



wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 21, 2022 – 12:05 AM JST

PDB ID : 7DRS
Title : Structure of SspE_40224
Authors : Haiyan, G.; Jinchuan, Z.; Chen, S.; Wang, L.; Wu, G.
Deposited on : 2020-12-29
Resolution : 3.42 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 types of validation reports are described at <http://www.wwpdb.org/validation/2017/FAQs#types>.

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	:	2.29
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)	:	2.29

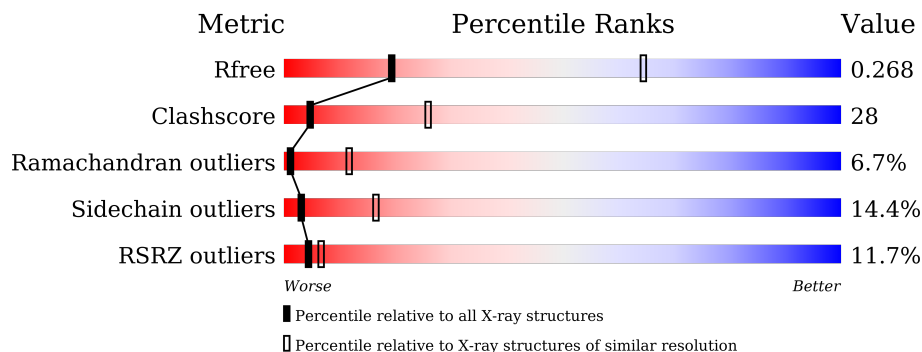
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.42 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	1486 (3.50-3.34)
Clashscore	141614	1572 (3.50-3.34)
Ramachandran outliers	138981	1534 (3.50-3.34)
Sidechain outliers	138945	1535 (3.50-3.34)
RSRZ outliers	127900	1395 (3.50-3.34)

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	771	<div> <div>7%</div> <div>46%</div> <div>43%</div> <div>10%</div> <div>.</div> </div>
1	B	771	<div> <div>23%</div> <div>48%</div> <div>39%</div> <div>11%</div> <div>.</div> </div>
1	C	771	<div> <div>7%</div> <div>45%</div> <div>41%</div> <div>12%</div> <div>.</div> </div>
1	D	771	<div> <div>9%</div> <div>45%</div> <div>42%</div> <div>11%</div> <div>.</div> </div>

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 24548 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.

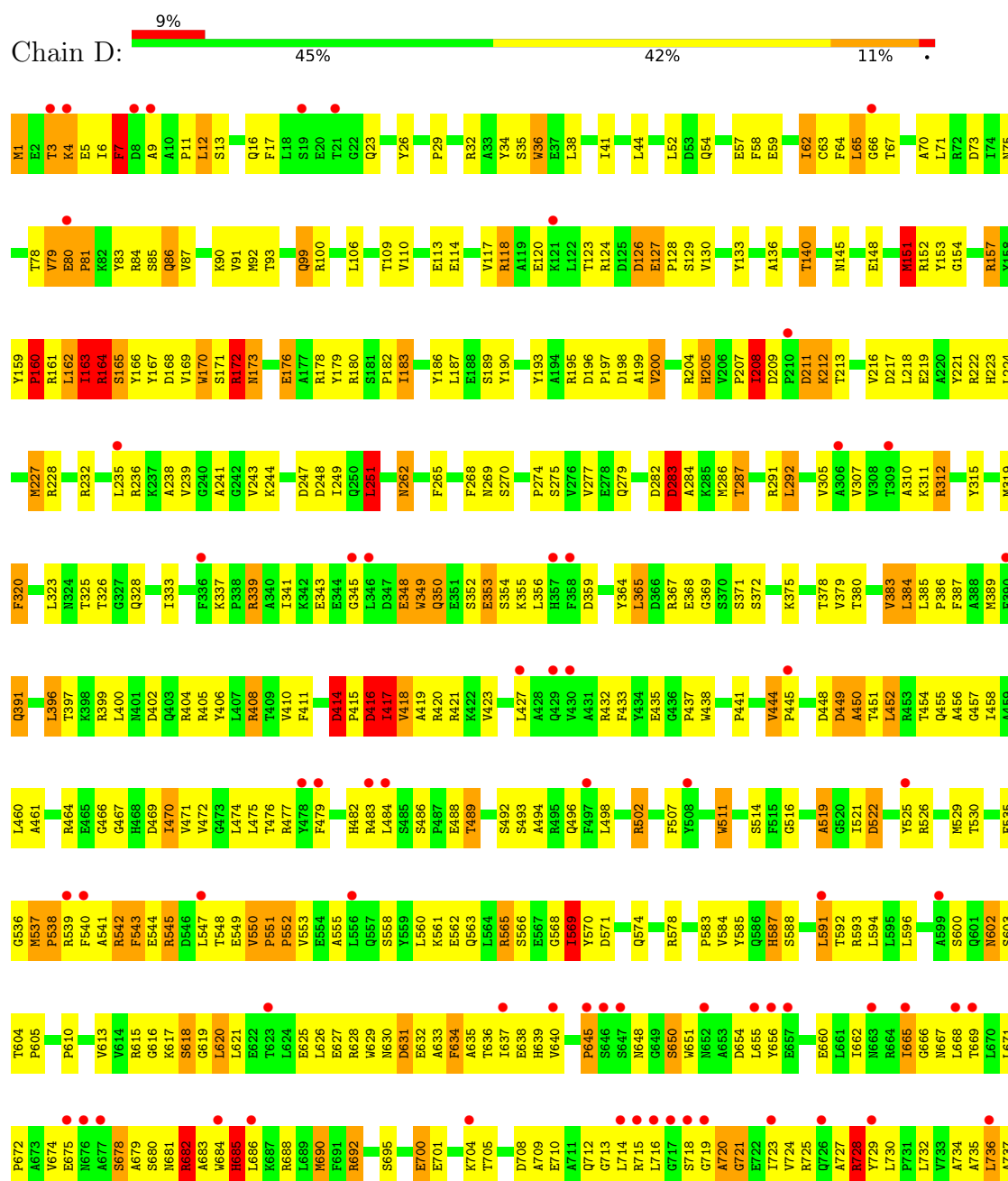
- Molecule 1 is a protein called SspE protein.

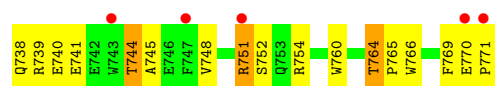
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	D	771	Total	C	N	O	S	0	0	0
			6137	3872	1079	1166	20			
1	A	771	Total	C	N	O	S	0	0	0
			6137	3872	1079	1166	20			
1	B	771	Total	C	N	O	S	0	0	0
			6137	3872	1079	1166	20			
1	C	771	Total	C	N	O	S	0	0	0
			6137	3872	1079	1166	20			

3 Residue-property plots

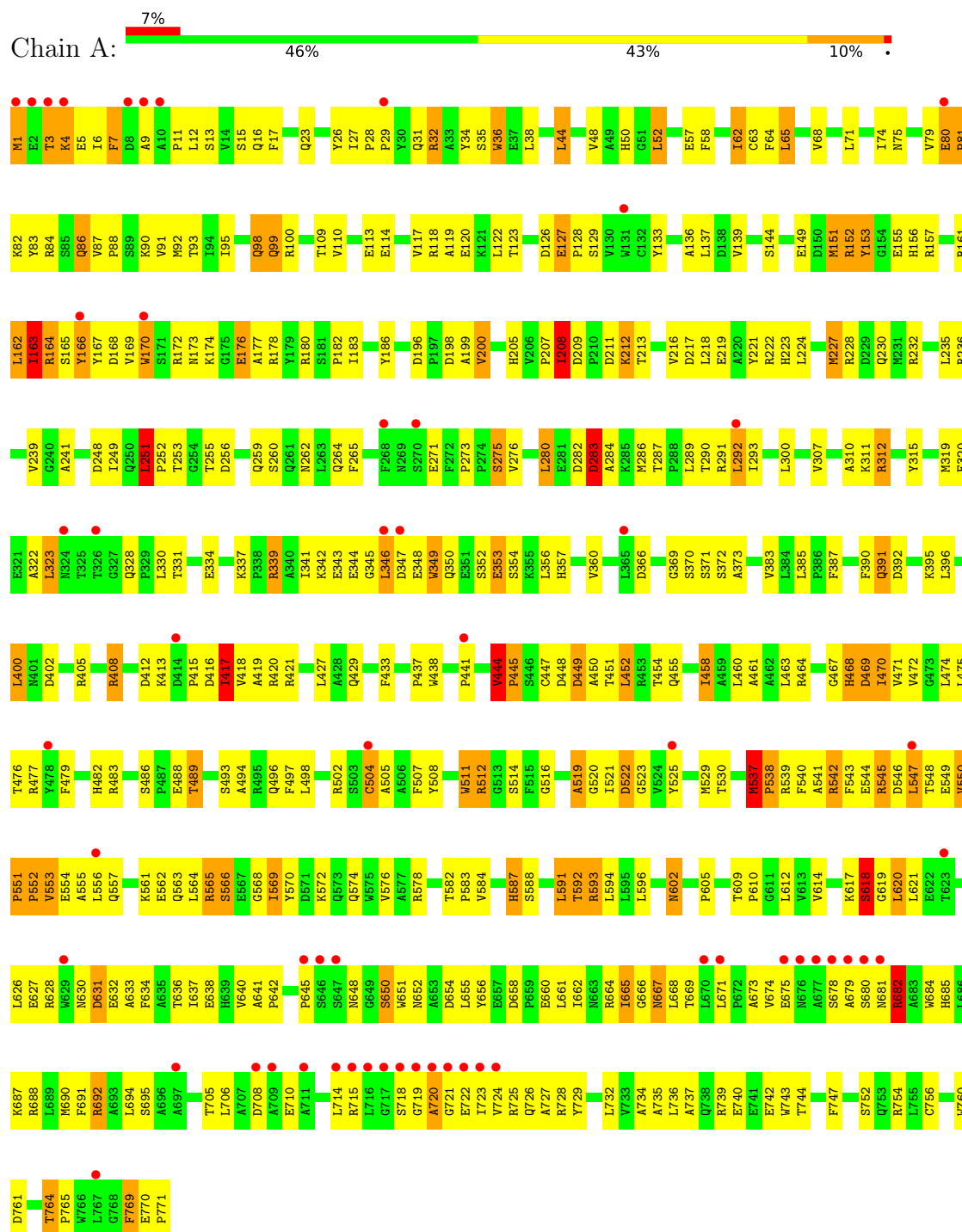
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density.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



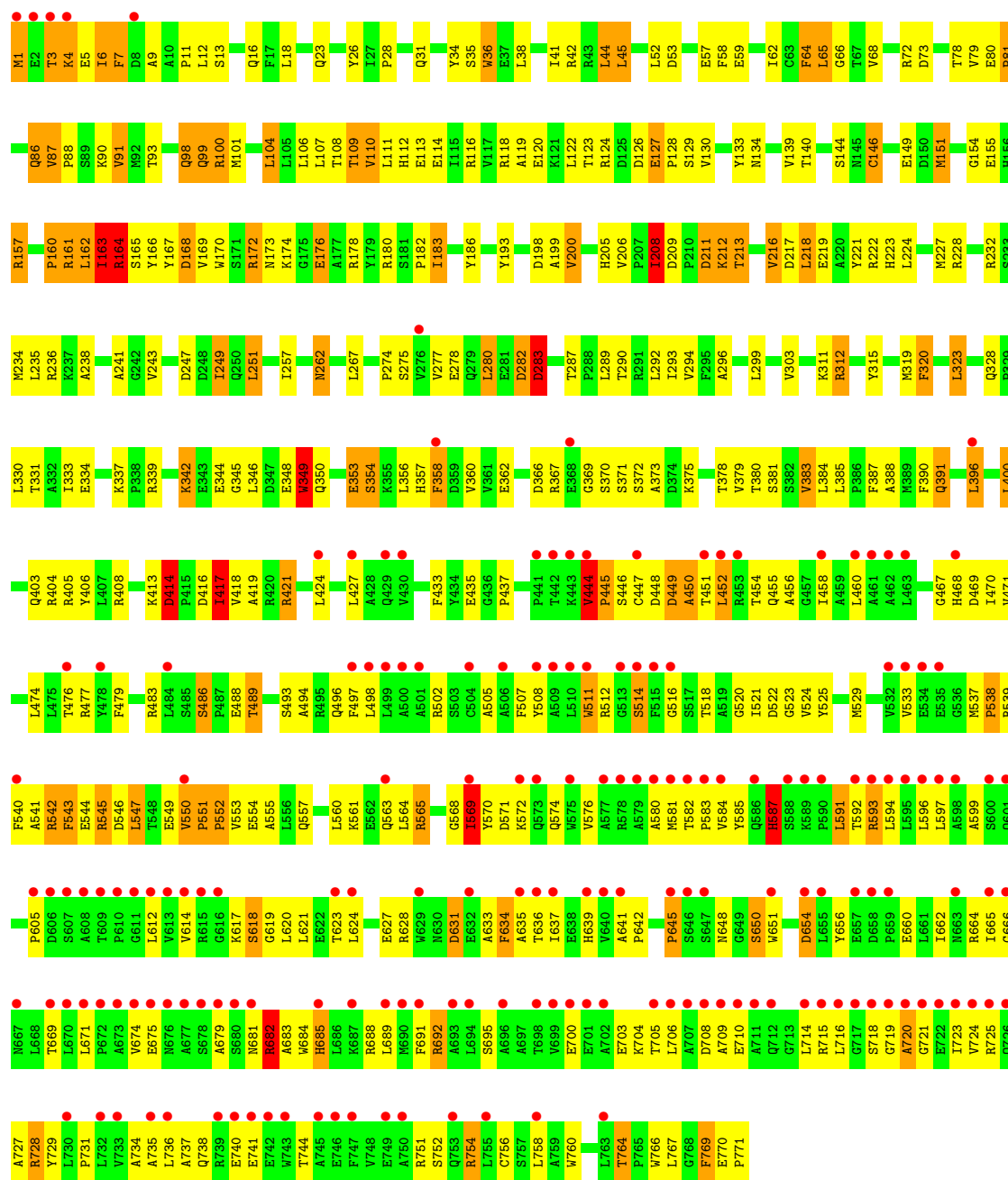


• Molecule 1: SspE protein

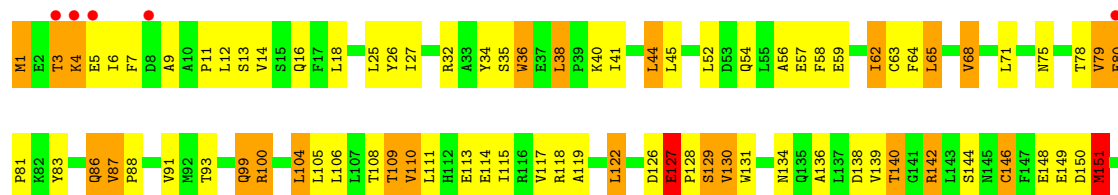
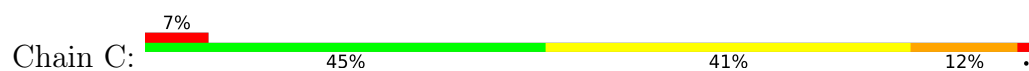


• Molecule 1: SspE protein





• Molecule 1: SspE protein





4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	111.08Å 138.15Å 293.69Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.14 – 3.42 51.95 – 3.42	Depositor EDS
% Data completeness (in resolution range)	99.7 (46.14-3.42) 99.7 (51.95-3.42)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.40 (at 3.40Å)	Xtriage
Refinement program	REFMAC 5.5	Depositor
R, R_{free}	0.212 , 0.264 0.221 , 0.268	Depositor DCC
R_{free} test set	3031 reflections (4.91%)	wwPDB-VP
Wilson B-factor (Å ²)	121.6	Xtriage
Anisotropy	0.343	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 100.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	24548	wwPDB-VP
Average B, all atoms (Å ²)	156.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.48% 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, 0.333 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	0.63	2/6272 (0.0%)	0.85	8/8515 (0.1%)
1	B	0.60	3/6272 (0.0%)	0.83	11/8515 (0.1%)
1	C	0.79	9/6272 (0.1%)	0.98	24/8515 (0.3%)
1	D	0.70	5/6272 (0.1%)	0.91	13/8515 (0.2%)
All	All	0.69	19/25088 (0.1%)	0.90	56/34060 (0.2%)

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	B	0	1
1	C	0	5
1	D	0	3
All	All	0	11

The worst 5 of 19 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	80	GLU	CB-CG	7.56	1.66	1.52
1	C	146	CYS	CB-SG	-6.49	1.71	1.82
1	C	185	TYR	CG-CD1	-6.38	1.30	1.39
1	C	63	CYS	CB-SG	-6.34	1.71	1.82
1	B	349	TRP	CB-CG	6.06	1.61	1.50

The worst 5 of 56 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	251	LEU	CA-CB-CG	10.36	139.12	115.30
1	B	346	LEU	CA-CB-CG	9.05	136.11	115.30
1	C	251	LEU	CA-CB-CG	8.93	135.85	115.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	251	LEU	CA-CB-CG	8.93	135.83	115.30
1	A	346	LEU	CA-CB-CG	8.91	135.79	115.30

There are no chirality outliers.

5 of 11 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	207	PRO	Peptide
1	A	553	VAL	Peptide
1	D	160	PRO	Peptide
1	D	173	ASN	Peptide
1	D	79	VAL	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	6137	0	6046	338	1
1	B	6137	0	6046	345	1
1	C	6137	0	6045	353	2
1	D	6137	0	6046	351	2
All	All	24548	0	24183	1365	3

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 28.

The worst 5 of 1365 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:C:628:ARG:HA	1:C:631:ASP:HB2	1.44	0.99
1:A:418:VAL:H	1:A:421:ARG:HD2	1.25	0.98
1:B:126:ASP:HB3	1:B:129:SER:HB3	1.45	0.98
1:A:126:ASP:HB3	1:A:129:SER:HB3	1.44	0.96
1:D:126:ASP:HB3	1:D:129:SER:HB3	1.47	0.95

All (3) symmetry-related close contacts are listed below. The label for Atom-2 includes the sym-

metry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:85:SER:N	1:C:211:ASP:OD2[4_457]	1.94	0.26
1:D:84:ARG:N	1:C:211:ASP:OD2[4_457]	2.13	0.07
1:A:236:ARG:NH1	1:B:278:GLU:OE1[4_447]	2.18	0.02

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	769/771 (100%)	613 (80%)	105 (14%)	51 (7%)	1	12
1	B	769/771 (100%)	605 (79%)	117 (15%)	47 (6%)	1	13
1	C	769/771 (100%)	612 (80%)	104 (14%)	53 (7%)	1	11
1	D	769/771 (100%)	609 (79%)	105 (14%)	55 (7%)	1	10
All	All	3076/3084 (100%)	2439 (79%)	431 (14%)	206 (7%)	1	12

5 of 206 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	4	LYS
1	D	36	TRP
1	D	127	GLU
1	D	200	VAL
1	D	205	HIS

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	651/651 (100%)	560 (86%)	91 (14%)	3	17
1	B	651/651 (100%)	558 (86%)	93 (14%)	3	17
1	C	651/651 (100%)	553 (85%)	98 (15%)	3	15
1	D	651/651 (100%)	558 (86%)	93 (14%)	3	17
All	All	2604/2604 (100%)	2229 (86%)	375 (14%)	3	16

5 of 375 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	323	LEU
1	C	75	ASN
1	B	358	PHE
1	B	547	LEU
1	C	164	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	401	ASN
1	C	391	GLN
1	A	667	ASN
1	C	455	GLN
1	C	223	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	771/771 (100%)	0.52	57 (7%) 14 17	84, 138, 228, 331	0
1	B	771/771 (100%)	1.33	178 (23%) 0 1	74, 180, 334, 405	0
1	C	771/771 (100%)	0.56	54 (7%) 16 19	57, 129, 243, 347	0
1	D	771/771 (100%)	0.63	71 (9%) 9 11	67, 143, 228, 341	0
All	All	3084/3084 (100%)	0.76	360 (11%) 4 7	57, 142, 295, 405	0

The worst 5 of 360 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	722	GLU	30.8
1	B	637	ILE	16.5
1	D	718	SER	15.5
1	C	696	ALA	14.8
1	B	721	GLY	13.5

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.