WORLDWIDE PDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 27, 2016 – 04:49 PM EDT

PDB ID	:	5E5N
Title	:	Ketosynthase from module 6 of the bacillaene synthase from Bacillus subtilis 168 (C167S mutant, crystal form 1)
Authors	:	Wagner, D.T.; Gay, D.C.; Keatinge-Clay, A.T.
Deposited on	:	2015-10-08
Resolution	:	2.00 Å(reported)

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

|		
MolProbity	:	4.02b-467
Mogul	:	unknown
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20027939
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)	:	rb-20027939

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 2.00 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	Percentile Ranks	Value
Rfree		0.243
Clashscore		9
Ramachandran outliers		0.1%
Sidechain outliers		1.9%
RSRZ outliers		19.1%

Worse .... *Better*

▌ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution
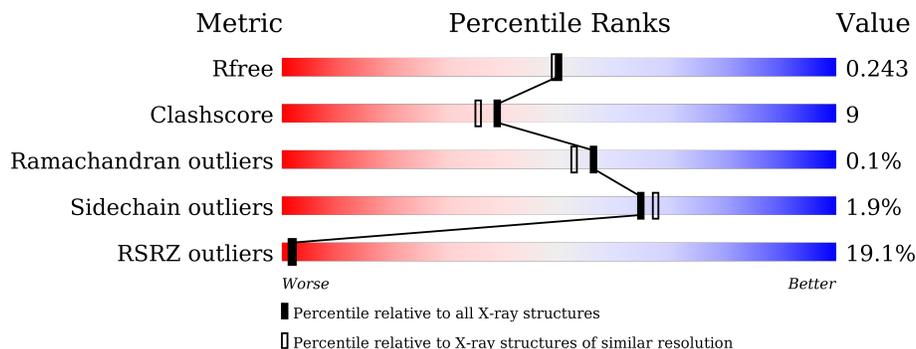
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	6249 (2.00-2.00)
Clashscore	102246	7340 (2.00-2.00)
Ramachandran outliers	100387	7248 (2.00-2.00)
Sidechain outliers	100360	7247 (2.00-2.00)
RSRZ outliers	91569	6262 (2.00-2.00)

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 <=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	617	17% / 78% 13% 9%
1	B	617	17% / 78% 13% • 8%
1	C	617	16% / 74% 12% • 14%
1	D	617	18% / 74% 11% • 14%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 18597 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 Polyketide synthase PksL.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	563	Total	C	N	O	S	0	0	0
			4419	2811	738	845	25			
1	B	568	Total	C	N	O	S	0	0	0
			4461	2837	746	852	26			
1	C	533	Total	C	N	O	S	0	0	0
			4185	2663	703	798	21			
1	D	532	Total	C	N	O	S	0	0	0
			4174	2657	700	796	21			

There are 84 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-21	MET	-	initiating methionine	UNP Q05470
A	-20	GLY	-	expression tag	UNP Q05470
A	-19	SER	-	expression tag	UNP Q05470
A	-18	SER	-	expression tag	UNP Q05470
A	-17	HIS	-	expression tag	UNP Q05470
A	-16	HIS	-	expression tag	UNP Q05470
A	-15	HIS	-	expression tag	UNP Q05470
A	-14	HIS	-	expression tag	UNP Q05470
A	-13	HIS	-	expression tag	UNP Q05470
A	-12	HIS	-	expression tag	UNP Q05470
A	-11	SER	-	expression tag	UNP Q05470
A	-10	SER	-	expression tag	UNP Q05470
A	-9	GLY	-	expression tag	UNP Q05470
A	-8	LEU	-	expression tag	UNP Q05470
A	-7	VAL	-	expression tag	UNP Q05470
A	-6	PRO	-	expression tag	UNP Q05470
A	-5	ARG	-	expression tag	UNP Q05470
A	-4	GLY	-	expression tag	UNP Q05470
A	-3	SER	-	expression tag	UNP Q05470
A	-2	SER	-	expression tag	UNP Q05470
A	169	SER	CYS	engineered mutation	UNP Q05470

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-21	MET	-	initiating methionine	UNP Q05470
B	-20	GLY	-	expression tag	UNP Q05470
B	-19	SER	-	expression tag	UNP Q05470
B	-18	SER	-	expression tag	UNP Q05470
B	-17	HIS	-	expression tag	UNP Q05470
B	-16	HIS	-	expression tag	UNP Q05470
B	-15	HIS	-	expression tag	UNP Q05470
B	-14	HIS	-	expression tag	UNP Q05470
B	-13	HIS	-	expression tag	UNP Q05470
B	-12	HIS	-	expression tag	UNP Q05470
B	-11	SER	-	expression tag	UNP Q05470
B	-10	SER	-	expression tag	UNP Q05470
B	-9	GLY	-	expression tag	UNP Q05470
B	-8	LEU	-	expression tag	UNP Q05470
B	-7	VAL	-	expression tag	UNP Q05470
B	-6	PRO	-	expression tag	UNP Q05470
B	-5	ARG	-	expression tag	UNP Q05470
B	-4	GLY	-	expression tag	UNP Q05470
B	-3	SER	-	expression tag	UNP Q05470
B	-2	SER	-	expression tag	UNP Q05470
B	169	SER	CYS	engineered mutation	UNP Q05470
C	-21	MET	-	initiating methionine	UNP Q05470
C	-20	GLY	-	expression tag	UNP Q05470
C	-19	SER	-	expression tag	UNP Q05470
C	-18	SER	-	expression tag	UNP Q05470
C	-17	HIS	-	expression tag	UNP Q05470
C	-16	HIS	-	expression tag	UNP Q05470
C	-15	HIS	-	expression tag	UNP Q05470
C	-14	HIS	-	expression tag	UNP Q05470
C	-13	HIS	-	expression tag	UNP Q05470
C	-12	HIS	-	expression tag	UNP Q05470
C	-11	SER	-	expression tag	UNP Q05470
C	-10	SER	-	expression tag	UNP Q05470
C	-9	GLY	-	expression tag	UNP Q05470
C	-8	LEU	-	expression tag	UNP Q05470
C	-7	VAL	-	expression tag	UNP Q05470
C	-6	PRO	-	expression tag	UNP Q05470
C	-5	ARG	-	expression tag	UNP Q05470
C	-4	GLY	-	expression tag	UNP Q05470
C	-3	SER	-	expression tag	UNP Q05470
C	-2	SER	-	expression tag	UNP Q05470
C	169	SER	CYS	engineered mutation	UNP Q05470

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	-21	MET	-	initiating methionine	UNP Q05470
D	-20	GLY	-	expression tag	UNP Q05470
D	-19	SER	-	expression tag	UNP Q05470
D	-18	SER	-	expression tag	UNP Q05470
D	-17	HIS	-	expression tag	UNP Q05470
D	-16	HIS	-	expression tag	UNP Q05470
D	-15	HIS	-	expression tag	UNP Q05470
D	-14	HIS	-	expression tag	UNP Q05470
D	-13	HIS	-	expression tag	UNP Q05470
D	-12	HIS	-	expression tag	UNP Q05470
D	-11	SER	-	expression tag	UNP Q05470
D	-10	SER	-	expression tag	UNP Q05470
D	-9	GLY	-	expression tag	UNP Q05470
D	-8	LEU	-	expression tag	UNP Q05470
D	-7	VAL	-	expression tag	UNP Q05470
D	-6	PRO	-	expression tag	UNP Q05470
D	-5	ARG	-	expression tag	UNP Q05470
D	-4	GLY	-	expression tag	UNP Q05470
D	-3	SER	-	expression tag	UNP Q05470
D	-2	SER	-	expression tag	UNP Q05470
D	169	SER	CYS	engineered mutation	UNP Q05470

- Molecule 2 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	371	Total 371	O 371	0	0
2	B	333	Total 333	O 333	0	0
2	C	320	Total 320	O 320	0	0
2	D	334	Total 334	O 334	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: Polyketide synthase PksL
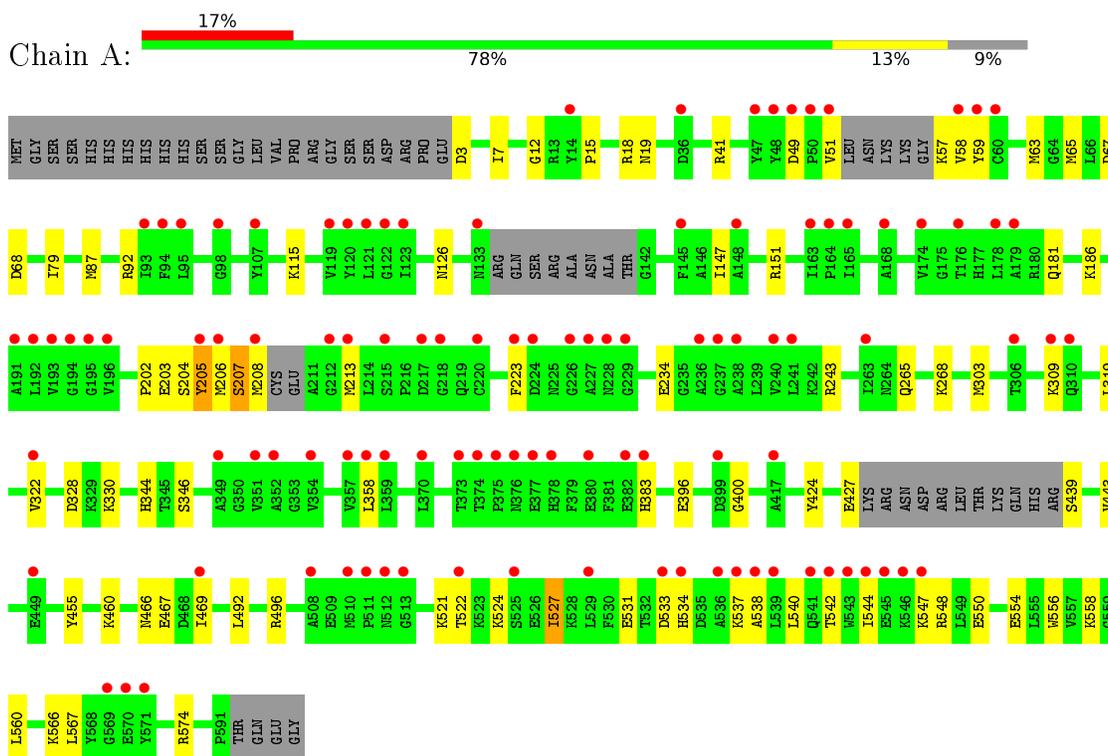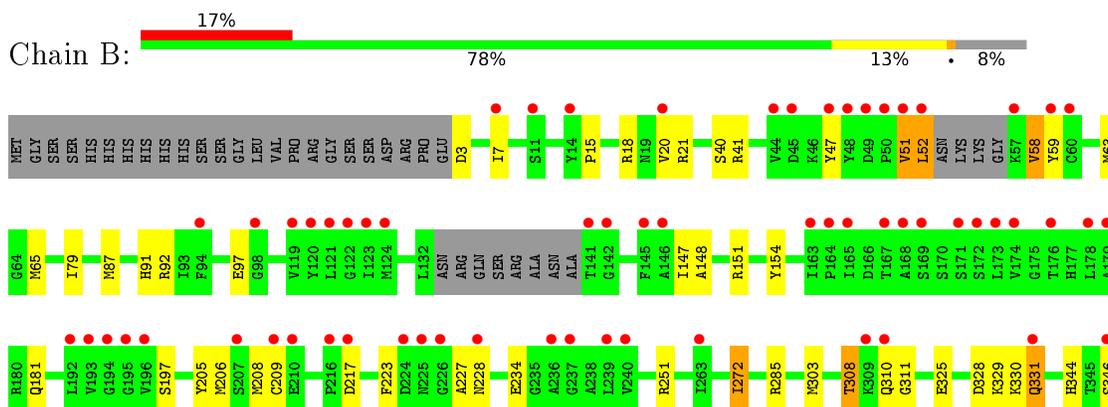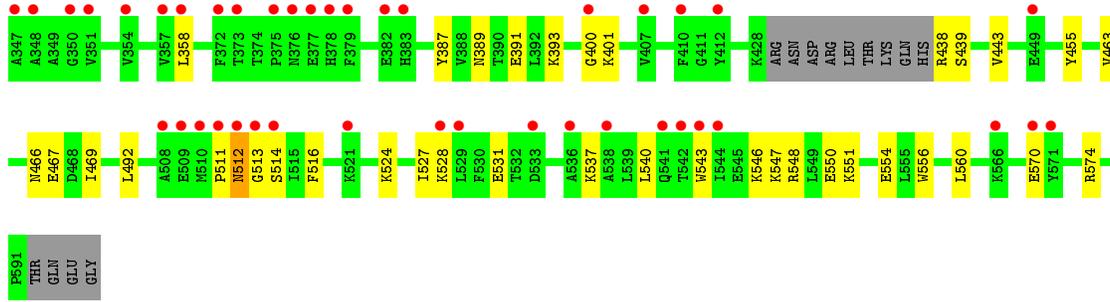
- Molecule 1: Polyketide synthase PksL

• Molecule 1: Polyketide synthase PksL

Chain C:

16%

74%　　　12%　•　14%

• Molecule 1: Polyketide synthase PksL

Chain D:

18%

74%　　　11%　•　14%
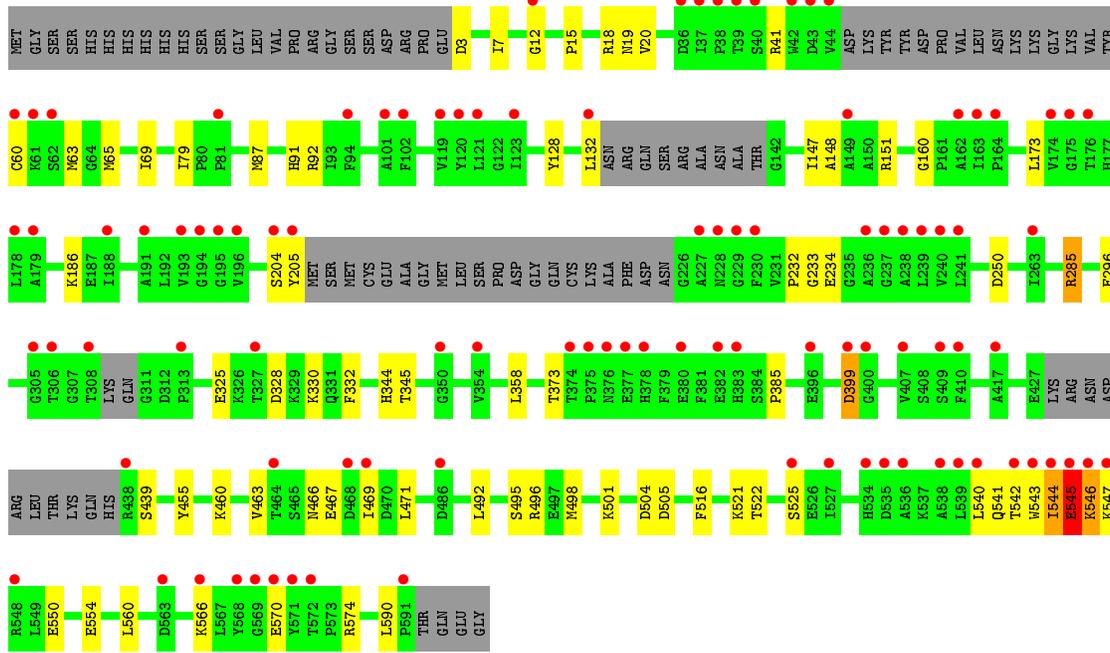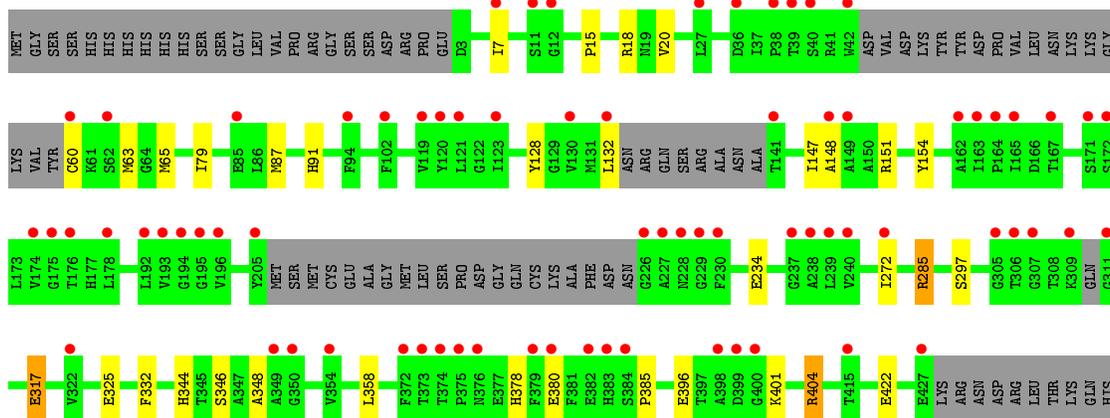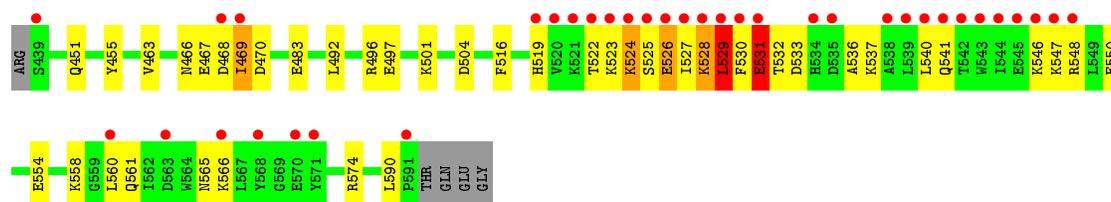
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	89.44Å 108.19Å 151.85Å 90.00° 96.43° 90.00°	Depositor
Resolution (Å)	150.00 – 2.00 44.92 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.2 (150.00-2.00) 99.2 (44.92-2.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.48 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.8.0107	Depositor
R, R_{free}	0.202 , 0.236 0.210 , 0.243	Depositor DCC
R_{free} test set	9700 reflections (5.30%)	DCC
Wilson B-factor (Å^2)	35.3	Xtriage
Anisotropy	0.029	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 52.3	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	18597	wwPDB-VP
Average B, all atoms (Å^2)	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.07% 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 (i)

5.1 Standard geometry (i)

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	0.94	0/4518	0.72	2/6110 (0.0%)
1	B	0.93	2/4561 (0.0%)	0.79	7/6168 (0.1%)
1	C	0.90	0/4277	0.71	7/5784 (0.1%)
1	D	0.98	6/4266 (0.1%)	0.77	10/5768 (0.2%)
All	All	0.94	8/17622 (0.0%)	0.75	26/23830 (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	2
1	B	0	3
1	C	0	2
1	D	0	2
All	All	0	9

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	531	GLU	CG-CD	-6.28	1.42	1.51
1	D	525	SER	CB-OG	6.25	1.50	1.42
1	D	297	SER	CB-OG	6.19	1.50	1.42
1	B	97	GLU	CD-OE1	5.93	1.32	1.25
1	D	404	ARG	CD-NE	-5.65	1.36	1.46

The worst 5 of 26 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	51	VAL	CB-CA-C	-18.49	76.28	111.40
1	B	514	SER	N-CA-CB	-14.74	88.39	110.50
1	B	52	LEU	N-CA-CB	-12.88	84.65	110.40

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	404	ARG	NE-CZ-NH2	12.12	126.36	120.30
1	D	529	LEU	CB-CA-C	-10.67	89.92	110.20

There are no chirality outliers.

5 of 9 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	400	GLY	Peptide
1	A	92	ARG	Sidechain
1	B	227	ALA	Peptide
1	B	400	GLY	Peptide
1	B	92	ARG	Sidechain

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	4419	0	4323	83	1
1	B	4461	0	4373	74	0
1	C	4185	0	4102	71	1
1	D	4174	0	4092	90	0
2	A	371	0	0	18	1
2	B	333	0	0	15	1
2	C	320	0	0	20	0
2	D	334	0	0	14	1
All	All	18597	0	16890	308	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 9.

The worst 5 of 308 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:D:531:GLU:O	1:D:537:LYS:NZ	1.62	1.29
1:C:542:THR:O	2:C:601:HOH:O	1.55	1.19

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:467:GLU:CD	2:B:601:HOH:O	1.81	1.18
1:D:528:LYS:O	1:D:531:GLU:HB3	1.51	1.09
1:D:529:LEU:CD1	1:D:529:LEU:O	2.03	1.07

All (3) 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)
2:A:636:HOH:O	2:A:679:HOH:O[2_745]	2.05	0.15
1:A:467:GLU:OE1	1:C:460:LYS:NZ[1_655]	2.08	0.12
2:B:899:HOH:O	2:D:889:HOH:O[1_455]	2.11	0.09

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	553/617 (90%)	539 (98%)	13 (2%)	1 (0%)	52	48
1	B	560/617 (91%)	545 (97%)	15 (3%)	0	100	100
1	C	521/617 (84%)	510 (98%)	10 (2%)	1 (0%)	52	48
1	D	520/617 (84%)	504 (97%)	16 (3%)	0	100	100
All	All	2154/2468 (87%)	2098 (97%)	54 (2%)	2 (0%)	56	53

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	205	TYR
1	C	544	ILE

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	468/515 (91%)	463 (99%)	5 (1%)	80	83
1	B	473/515 (92%)	465 (98%)	8 (2%)	68	71
1	C	442/515 (86%)	434 (98%)	8 (2%)	66	69
1	D	441/515 (86%)	428 (97%)	13 (3%)	50	49
All	All	1824/2060 (88%)	1790 (98%)	34 (2%)	65	67

5 of 34 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	439	SER
1	C	525	SER
1	D	529	LEU
1	C	504	ASP
1	B	58	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	344	HIS
1	B	270	ASN
1	B	344	HIS
1	B	378	HIS
1	C	541	GLN

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)

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(Å²)	Q<0.9
1	A	563/617 (91%)	0.90	107 (19%)	2	2	23, 37, 68, 106	0
1	B	568/617 (92%)	0.89	104 (18%)	2	2	25, 39, 71, 109	0
1	C	533/617 (86%)	0.90	99 (18%)	2	2	24, 40, 76, 108	0
1	D	532/617 (86%)	1.02	110 (20%)	1	1	25, 39, 71, 106	0
All	All	2196/2468 (88%)	0.93	420 (19%)	2	2	23, 39, 73, 109	0

The worst 5 of 420 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	527	ILE	12.4
1	B	511	PRO	9.3
1	D	535	ASP	8.0
1	D	571	TYR	8.0
1	A	374	THR	7.6

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)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.