



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 05:25 AM GMT

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](mailto: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

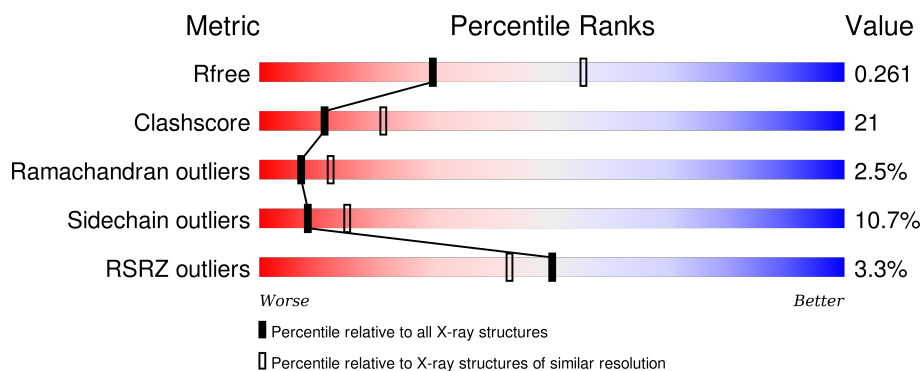
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.59 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))
$R_{free}$	91344	2328 (2.60-2.60)
Clashscore	102246	2679 (2.60-2.60)
Ramachandran outliers	100387	2635 (2.60-2.60)
Sidechain outliers	100360	2635 (2.60-2.60)
RSRZ outliers	91569	2334 (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  $\geq 3$ , 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	915	<div> <div>2%</div> <div>60%</div> <div>30%</div> <div>5% • 5%</div> </div>
1	B	915	<div> <div>4%</div> <div>61%</div> <div>27%</div> <div>6% • 5%</div> </div>

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	ACT	B	950	-	-	-	X

## 2 Entry composition

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 EryAI Erythromycin polyketide synthase modules 3 and 4.

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

There are 38 discrepancies between the modelled and reference sequences:

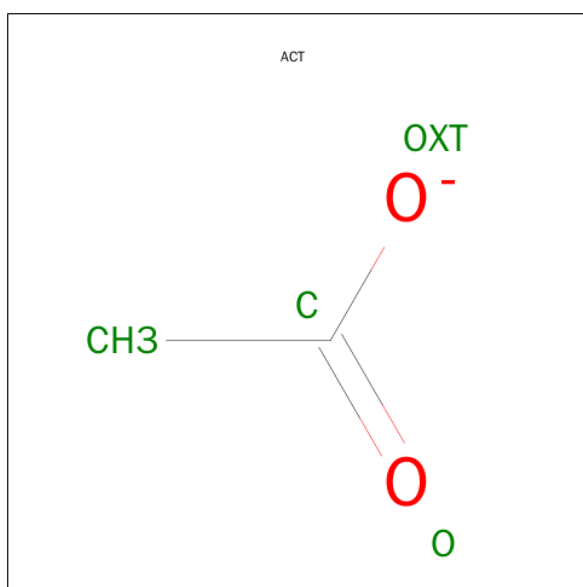
Chain	Residue	Modelled	Actual	Comment	Reference
A	26	MET	-	INITIATING METHIONINE	UNP A4F7P0
A	923	SER	-	EXPRESSION TAG	UNP A4F7P0
A	924	SER	-	EXPRESSION TAG	UNP A4F7P0
A	925	SER	-	EXPRESSION TAG	UNP A4F7P0
A	926	VAL	-	EXPRESSION TAG	UNP A4F7P0
A	927	ASP	-	EXPRESSION TAG	UNP A4F7P0
A	928	LYS	-	EXPRESSION 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:  $C_2H_3O_2$ ).

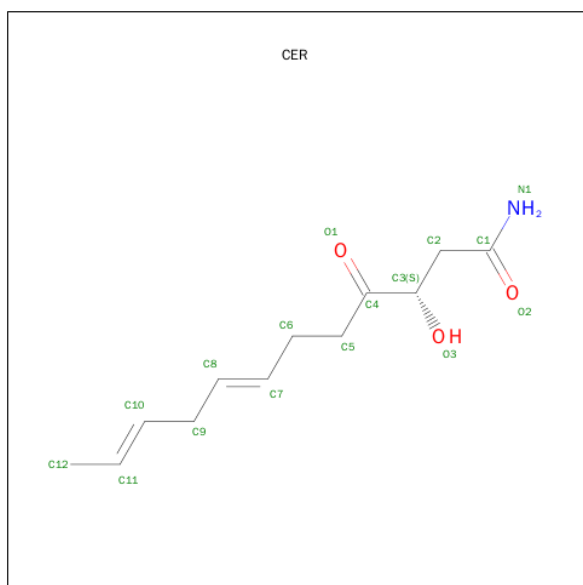


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			4	2	2		
2	B	1	Total	C	O	0	0
			4	2	2		

- 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: C<sub>12</sub>H<sub>19</sub>NO<sub>3</sub>).

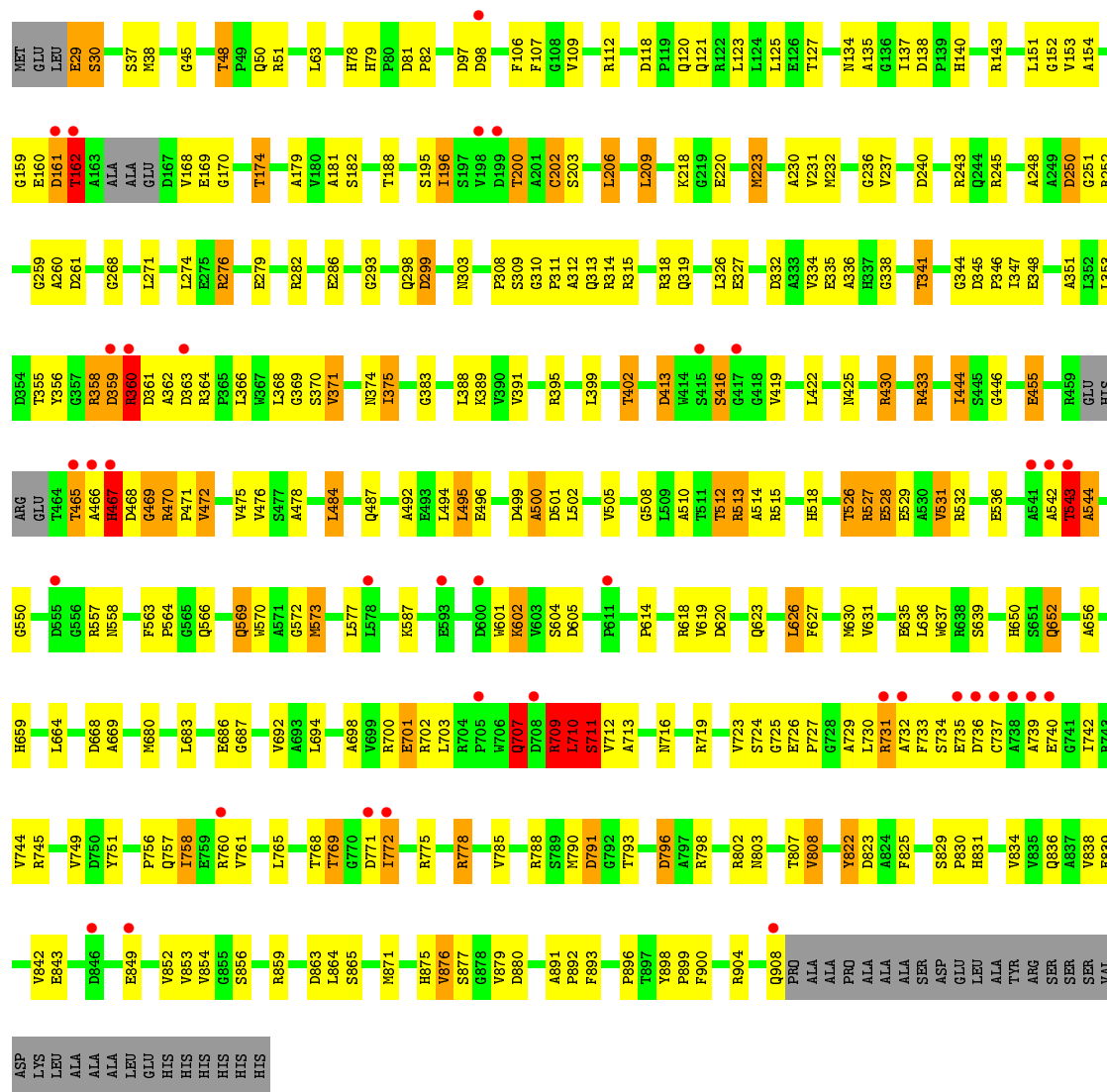


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C N O 8 4 1 3	0	0
4	B	1	Total C N O 8 4 1 3	0	0

- 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







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	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
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.12 (at 2.61Å)	Xtriage
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 (5.34%)	DCC
Wilson B-factor (Å <sup>2</sup> )	37.7	Xtriage
Anisotropy	0.062	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 39.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	1 of 62055 reflections (0.002%)	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	13148	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	21.0	wwPDB-VP

Xtriage'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.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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 [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

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

### 5.3.1 Protein backbone

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%)	8	14
1	B	867/915 (95%)	781 (90%)	62 (7%)	24 (3%)	6	10
All	All	1730/1830 (94%)	1560 (90%)	126 (7%)	44 (2%)	7	12

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

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%)	11	20
1	B	657/688 (96%)	579 (88%)	78 (12%)	6	11
All	All	1311/1376 (95%)	1171 (89%)	140 (11%)	8	15

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
2	ACT	A	950	-	1,3,3	1.24	0	0,3,3	0.00	-
4	CER	A	960	1	7,7,15	0.72	0	7,8,17	0.81	0
2	ACT	B	950	-	1,3,3	0.80	0	0,3,3	0.00	-
4	CER	B	960	1	7,7,15	0.70	0	7,8,17	0.81	0

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
2	ACT	A	950	-	-	0/0/0/0	0/0/0/0
4	CER	A	960	1	-	0/4/6/16	0/0/0/0
2	ACT	B	950	-	-	0/0/0/0	0/0/0/0
4	CER	B	960	1	-	0/4/6/16	0/0/0/0

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.

3 monomers are involved in 5 short contacts:

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

## 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 [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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	869/915 (94%)	0.03	21 (2%) 62 56	7, 20, 38, 50	0
1	B	873/915 (95%)	0.08	37 (4%) 40 32	7, 19, 40, 61	0
All	All	1742/1830 (95%)	0.06	58 (3%) 50 43	7, 20, 39, 61	0

The worst 5 of 58 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	162	THR	5.9
1	A	499	ASP	4.8
1	B	415	SER	4.2
1	B	908	GLN	4.1
1	B	600	ASP	4.1

### 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, 95<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
2	ACT	B	950	4/4	0.96	0.31	4.13	37,37,37,37	0
2	ACT	A	950	4/4	0.90	0.20	1.47	30,30,31,31	0
4	CER	A	960	8/16	0.88	0.15	-0.09	38,40,42,43	0
4	CER	B	960	8/16	0.93	0.15	-0.57	40,42,47,49	0
3	CL	B	1	1/1	0.93	0.06	-	29,29,29,29	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.