



wwPDB X-ray Structure Validation Summary Report ⓘ

May 21, 2020 – 11:40 pm BST

PDB ID : 2QO3
Title : Crystal Structure of [KS3][AT3] didomain from module 3 of 6-deoxyerthronolide B synthase
Authors : Khosla, C.; Cane, E.D.; Tang, Y.; Chen, Y.A.; Kim, C.Y.
Deposited on : 2007-07-19
Resolution : 2.59 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
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

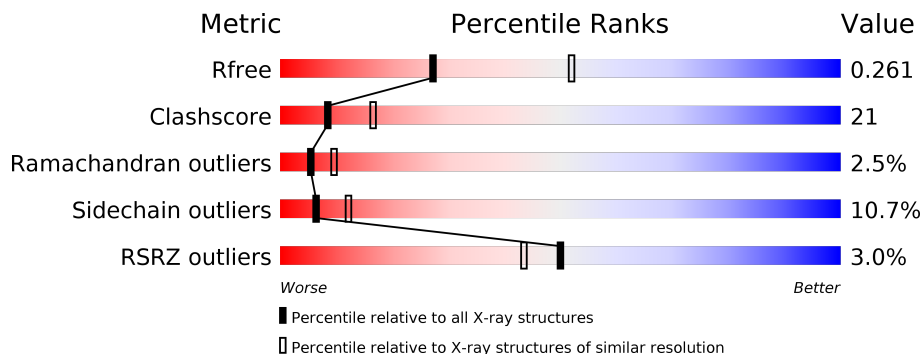
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.59 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 $\leq 5\%$. 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	915	 2% 60% 30% 5% • 5%
1	B	915	 4% 61% 27% 6% • 5%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	ACT	A	950	-	-	X	-

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 13148 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 EryAII Erythromycin polyketide synthase modules 3 and 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	869	6449	4003	1178	1250	18	0	0	0
1	B	873	6479	4020	1185	1256	18	0	0	0

There are 38 discrepancies between the modelled and reference sequences:

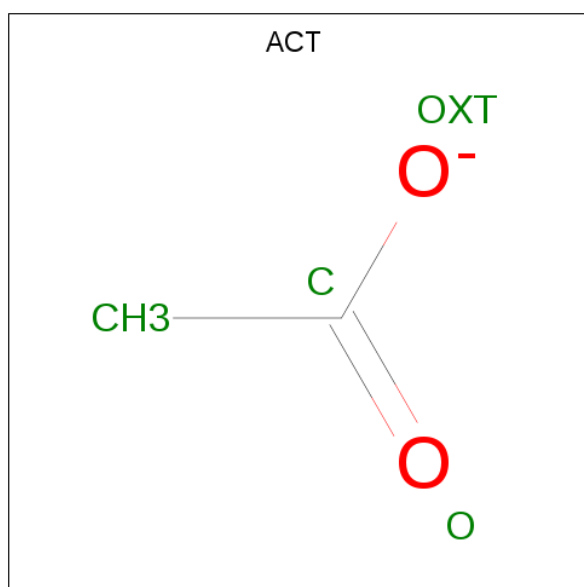
Chain	Residue	Modelled	Actual	Comment	Reference
A	26	MET	-	INITIATING METHIONINE	UNP A4F7P0
A	923	SER	-	EXPRESSION TAG	UNP A4F7P0
A	924	SER	-	EXPRESSION TAG	UNP A4F7P0
A	925	SER	-	EXPRESSION TAG	UNP A4F7P0
A	926	VAL	-	EXPRESSION TAG	UNP A4F7P0
A	927	ASP	-	EXPRESSION TAG	UNP A4F7P0
A	928	LYS	-	EXPRESSION TAG	UNP A4F7P0
A	929	LEU	-	EXPRESSION TAG	UNP A4F7P0
A	930	ALA	-	EXPRESSION TAG	UNP A4F7P0
A	931	ALA	-	EXPRESSION TAG	UNP A4F7P0
A	932	ALA	-	EXPRESSION TAG	UNP A4F7P0
A	933	LEU	-	EXPRESSION TAG	UNP A4F7P0
A	934	GLU	-	EXPRESSION TAG	UNP A4F7P0
A	935	HIS	-	EXPRESSION TAG	UNP A4F7P0
A	936	HIS	-	EXPRESSION TAG	UNP A4F7P0
A	937	HIS	-	EXPRESSION TAG	UNP A4F7P0
A	938	HIS	-	EXPRESSION TAG	UNP A4F7P0
A	939	HIS	-	EXPRESSION TAG	UNP A4F7P0
A	940	HIS	-	EXPRESSION TAG	UNP A4F7P0
B	26	MET	-	INITIATING METHIONINE	UNP A4F7P0
B	923	SER	-	EXPRESSION TAG	UNP A4F7P0
B	924	SER	-	EXPRESSION TAG	UNP A4F7P0
B	925	SER	-	EXPRESSION TAG	UNP A4F7P0
B	926	VAL	-	EXPRESSION TAG	UNP A4F7P0
B	927	ASP	-	EXPRESSION TAG	UNP A4F7P0

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	928	LYS	-	EXPRESSION TAG	UNP A4F7P0
B	929	LEU	-	EXPRESSION TAG	UNP A4F7P0
B	930	ALA	-	EXPRESSION TAG	UNP A4F7P0
B	931	ALA	-	EXPRESSION TAG	UNP A4F7P0
B	932	ALA	-	EXPRESSION TAG	UNP A4F7P0
B	933	LEU	-	EXPRESSION TAG	UNP A4F7P0
B	934	GLU	-	EXPRESSION TAG	UNP A4F7P0
B	935	HIS	-	EXPRESSION TAG	UNP A4F7P0
B	936	HIS	-	EXPRESSION TAG	UNP A4F7P0
B	937	HIS	-	EXPRESSION TAG	UNP A4F7P0
B	938	HIS	-	EXPRESSION TAG	UNP A4F7P0
B	939	HIS	-	EXPRESSION TAG	UNP A4F7P0
B	940	HIS	-	EXPRESSION TAG	UNP A4F7P0

- Molecule 2 is ACETATE ION (three-letter code: ACT) (formula: C₂H₃O₂).

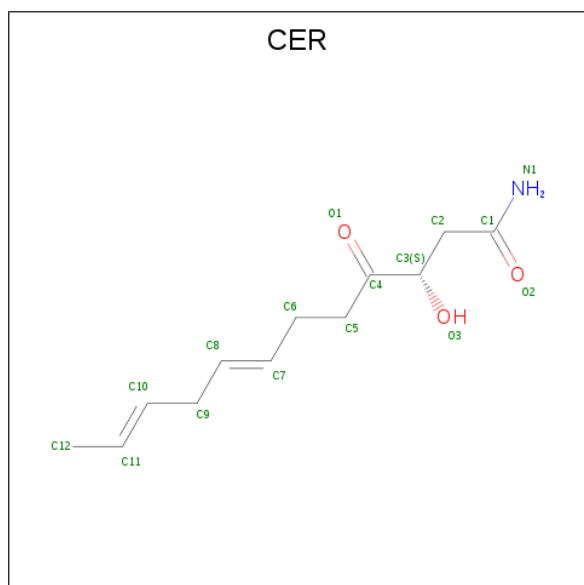


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			4	2	2		
2	B	1	Total	C	O	0	0
			4	2	2		

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total Cl 1 1	0	0
3	A	1	Total Cl 1 1	0	0

- Molecule 4 is (2S, 3R)-3-HYDROXY-4-OXO-7,10-TRANS,TRANS-DODECADIENAMIDE (three-letter code: CER) (formula: C₁₂H₁₉NO₃).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C N O 8 4 1 3	0	0
4	B	1	Total C N O 8 4 1 3	0	0

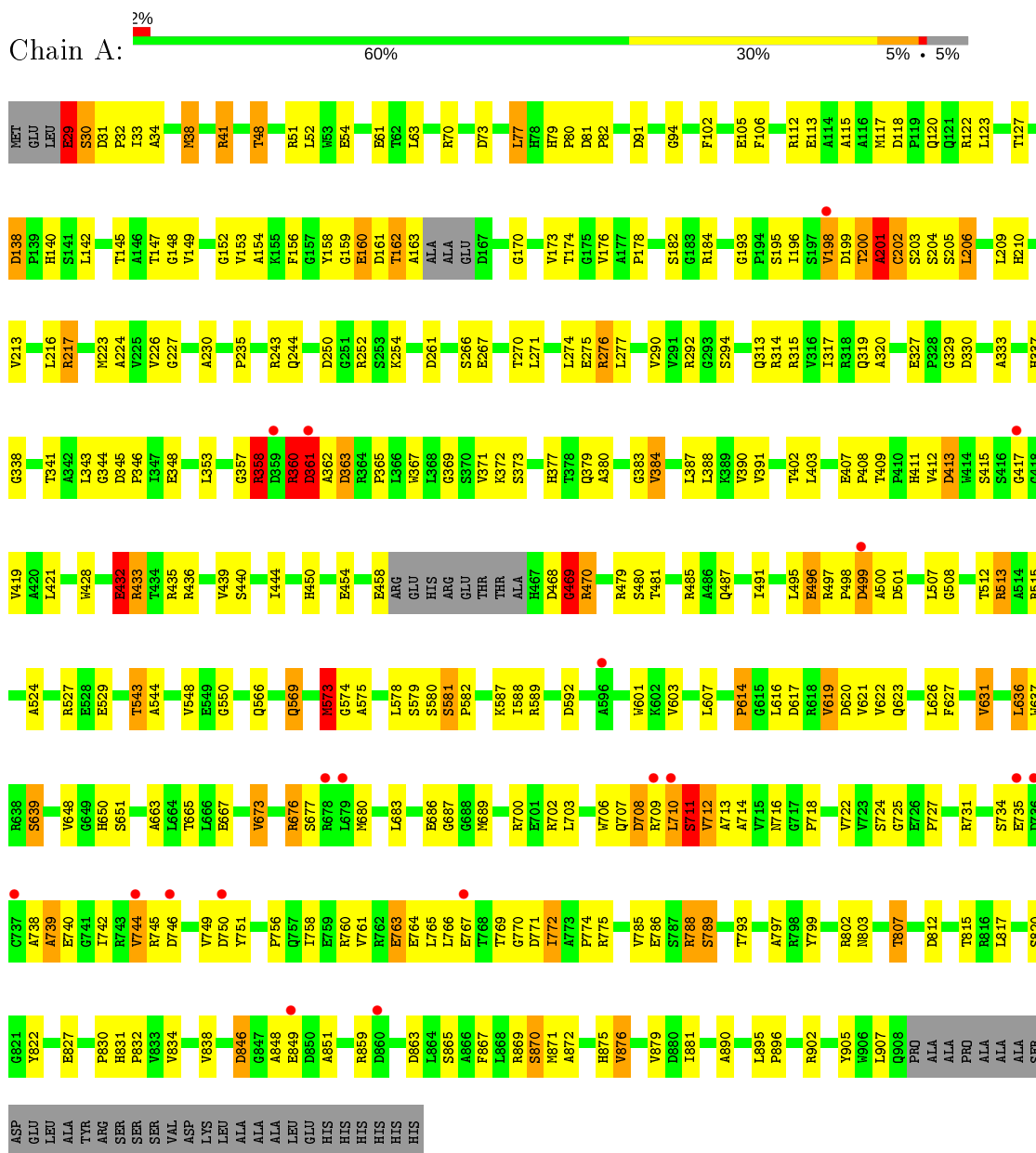
- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	95	Total O 95 95	0	0
5	B	99	Total O 99 99	0	0

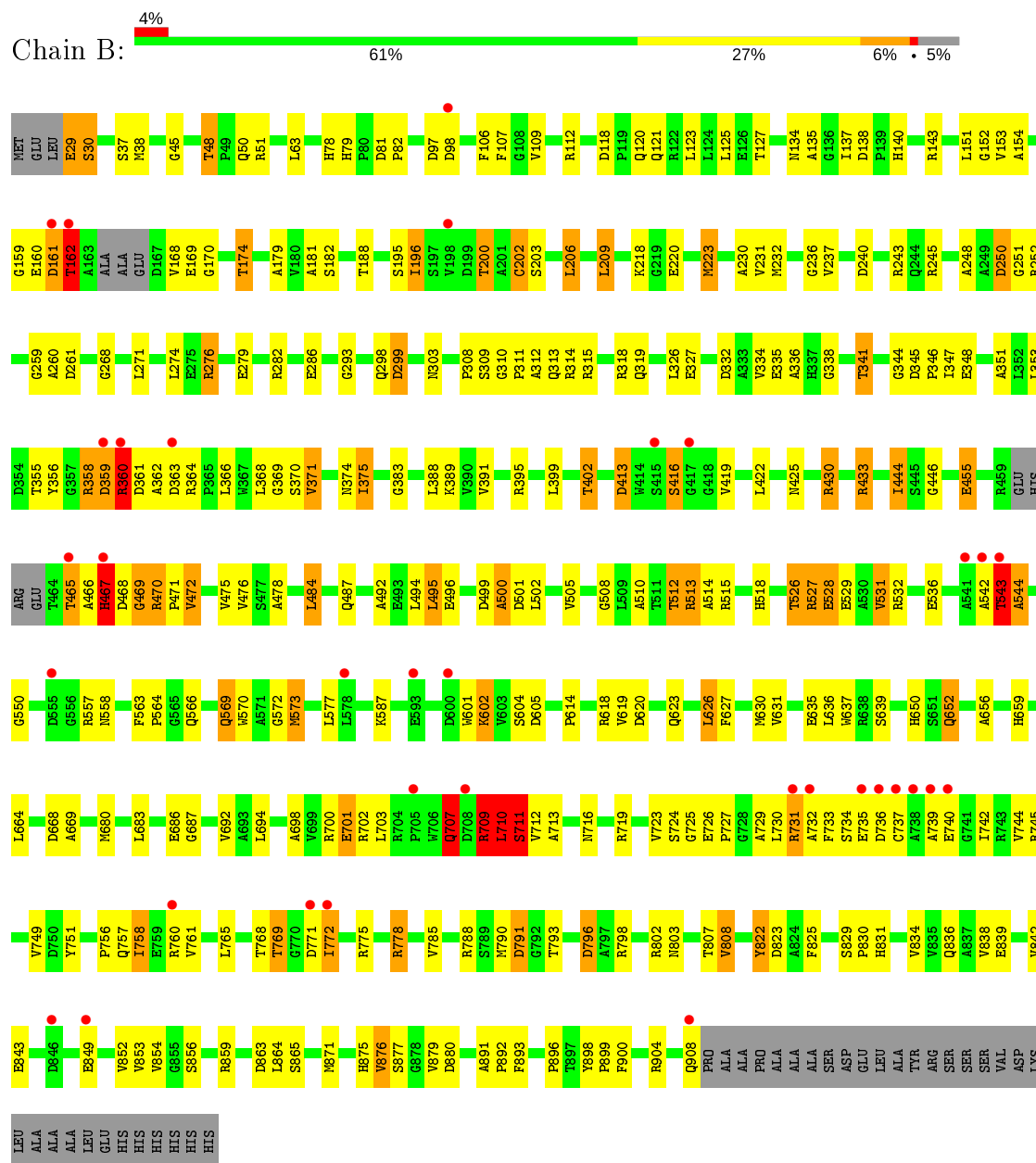
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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: EryAII Erythromycin polyketide synthase modules 3 and 4



- Molecule 1: EryAII Erythromycin polyketide synthase modules 3 and 4



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	75.20Å 139.00Å 102.34Å 90.00° 106.14° 90.00°	Depositor
Resolution (Å)	41.52 – 2.59 41.51 – 2.59	Depositor EDS
% Data completeness (in resolution range)	99.5 (41.52-2.59) 99.4 (41.51-2.59)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.12 (at 2.61Å)	Xtrriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.214 , 0.268 0.210 , 0.261	Depositor DCC
R_{free} test set	3145 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å ²)	37.7	Xtrriage
Anisotropy	0.062	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 39.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	13148	wwPDB-VP
Average B, all atoms (Å ²)	21.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.39% of the height of the origin peak. No significant pseudotranslation is detected.*

¹ Intensities estimated from amplitudes.

² Theoretical 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 [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: CL, CER, ACT

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	0.83	0/6573	0.97	18/8942 (0.2%)
1	B	0.86	1/6603 (0.0%)	0.95	13/8983 (0.1%)
All	All	0.84	1/13176 (0.0%)	0.96	31/17925 (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	11
1	B	0	9
All	All	0	20

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	327	GLU	CG-CD	5.68	1.60	1.51

The worst 5 of 31 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	201	ALA	O-C-N	-10.90	105.26	122.70
1	A	201	ALA	C-N-CA	10.14	147.05	121.70
1	B	202	CYS	O-C-N	-9.46	107.57	122.70
1	A	360	ARG	NE-CZ-NH1	8.20	124.40	120.30
1	A	202	CYS	C-N-CA	-7.67	102.53	121.70

There are no chirality outliers.

5 of 20 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	160	GLU	Peptide
1	A	201	ALA	Mainchain
1	A	202	CYS	Mainchain
1	A	29	GLU	Peptide
1	A	361	ASP	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	6449	0	6307	269	0
1	B	6479	0	6340	270	0
2	A	4	0	3	2	0
2	B	4	0	3	0	0
3	A	1	0	0	1	0
3	B	1	0	0	0	0
4	A	8	0	5	1	0
4	B	8	0	5	2	0
5	A	95	0	0	22	0
5	B	99	0	0	21	0
All	All	13148	0	12663	532	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 21.

The worst 5 of 532 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:B:711:SER:HB3	1:B:724:SER:N	1.40	1.35
1:B:711:SER:CB	1:B:724:SER:H	1.56	1.19
1:A:711:SER:HB2	1:A:724:SER:H	0.98	1.12
1:A:709:ARG:NH1	1:A:709:ARG:HB2	1.64	1.12
1:B:37:SER:HB2	1:B:135:ALA:HB2	1.34	1.07

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	863/915 (94%)	779 (90%)	64 (7%)	20 (2%)	6	11
1	B	867/915 (95%)	781 (90%)	62 (7%)	24 (3%)	5	7
All	All	1730/1830 (94%)	1560 (90%)	126 (7%)	44 (2%)	5	9

5 of 44 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	30	SER
1	A	358	ARG
1	A	415	SER
1	A	432	GLU
1	A	470	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	654/688 (95%)	592 (90%)	62 (10%)	8	16
1	B	657/688 (96%)	579 (88%)	78 (12%)	5	9
All	All	1311/1376 (95%)	1171 (89%)	140 (11%)	6	12

5 of 140 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	48	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	358	ARG
1	B	793	THR
1	B	112	ARG
1	B	202	CYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 24 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	875	HIS
1	B	134	ASN
1	B	875	HIS
1	B	79	HIS
1	B	120	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](#)

Of 6 ligands modelled in this entry, 2 are monoatomic - leaving 4 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
4	CER	A	960	1	6,7,15	0.59	0	6,8,17	0.52	0
4	CER	B	960	1	6,7,15	0.62	0	6,8,17	0.60	0
2	ACT	B	950	-	1,3,3	0.94	0	0,3,3	0.00	-
2	ACT	A	950	-	1,3,3	1.42	0	0,3,3	0.00	-

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	CER	A	960	1	-	0/5/6/16	-
4	CER	B	960	1	-	2/5/6/16	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	B	960	CER	C2-C3-C4-O1
4	B	960	CER	C1-C2-C3-O3

There are no ring outliers.

3 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	960	CER	1	0
4	B	960	CER	2	0
2	A	950	ACT	2	0

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	869/915 (94%)	-0.03	19 (2%) 62 56	7, 20, 38, 50	0
1	B	873/915 (95%)	0.02	34 (3%) 39 32	7, 19, 40, 61	0
All	All	1742/1830 (95%)	-0.00	53 (3%) 50 43	7, 20, 39, 61	0

The worst 5 of 53 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	162	THR	5.6
1	A	499	ASP	4.6
1	B	415	SER	4.0
1	B	908	GLN	4.0
1	B	600	ASP	3.9

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. 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	B-factors(\AA^2)	Q<0.9
4	CER	A	960	8/16	0.88	0.15	38,40,42,43	0
2	ACT	A	950	4/4	0.90	0.20	30,30,31,31	0
3	CL	B	1	1/1	0.93	0.06	29,29,29,29	0
4	CER	B	960	8/16	0.93	0.16	40,42,47,49	0
2	ACT	B	950	4/4	0.96	0.31	37,37,37,37	0
3	CL	A	2	1/1	0.96	0.06	36,36,36,36	0

6.5 Other polymers [i](#)

There are no such residues in this entry.