



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 07:15 PM GMT

PDB ID : 4O9C
Title : Crystal structure of Beta-ketothiolase (PhaA) from Ralstonia eutropha H16
Authors : Kim, E.J.; Kim, J.; Kim, S.; Kim, K.J.
Deposited on : 2014-01-02
Resolution : 2.00 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
<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

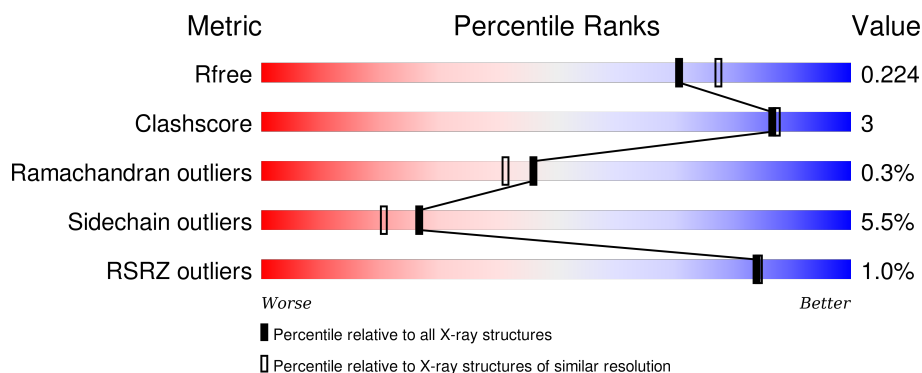
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 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.




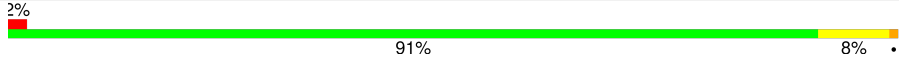
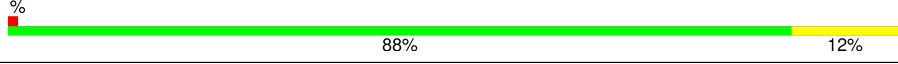
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	6249 (2.00-2.00)
Clashscore	102246	7340 (2.00-2.00)
Ramachandran outliers	100387	7248 (2.00-2.00)
Sidechain outliers	100360	7247 (2.00-2.00)
RSRZ outliers	91569	6262 (2.00-2.00)

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 $\leq 5\%$. 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	393	<div> <div></div> <div>89%10% .</div> </div>
1	B	393	<div> <div></div> <div>91%9% .</div> </div>
1	C	393	<div> <div></div> <div>89%10% .</div> </div>
1	D	393	<div> <div></div> <div>90%9% .</div> </div>
1	E	393	<div> <div></div> <div>87%12% .</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	393	
1	G	393	
1	H	393	

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	COA	C	401	-	-	-	X

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 23650 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.

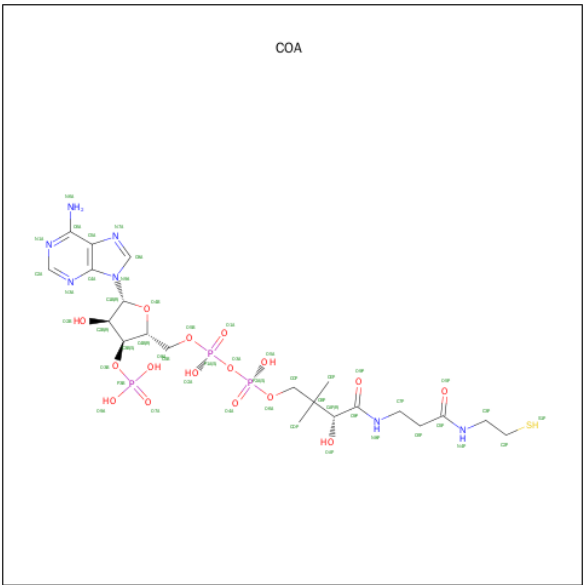
- Molecule 1 is a protein called Acetyl-CoA acetyltransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	393	Total	C	N	O	S	0	0	0
			2835	1774	502	539	20			
1	B	393	Total	C	N	O	S	0	0	0
			2835	1774	502	539	20			
1	C	393	Total	C	N	O	S	0	0	0
			2835	1774	502	539	20			
1	D	393	Total	C	N	O	S	0	0	0
			2835	1774	502	539	20			
1	E	393	Total	C	N	O	S	0	0	0
			2835	1774	502	539	20			
1	F	393	Total	C	N	O	S	0	0	0
			2835	1774	502	539	20			
1	G	393	Total	C	N	O	S	0	0	0
			2835	1774	502	539	20			
1	H	393	Total	C	N	O	S	0	0	0
			2835	1774	502	539	20			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	88	SER	CYS	ENGINEERED MUTATION	UNP P14611
B	88	SER	CYS	ENGINEERED MUTATION	UNP P14611
C	88	SER	CYS	ENGINEERED MUTATION	UNP P14611
D	88	SER	CYS	ENGINEERED MUTATION	UNP P14611
E	88	SER	CYS	ENGINEERED MUTATION	UNP P14611
F	88	SER	CYS	ENGINEERED MUTATION	UNP P14611
G	88	SER	CYS	ENGINEERED MUTATION	UNP P14611
H	88	SER	CYS	ENGINEERED MUTATION	UNP P14611

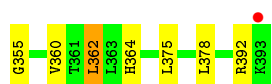
- Molecule 2 is COENZYME A (three-letter code: COA) (formula: C₂₁H₃₆N₇O₁₆P₃S).



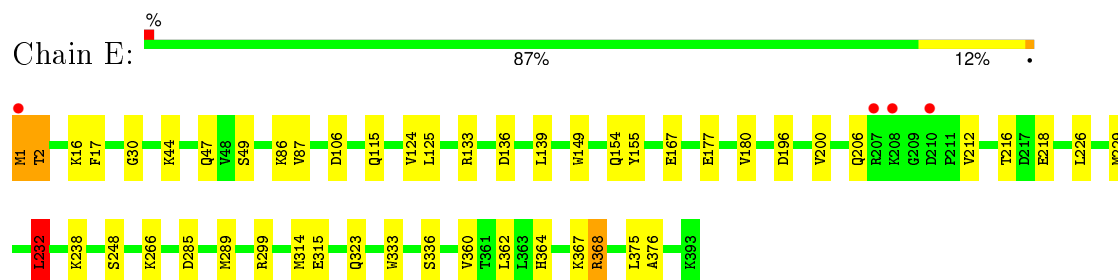
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
2	C	1	48	21	7	16	3	0	0

- Molecule 3 is water.

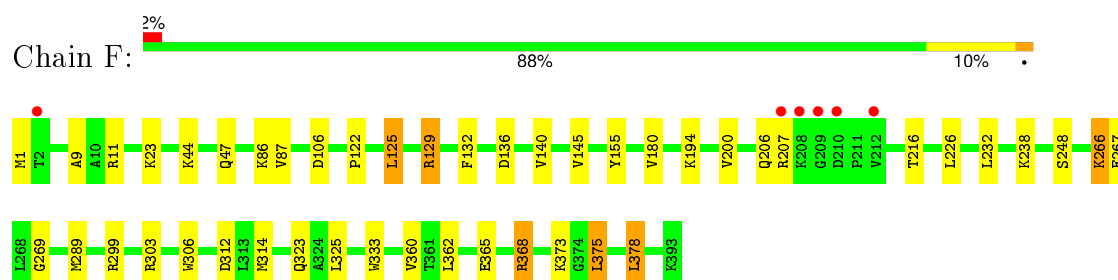
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	131	Total	O	0	0
			131	131		
3	B	119	Total	O	0	0
			119	119		
3	C	134	Total	O	0	0
			134	134		
3	D	122	Total	O	0	0
			122	122		
3	E	114	Total	O	0	0
			114	114		
3	F	98	Total	O	0	0
			98	98		
3	G	95	Total	O	0	0
			95	95		
3	H	109	Total	O	0	0
			109	109		



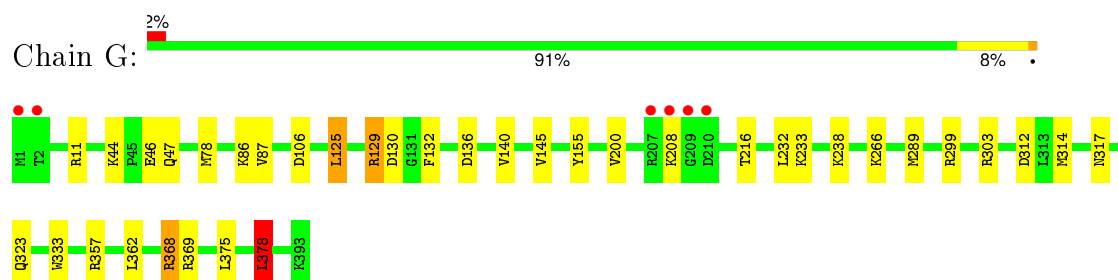
• Molecule 1: Acetyl-CoA acetyltransferase



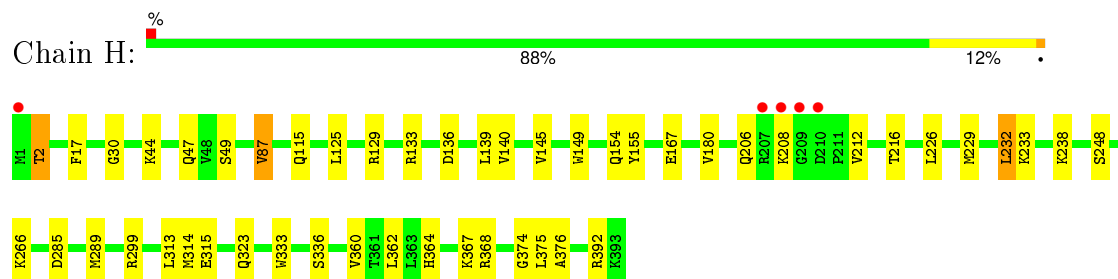
• Molecule 1: Acetyl-CoA acetyltransferase



• Molecule 1: Acetyl-CoA acetyltransferase



• Molecule 1: Acetyl-CoA acetyltransferase



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	102.09Å 157.44Å 114.87Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	102.09 – 2.00 27.71 – 2.00	Depositor EDS
% Data completeness (in resolution range)	98.4 (102.09-2.00) 98.4 (27.71-2.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	16.72 (at 2.01Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, R_{free}	0.173 , 0.213 0.186 , 0.224	Depositor DCC
R_{free} test set	11968 reflections (5.29%)	DCC
Wilson B-factor (Å ²)	19.7	Xtriage
Anisotropy	0.058	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.40 , 28.3	EDS
Estimated twinning fraction	0.478 for h,-k,-l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.35$	Xtriage
Outliers	0 of 238332 reflections	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	23650	wwPDB-VP
Average B, all atoms (Å ²)	22.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.16% 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, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: COA

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	0.91	2/2875 (0.1%)	0.97	8/3886 (0.2%)
1	B	0.90	1/2875 (0.0%)	0.97	6/3886 (0.2%)
1	C	0.92	3/2875 (0.1%)	0.97	10/3886 (0.3%)
1	D	0.91	1/2875 (0.0%)	0.97	12/3886 (0.3%)
1	E	0.93	2/2875 (0.1%)	0.99	9/3886 (0.2%)
1	F	0.92	0/2875	1.00	15/3886 (0.4%)
1	G	0.93	0/2875	1.02	18/3886 (0.5%)
1	H	0.93	1/2875 (0.0%)	0.98	6/3886 (0.2%)
All	All	0.92	10/23000 (0.0%)	0.98	84/31088 (0.3%)

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	E	0	1

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	86	LYS	CD-CE	6.33	1.67	1.51
1	D	86	LYS	CD-CE	5.87	1.66	1.51
1	E	336	SER	CB-OG	-5.76	1.34	1.42
1	C	86	LYS	CD-CE	5.71	1.65	1.51
1	A	86	LYS	CD-CE	5.60	1.65	1.51

The worst 5 of 84 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	375	LEU	CA-CB-CG	8.56	134.99	115.30
1	D	299	ARG	NE-CZ-NH1	8.23	124.42	120.30
1	F	303	ARG	NE-CZ-NH2	-7.93	116.34	120.30
1	B	299	ARG	NE-CZ-NH1	7.77	124.19	120.30
1	G	303	ARG	NE-CZ-NH2	-7.67	116.47	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	E	1	MET	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	2835	0	2907	23	0
1	B	2835	0	2907	17	1
1	C	2835	0	2907	18	0
1	D	2835	0	2907	13	0
1	E	2835	0	2907	18	0
1	F	2835	0	2907	17	1
1	G	2835	0	2907	13	0
1	H	2835	0	2907	19	0
2	C	48	0	32	0	0
3	A	131	0	0	2	0
3	B	119	0	0	2	0
3	C	134	0	0	2	0
3	D	122	0	0	3	0
3	E	114	0	0	3	0
3	F	98	0	0	1	0
3	G	95	0	0	1	0
3	H	109	0	0	2	0
All	All	23650	0	23288	125	1

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 3.

The worst 5 of 125 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:A:200:VAL:CG2	3:A:605:HOH:O	2.29	0.81
1:C:200:VAL:CG2	3:C:602:HOH:O	2.30	0.80
1:E:2:THR:HG22	3:E:575:HOH:O	1.88	0.74
1:C:368:ARG:HB3	1:C:368:ARG:CZ	2.17	0.74
1:A:368:ARG:CZ	1:A:368:ARG:HB3	2.19	0.71

All (1) 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)
1:B:368:ARG:NH2	1:F:269:GLY:O[2_556]	2.16	0.04

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	391/393 (100%)	378 (97%)	11 (3%)	2 (0%)	34	26
1	B	391/393 (100%)	381 (97%)	9 (2%)	1 (0%)	46	41
1	C	391/393 (100%)	380 (97%)	9 (2%)	2 (0%)	34	26
1	D	391/393 (100%)	381 (97%)	9 (2%)	1 (0%)	46	41
1	E	391/393 (100%)	380 (97%)	10 (3%)	1 (0%)	46	41
1	F	391/393 (100%)	380 (97%)	10 (3%)	1 (0%)	46	41
1	G	391/393 (100%)	379 (97%)	11 (3%)	1 (0%)	46	41
1	H	391/393 (100%)	382 (98%)	8 (2%)	1 (0%)	46	41
All	All	3128/3144 (100%)	3041 (97%)	77 (2%)	10 (0%)	46	41

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	207	ARG
1	A	87	VAL
1	B	87	VAL
1	C	87	VAL
1	D	87	VAL

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	285/285 (100%)	268 (94%)	17 (6%)	24	17
1	B	285/285 (100%)	274 (96%)	11 (4%)	39	35
1	C	285/285 (100%)	270 (95%)	15 (5%)	28	22
1	D	285/285 (100%)	271 (95%)	14 (5%)	31	25
1	E	285/285 (100%)	265 (93%)	20 (7%)	19	12
1	F	285/285 (100%)	269 (94%)	16 (6%)	26	20
1	G	285/285 (100%)	271 (95%)	14 (5%)	31	25
1	H	285/285 (100%)	267 (94%)	18 (6%)	22	16
All	All	2280/2280 (100%)	2155 (94%)	125 (6%)	27	21

5 of 125 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	362	LEU
1	E	266	LYS
1	H	232	LEU
1	E	2	THR
1	E	167	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 38 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	364	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	E	184	ASN
1	H	154	GLN
1	E	115	GLN
1	E	364	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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

1 ligand is modelled in this entry.

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	COA	C	401	-	40,50,50	1.53	5 (12%)	50,75,75	2.31	13 (26%)

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	COA	C	401	-	-	0/44/64/64	0/3/3/3

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	401	COA	C2A-N3A	2.59	1.36	1.32
2	C	401	COA	O4B-C1B	2.76	1.44	1.41
2	C	401	COA	C5A-C4A	3.33	1.48	1.40
2	C	401	COA	C2P-S1P	3.99	1.94	1.80
2	C	401	COA	C3P-N4P	5.34	1.58	1.46

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	401	COA	N3A-C2A-N1A	-8.31	122.53	128.89
2	C	401	COA	CDP-CBP-CCP	-5.61	101.23	108.50
2	C	401	COA	O5P-C5P-C6P	-5.37	112.72	121.98
2	C	401	COA	C7P-C6P-C5P	-4.00	105.72	112.31
2	C	401	COA	C3B-C2B-C1B	2.00	104.78	99.98

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	393/393 (100%)	-0.21	3 (0%) 87 88	12, 19, 32, 83	0
1	B	393/393 (100%)	-0.28	1 (0%) 94 94	12, 19, 32, 70	0
1	C	393/393 (100%)	-0.26	3 (0%) 87 88	12, 19, 33, 77	0
1	D	393/393 (100%)	-0.23	3 (0%) 87 88	12, 19, 32, 71	0
1	E	393/393 (100%)	-0.12	4 (1%) 84 84	12, 20, 36, 77	0
1	F	393/393 (100%)	-0.05	6 (1%) 76 77	13, 22, 41, 80	0
1	G	393/393 (100%)	-0.05	6 (1%) 76 77	13, 22, 42, 82	0
1	H	393/393 (100%)	-0.13	5 (1%) 79 80	12, 20, 37, 81	0
All	All	3144/3144 (100%)	-0.17	31 (0%) 84 84	12, 20, 36, 83	0

The worst 5 of 31 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1	MET	7.7
1	E	207	ARG	6.0
1	C	1	MET	5.5
1	F	2	THR	4.9
1	F	209	GLY	4.8

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 [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, 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	LLDF	B-factors(\AA^2)	Q<0.9
2	COA	C	401	48/48	0.76	0.43	17.46	24,125,158,164	0

6.5 Other polymers [i](#)

There are no such residues in this entry.