69	86
All	All	2263/2466 (92%)	2233 (99%)	30 (1%)	69	86

All (30) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	226	PHE
1	A	274	LEU
1	A	351	TYR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	398	ASP
1	A	496	LEU
1	B	110	THR
1	B	135	VAL
1	B	226	PHE
1	B	274	LEU
1	B	351	TYR
1	B	398	ASP
1	B	496	LEU
1	C	110	THR
1	C	226	PHE
1	C	274	LEU
1	C	351	TYR
1	C	398	ASP
1	D	226	PHE
1	D	274	LEU
1	D	351	TYR
1	D	496	LEU
1	E	5	LEU
1	E	226	PHE
1	E	274	LEU
1	E	351	TYR
1	F	226	PHE
1	F	274	LEU
1	F	351	TYR
1	F	398	ASP
1	F	496	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 carbohydrates 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(Å^2)	Q<0.9
1	A	507/521 (97%)	-0.16	5 (0%)	82	80	26, 35, 68, 115	0
1	B	506/521 (97%)	-0.05	5 (0%)	82	80	28, 40, 70, 130	0
1	C	506/521 (97%)	-0.32	6 (1%)	79	76	22, 30, 63, 121	0
1	D	506/521 (97%)	-0.20	7 (1%)	75	71	28, 37, 70, 116	0
1	E	506/521 (97%)	-0.20	7 (1%)	75	71	24, 34, 63, 133	0
1	F	508/521 (97%)	-0.17	5 (0%)	82	80	25, 36, 64, 116	0
All	All	3039/3126 (97%)	-0.18	35 (1%)	79	76	22, 36, 68, 133	0

All (35) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	346	ILE	6.1
1	A	347	GLY	6.0
1	B	306	MET	4.8
1	E	348	PRO	4.7
1	E	346	ILE	4.7
1	A	306	MET	4.5
1	E	347	GLY	4.5
1	F	306	MET	4.2
1	E	350	SER	4.1
1	F	348	PRO	3.9
1	C	265	PHE	3.7
1	D	347	GLY	3.4
1	D	306	MET	3.4
1	F	4	THR	3.4
1	C	348	PRO	3.4
1	E	306	MET	3.3
1	D	265	PHE	3.2
1	F	280	THR	3.0
1	C	306	MET	3.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	294	TYR	2.7
1	B	390	SER	2.7
1	C	293	VAL	2.6
1	C	347	GLY	2.6
1	D	278	LYS	2.6
1	F	346	ILE	2.5
1	D	346	ILE	2.4
1	E	265	PHE	2.4
1	A	8	ALA	2.4
1	D	345	ALA	2.4
1	E	263	ALA	2.3
1	A	253	GLY	2.3
1	A	280	THR	2.1
1	B	237	LEU	2.0
1	D	251	GLU	2.0
1	B	348	PRO	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 carbohydrates 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.