



# Full wwPDB X-ray Structure Validation Report

Oct 9, 2014 – 08:42 AM EDT

PDB ID : 4R1X  
Title : Crystal structure of a Hypothetical Acyl-CoA ligase (BT\_0428) from *Bacteroides thetaiotaomicron* VPI-5482 at 2.20 Å resolution  
Authors : Joint Center for Structural Genomics (JCSG)  
Deposited on : 2014-08-07  
Resolution : 2.20 Å (reported)

This is a full wwPDB validation 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/ValidationPDFNotes.html>

---

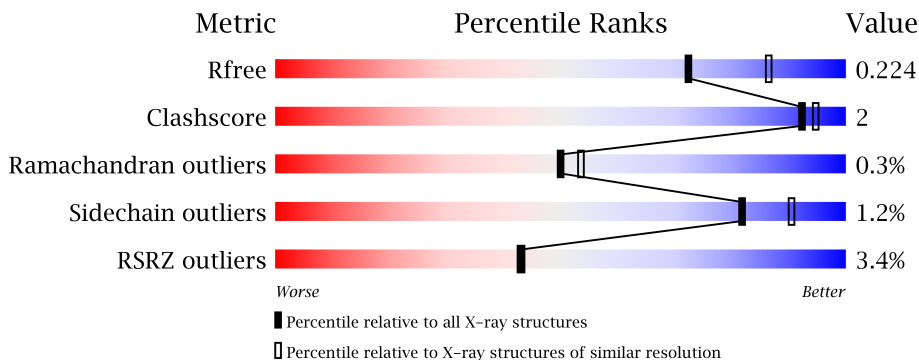
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.16 November 2013  
Xtriage (Phenix) : dev-1439  
EDS : stable23828  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable23828

# 1 Overall quality at a glance

The reported resolution of this entry is 2.20 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))
$R_{free}$	66092	2938 (2.20-2.20)
Clashscore	79885	3751 (2.20-2.20)
Ramachandran outliers	78287	3681 (2.20-2.20)
Sidechain outliers	78261	3682 (2.20-2.20)
RSRZ outliers	66119	2939 (2.20-2.20)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	436	
1	B	436	
1	C	436	
1	D	436	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
5	COA	A	504[A]	-	X
5	COA	C	503[A]	-	X
7	SO4	A	509[B]	-	X
7	SO4	B	504	-	X

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Geometry	Electron density
7	SO4	C	507[B]	-	X
8	EDO	A	512	-	X
8	EDO	A	513	-	X
8	EDO	A	514	-	X
8	EDO	A	515	-	X
8	EDO	A	516	-	X
8	EDO	B	505	-	X
8	EDO	C	509	-	X
8	EDO	C	510	-	X
8	EDO	C	511	-	X
8	EDO	D	505	-	X

## 2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 14649 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Phenylacetate-coenzymeA ligase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	433	Total	C	N	O	S	Se	0	5	0
			3439	2169	593	656	6	15			
1	B	424	Total	C	N	O	S	Se	0	7	0
			3353	2113	581	638	6	15			
1	C	434	Total	C	N	O	S	Se	0	3	0
			3427	2162	597	647	6	15			
1	D	426	Total	C	N	O	S	Se	0	6	0
			3362	2124	577	640	6	15			

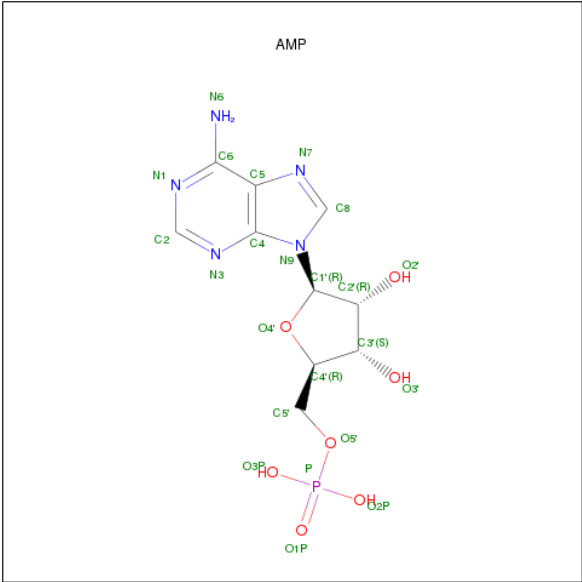
There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	GLY	-	leader sequence	UNP Q8AAN6
B	0	GLY	-	leader sequence	UNP Q8AAN6
C	0	GLY	-	leader sequence	UNP Q8AAN6
D	0	GLY	-	leader sequence	UNP Q8AAN6

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

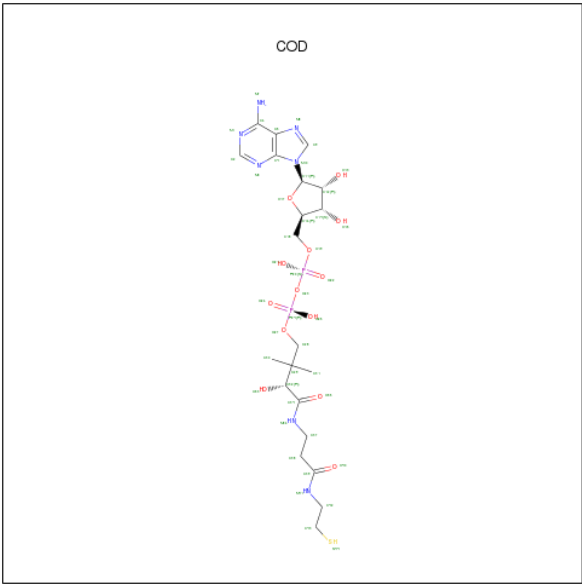
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Zn	0	0
			1	1		
2	A	1	Total	Zn	0	0
			1	1		
2	D	1	Total	Zn	0	0
			1	1		
2	C	1	Total	Zn	0	0
			1	1		

- Molecule 3 is ADENOSINE MONOPHOSPHATE (three-letter code: AMP) (formula: C<sub>10</sub>H<sub>14</sub>N<sub>5</sub>O<sub>7</sub>P).



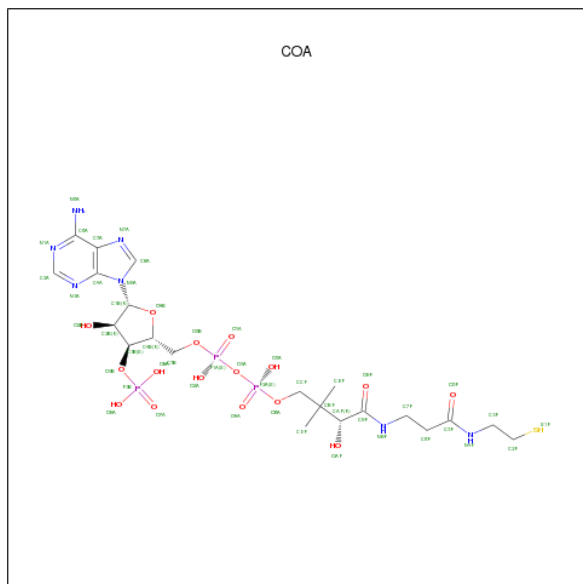
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	1
			23	10	5	7	1		
3	B	1	Total	C	N	O	P	0	1
			23	10	5	7	1		
3	C	1	Total	C	N	O	P	0	0
			23	10	5	7	1		
3	D	1	Total	C	N	O	P	0	1
			23	10	5	7	1		

- Molecule 4 is DEPHOSPHO COENZYME A (three-letter code: COD) (formula: C<sub>21</sub>H<sub>35</sub>N<sub>7</sub>O<sub>13</sub>P<sub>2</sub>S).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
4	A	1	Total 44	C 21	N 7	O 13	P 2	S 1	0	1
4	B	1	Total 44	C 21	N 7	O 13	P 2	S 1	0	1
4	D	1	Total 44	C 21	N 7	O 13	P 2	S 1	0	1

- Molecule 5 is COENZYME A (three-letter code: COA) (formula:  $C_{21}H_{36}N_7O_{16}P_3S$ ).

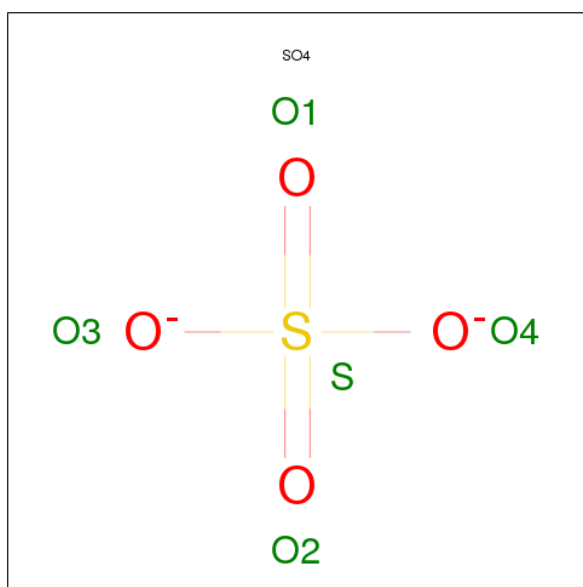


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	1	Total	C	N	O	P	0	1
			31	10	5	13	3		
5	C	1	Total	C	N	O	P	0	1
			31	10	5	13	3		

- Molecule 6 is POTASSIUM ION (three-letter code: K) (formula: K).

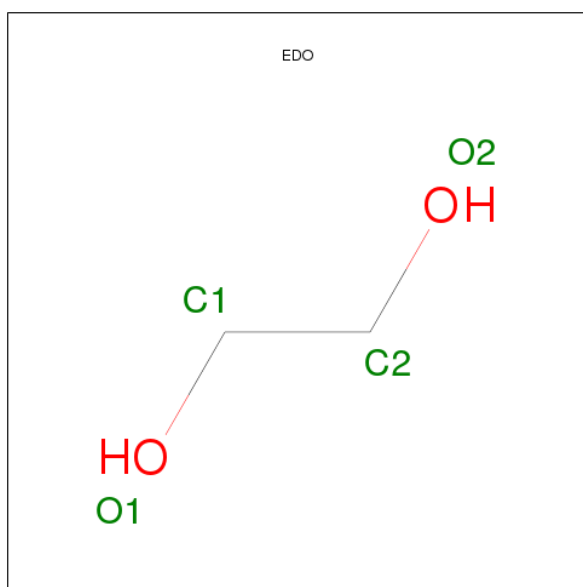
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	4	Total K 4 4	0	0
6	C	3	Total K 3 3	0	0

- Molecule 7 is SULFATE ION (three-letter code: SO4) (formula:  $O_4S$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total	O	S	0	1
			5	4	1		
7	A	1	Total	O	S	0	0
			5	4	1		
7	A	1	Total	O	S	0	0
			5	4	1		
7	B	1	Total	O	S	0	0
			5	4	1		
7	C	1	Total	O	S	0	1
			5	4	1		
7	C	1	Total	O	S	0	0
			5	4	1		
7	D	1	Total	O	S	0	0
			5	4	1		

- Molecule 8 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



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

- Molecule 9 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	223	Total	O	0	1
			224	224		

*Continued on next page...*



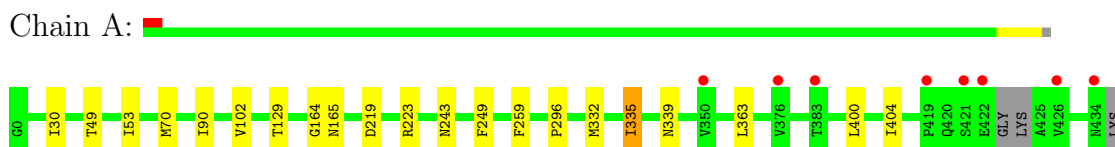
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	B	147	Total 149	O 149	0	5
9	C	140	Total 141	O 141	0	1
9	D	178	Total 178	O 178	0	3

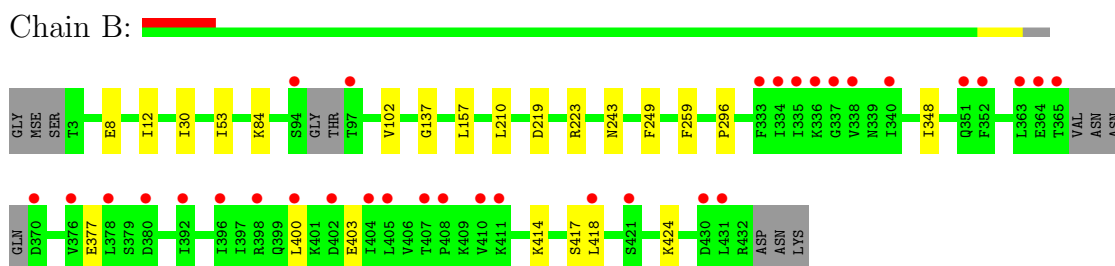
### 3 Residue-property plots

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.

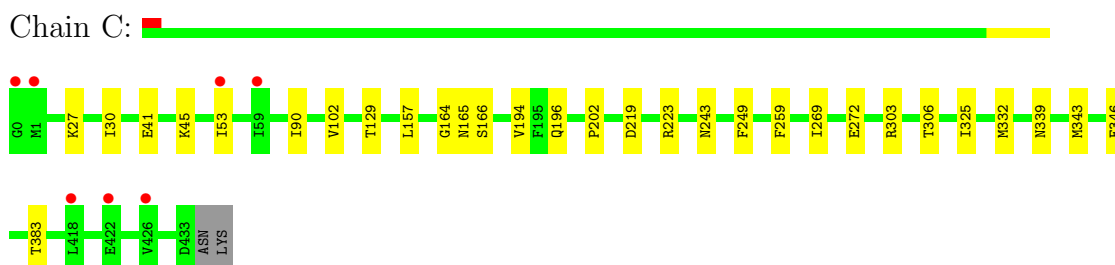
- Molecule 1: Phenylacetate-coenzymeA ligase



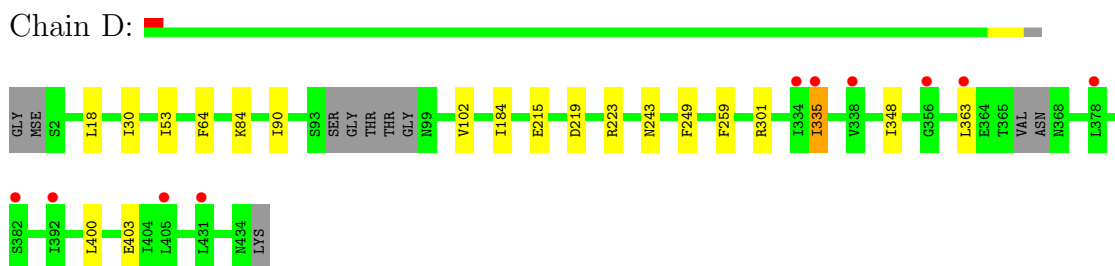
- Molecule 1: Phenylacetate-coenzymeA ligase



- Molecule 1: Phenylacetate-coenzymeA ligase



- Molecule 1: Phenylacetate-coenzymeA ligase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	127.78Å 211.30Å 71.98Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.78 – 2.20 47.78 – 2.20	Depositor EDS
% Data completeness (in resolution range)	99.9 (47.78-2.20) 99.9 (47.78-2.20)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.11 (at 2.20Å)	Xtriage
Refinement program	BUSTER 2.10.0	Depositor
R, $R_{free}$	0.176 , 0.213 0.186 , 0.224	Depositor DCC
$R_{free}$ test set	4990 reflections (5.28%)	DCC
Wilson B-factor (Å <sup>2</sup> )	38.1	Xtriage
Anisotropy	0.564	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 49.9	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	1 of 99569 reflections (0.001%)	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	14649	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	57.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.70% of the height of the origin peak. No significant pseudotranslation is detected.*

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

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, K, EDO, COA, COD, AMP, SO4

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	0.53	0/3497	0.66	0/4697
1	B	0.50	0/3417	0.65	0/4593
1	C	0.50	0/3480	0.64	0/4675
1	D	0.53	0/3423	0.66	0/4602
All	All	0.52	0/13817	0.65	0/18567

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3439	0	3428	13	0
1	B	3353	0	3304	13	0
1	C	3427	0	3417	15	0
1	D	3362	0	3307	12	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	23	0	12	0	0
3	B	23	0	12	0	0
3	C	23	0	12	0	0
3	D	23	0	12	0	0
4	A	44	0	33	0	0
4	B	44	0	33	1	0
4	D	44	0	33	0	0
5	A	31	0	11	1	0
5	C	31	0	11	1	0
6	A	4	0	0	0	0
6	C	3	0	0	0	0
7	A	15	0	0	0	0
7	B	5	0	0	0	0
7	C	10	0	0	0	0
7	D	5	0	0	0	0
8	A	20	0	30	2	0
8	B	4	0	6	0	0
8	C	12	0	18	2	0
8	D	8	0	12	0	0
9	A	224	0	0	1	0
9	B	149	0	0	2	0
9	C	141	0	0	1	0
9	D	178	0	0	0	0
All	All	14649	0	13691	50	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 2.

All (50) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:D:30:ILE:HD13	1:D:53:ILE:HG13	1.76	0.67
1:A:30:ILE:HD13	1:A:53:ILE:HG13	1.76	0.67
1:B:30:ILE:HD13	1:B:53:ILE:HG13	1.78	0.66
1:C:30:ILE:HD13	1:C:53:ILE:HG13	1.80	0.64
1:A:129:THR:HB	1:B:84:LYS:HE2	1.81	0.63
1:D:348:ILE:HD11	1:D:403:GLU:CB	2.29	0.62
1:C:306:THR:HB	1:C:325:ILE:HD13	1.84	0.60
1:B:348:ILE:HD11	1:B:403:GLU:HB3	1.85	0.59
1:D:90:ILE:HG23	1:D:102:VAL:HG13	1.84	0.58
1:A:296:PRO:HD2	9:A:606:HOH:O	2.08	0.54
1:B:348:ILE:HD13	1:B:400:LEU:HA	1.92	0.51
1:B:377:GLU:HG3	1:B:418:LEU:HD12	1.92	0.51

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:C:41:GLU:HG3	1:C:45:LYS:HE2	1.93	0.51
1:C:343:MSE:HE1	1:C:346:GLU:OE1	2.12	0.50
8:C:511:EDO:H11	9:C:646:HOH:O	2.12	0.50
1:D:348:ILE:HD11	1:D:403:GLU:HB3	1.94	0.49
1:A:30:ILE:CD1	1:A:53:ILE:HG13	2.42	0.49
1:D:335:ILE:HD13	1:D:363:LEU:HD22	1.94	0.49
1:C:129:THR:HB	1:D:84:LYS:HE2	1.94	0.49
1:D:30:ILE:CD1	1:D:53:ILE:HG13	2.41	0.49
1:A:90:ILE:HG23	1:A:102:VAL:HG13	1.93	0.49
1:A:49:THR:HA	8:A:512:EDO:H21	1.94	0.49
1:B:414:LYS:HB3	1:B:417:SER:OG	2.13	0.48
1:A:219:ASP:O	1:A:223:ARG:HG2	2.14	0.48
1:A:332:MSE:SE	1:A:339:ASN:HB3	2.63	0.48
1:A:164:GLY:HA2	5:A:504[A]:COA:H1B	1.96	0.46
1:C:219:ASP:O	1:C:223:ARG:HG2	2.15	0.46
1:D:348:ILE:HD13	1:D:400:LEU:HA	1.98	0.46
1:A:70:MSE:O	8:A:514:EDO:H11	2.16	0.46
1:C:90:ILE:HG23	1:C:102:VAL:HG13	1.98	0.46
1:B:30:ILE:CD1	1:B:53:ILE:HG13	2.43	0.45
1:A:90:ILE:HD11	1:B:157:LEU:HD21	1.99	0.44
1:C:30:ILE:CD1	1:C:53:ILE:HG13	2.45	0.44
1:D:219:ASP:O	1:D:223:ARG:HG2	2.18	0.44
1:A:335:ILE:HD13	1:A:363:LEU:HD22	2.00	0.44
1:B:8:GLU:O	1:B:12:ILE:HG12	2.18	0.43
1:B:219:ASP:O	1:B:223:ARG:HG2	2.18	0.43
1:C:164:GLY:HA2	5:C:503[A]:COA:H1B	2.00	0.43
1:C:166:SER:HB3	1:C:194:VAL:HG21	2.01	0.43
1:C:303:ARG:HH11	8:C:511:EDO:H12	1.83	0.43
1:B:137:GLY:H	4:B:503[B]:COD:H431	1.84	0.42
1:B:296:PRO:HD2	9:B:617:HOH:O	2.18	0.42
1:D:184:ILE:HD12	1:D:215:GLU:OE2	2.20	0.42
1:D:64:PHE:HB3	1:D:301:ARG:HG3	2.02	0.42
1:C:27:LYS:HE2	1:C:53:ILE:O	2.20	0.41
1:C:157:LEU:HD21	1:D:90:ILE:HD11	2.02	0.41
1:C:332:MSE:SE	1:C:339:ASN:HB3	2.71	0.41
1:B:424:LYS:HE3	9:B:683:HOH:O	2.21	0.41
1:A:400:LEU:O	1:A:404:ILE:HG12	2.21	0.40
1:C:196:GLN:HG2	1:C:202:PRO:HD3	2.04	0.40

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	434/436 (100%)	417 (96%)	16 (4%)	1 (0%)	56	62
1	B	425/436 (98%)	412 (97%)	12 (3%)	1 (0%)	56	62
1	C	435/436 (100%)	417 (96%)	16 (4%)	2 (0%)	38	38
1	D	426/436 (98%)	408 (96%)	17 (4%)	1 (0%)	56	62
All	All	1720/1744 (99%)	1654 (96%)	61 (4%)	5 (0%)	50	53

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	243	ASN
1	B	243	ASN
1	C	243	ASN
1	D	243	ASN
1	C	383	THR

### 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	376/370 (102%)	372 (99%)	4 (1%)	84	92
1	B	362/370 (98%)	358 (99%)	4 (1%)	84	92
1	C	371/370 (100%)	366 (99%)	5 (1%)	80	89
1	D	361/370 (98%)	357 (99%)	4 (1%)	84	92
All	All	1470/1480 (99%)	1453 (99%)	17 (1%)	82	90

All (17) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	165	ASN
1	A	249	PHE
1	A	259	PHE
1	A	335	ILE
1	B	102	VAL
1	B	210	LEU
1	B	249	PHE
1	B	259	PHE
1	C	165	ASN
1	C	249	PHE
1	C	259	PHE
1	C	269	ILE
1	C	272	GLU
1	D	18	LEU
1	D	249	PHE
1	D	259	PHE
1	D	335	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA ⓘ

There are no RNA chains in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 38 ligands modelled in this entry, 11 are monoatomic - leaving 27 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
3	AMP	A	502[A]	6	25,25,25	0.78	0	38,38,38	0.75	0
4	COD	A	503[B]	6	46,46,46	0.83	1 (2%)	68,68,68	1.23	7 (10%)
5	COA	A	504[A]	-	33,33,50	0.90	2 (6%)	52,52,75	1.52	7 (13%)
7	SO4	A	509[B]	-	4,4,4	0.18	0	6,6,6	0.16	0
7	SO4	A	510	-	4,4,4	0.48	0	6,6,6	0.58	0
7	SO4	A	511	-	4,4,4	0.10	0	6,6,6	0.07	0
8	EDO	A	512	-	3,3,3	0.78	0	2,2,2	0.15	0
8	EDO	A	513	-	3,3,3	0.60	0	2,2,2	0.40	0
8	EDO	A	514	-	3,3,3	0.44	0	2,2,2	0.39	0
8	EDO	A	515	-	3,3,3	0.96	0	2,2,2	0.20	0
8	EDO	A	516	-	3,3,3	0.60	0	2,2,2	0.52	0
3	AMP	B	502[A]	-	25,25,25	0.55	0	38,38,38	0.84	1 (2%)
4	COD	B	503[B]	-	46,46,46	0.72	1 (2%)	68,68,68	1.15	7 (10%)
7	SO4	B	504	-	4,4,4	0.20	0	6,6,6	0.16	0
8	EDO	B	505	-	3,3,3	0.54	0	2,2,2	0.39	0
3	AMP	C	502	6	25,25,25	0.63	0	38,38,38	0.90	0
5	COA	C	503[A]	-	33,33,50	1.07	2 (6%)	52,52,75	1.50	5 (9%)
7	SO4	C	507[B]	-	4,4,4	0.26	0	6,6,6	0.13	0
7	SO4	C	508	-	4,4,4	0.08	0	6,6,6	0.14	0
8	EDO	C	509	-	3,3,3	0.59	0	2,2,2	0.46	0
8	EDO	C	510	-	3,3,3	0.69	0	2,2,2	0.24	0
8	EDO	C	511	-	3,3,3	0.93	0	2,2,2	0.16	0
3	AMP	D	502[A]	-	25,25,25	0.58	0	38,38,38	0.83	1 (2%)
4	COD	D	503[B]	-	46,46,46	0.77	1 (2%)	68,68,68	1.16	7 (10%)
7	SO4	D	504	-	4,4,4	0.26	0	6,6,6	0.26	0
8	EDO	D	505	-	3,3,3	0.75	0	2,2,2	0.19	0
8	EDO	D	506	-	3,3,3	0.63	0	2,2,2	0.41	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
3	AMP	A	502[A]	6	-	0/10/26/26	0/3/3/3

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	COD	A	503[B]	6	-	0/43/59/59	0/3/3/3
5	COA	A	504[A]	-	-	0/21/37/64	0/3/3/3
7	SO4	A	509[B]	-	-	0/0/0/0	0/0/0/0
7	SO4	A	510	-	-	0/0/0/0	0/0/0/0
7	SO4	A	511	-	-	0/0/0/0	0/0/0/0
8	EDO	A	512	-	-	0/1/1/1	0/0/0/0
8	EDO	A	513	-	-	0/1/1/1	0/0/0/0
8	EDO	A	514	-	-	0/1/1/1	0/0/0/0
8	EDO	A	515	-	-	0/1/1/1	0/0/0/0
8	EDO	A	516	-	-	0/1/1/1	0/0/0/0
3	AMP	B	502[A]	-	-	0/10/26/26	0/3/3/3
4	COD	B	503[B]	-	-	0/43/59/59	0/3/3/3
7	SO4	B	504	-	-	0/0/0/0	0/0/0/0
8	EDO	B	505	-	-	0/1/1/1	0/0/0/0
3	AMP	C	502	6	-	0/10/26/26	0/3/3/3
5	COA	C	503[A]	-	-	0/21/37/64	0/3/3/3
7	SO4	C	507[B]	-	-	0/0/0/0	0/0/0/0
7	SO4	C	508	-	-	0/0/0/0	0/0/0/0
8	EDO	C	509	-	-	0/1/1/1	0/0/0/0
8	EDO	C	510	-	-	0/1/1/1	0/0/0/0
8	EDO	C	511	-	-	0/1/1/1	0/0/0/0
3	AMP	D	502[A]	-	-	0/10/26/26	0/3/3/3
4	COD	D	503[B]	-	-	0/43/59/59	0/3/3/3
7	SO4	D	504	-	-	0/0/0/0	0/0/0/0
8	EDO	D	505	-	-	0/1/1/1	0/0/0/0
8	EDO	D	506	-	-	0/1/1/1	0/0/0/0

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	C	503[A]	COA	P2A-O6A	3.85	1.63	1.50
4	A	503[B]	COD	P20-O23	-3.73	1.53	1.59
5	A	504[A]	COA	P2A-O6A	3.36	1.61	1.50
4	D	503[B]	COD	C28-C29	3.35	1.57	1.52
4	B	503[B]	COD	C28-C29	2.59	1.56	1.52
5	C	503[A]	COA	P3B-O9A	-2.38	1.46	1.54
5	A	504[A]	COA	P1A-O3A	2.22	1.63	1.59

All (35) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	C	503[A]	COA	O3A-P1A-O5B	5.99	118.78	102.91

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	D	503[B]	COD	O23-P24-O27	5.70	118.02	102.91
5	A	504[A]	COA	O3A-P1A-O5B	4.84	115.75	102.91
4	B	503[B]	COD	O23-P24-O27	4.22	114.09	102.91
4	A	503[B]	COD	O25-P24-O27	-4.22	87.25	108.51
5	A	504[A]	COA	O6A-P2A-O3A	4.15	119.51	106.53
5	A	504[A]	COA	O3B-P3B-O7A	3.69	117.39	106.80
4	A	503[B]	COD	O23-P24-O27	3.57	112.38	102.91
4	A	503[B]	COD	C31-C29-C28	3.54	113.61	108.71
5	C	503[A]	COA	O9A-P3B-O3B	3.39	116.84	107.09
5	C	503[A]	COA	O6A-P2A-O3A	3.26	116.74	106.53
5	A	504[A]	COA	O5A-P2A-O6A	-3.13	98.18	110.63
4	B	503[B]	COD	O33-C32-C29	3.13	117.62	110.60
5	C	503[A]	COA	O6A-P2A-O4A	-3.11	99.65	114.60
5	A	504[A]	COA	O5A-P2A-O4A	2.96	121.58	110.88
4	B	503[B]	COD	O25-P24-O26	2.84	127.86	112.14
4	A	503[B]	COD	O23-P20-O19	2.73	110.15	102.91
4	B	503[B]	COD	C29-C32-C34	2.62	117.79	113.66
3	D	502[A]	AMP	P-O5'-C5'	2.58	125.23	118.63
5	C	503[A]	COA	O5A-P2A-O4A	2.58	120.21	110.88
5	A	504[A]	COA	O9A-P3B-O3B	-2.56	99.71	107.09
4	A	503[B]	COD	O25-P24-O26	2.53	126.16	112.14
4	A	503[B]	COD	O25-P24-O23	-2.50	93.28	105.14
5	A	504[A]	COA	O3A-P2A-O4A	-2.49	100.25	107.65
4	D	503[B]	COD	O25-P24-O23	-2.47	93.42	105.14
4	B	503[B]	COD	O25-P24-O23	-2.47	93.44	105.14
4	A	503[B]	COD	O35-C34-C32	-2.42	116.69	120.05
4	D	503[B]	COD	O25-P24-O26	2.35	125.17	112.14
4	B	503[B]	COD	C31-C29-C28	2.29	111.87	108.71
4	D	503[B]	COD	C29-C32-C34	2.27	117.23	113.66
4	D	503[B]	COD	O25-P24-O27	-2.26	97.13	108.51
4	D	503[B]	COD	O33-C32-C29	2.25	115.65	110.60
4	D	503[B]	COD	C31-C29-C28	2.15	111.69	108.71
4	B	503[B]	COD	O23-P20-O19	2.07	108.41	102.91
3	B	502[A]	AMP	P-O5'-C5'	2.00	123.76	118.63

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 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, 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	433/436 (99%)	-0.28	8 (1%) 65 66	27, 46, 86, 131	0
1	B	424/436 (97%)	0.19	33 (7%) 13 12	32, 56, 107, 129	0
1	C	434/436 (99%)	-0.10	7 (1%) 68 69	34, 58, 98, 130	0
1	D	426/436 (97%)	-0.10	10 (2%) 57 58	31, 49, 94, 117	0
All	All	1717/1744 (98%)	-0.07	58 (3%) 43 43	27, 53, 99, 131	0

All (58) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	426	VAL	5.9
1	A	421	SER	5.6
1	C	53	ILE	5.3
1	B	400	LEU	5.2
1	B	335	ILE	4.5
1	C	1	MSE	4.3
1	B	94	SER	4.2
1	B	370	ASP	4.2
1	D	338	VAL	4.1
1	B	334	ILE	4.1
1	B	340	ILE	4.1
1	B	363	LEU	3.9
1	B	431	LEU	3.9
1	B	408	PRO	3.9
1	B	398[A]	ARG	3.8
1	B	407	THR	3.7
1	B	364	GLU	3.7
1	D	378	LEU	3.6
1	B	333	PHE	3.6
1	B	337	GLY	3.6
1	B	365	THR	3.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	419	PRO	3.5
1	D	363	LEU	3.5
1	B	338	VAL	3.5
1	B	396	ILE	3.3
1	B	402	ASP	3.2
1	B	405	LEU	3.2
1	D	335	ILE	3.2
1	B	352	PHE	3.2
1	C	0	GLY	3.2
1	B	410	VAL	3.1
1	B	336	LYS	3.0
1	B	392	ILE	2.9
1	D	334	ILE	2.9
1	B	430	ASP	2.8
1	B	404	ILE	2.8
1	A	434	ASN	2.7
1	B	411	LYS	2.7
1	C	426	VAL	2.7
1	A	422	GLU	2.7
1	B	378	LEU	2.6
1	B	97	THR	2.6
1	D	431	LEU	2.6
1	B	376	VAL	2.5
1	D	405	LEU	2.5
1	A	350	VAL	2.5
1	D	392	ILE	2.4
1	C	422	GLU	2.3
1	B	351[A]	GLN	2.3
1	D	356	GLY	2.2
1	D	382	SER	2.2
1	B	418	LEU	2.2
1	B	380	ASP	2.2
1	A	376	VAL	2.1
1	C	59	ILE	2.1
1	C	418	LEU	2.1
1	A	383	THR	2.0
1	B	421	SER	2.0

## 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 ⓘ

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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
8	EDO	A	516	4/4	0.19	31.92	72,73,74,75	0
8	EDO	A	513	4/4	0.22	11.56	68,71,73,76	0
8	EDO	A	512	4/4	0.26	7.66	53,54,54,55	0
5	COA	A	504[A]	31/48	0.31	6.41	39,55,68,69	31
8	EDO	C	510	4/4	0.20	6.32	61,63,64,66	0
5	COA	C	503[A]	31/48	0.28	5.61	31,43,49,49	31
7	SO4	C	507[B]	5/5	0.15	4.94	55,56,57,57	5
8	EDO	D	505	4/4	0.18	4.12	73,76,78,79	0
8	EDO	C	511	4/4	0.14	3.65	50,52,52,52	0
8	EDO	A	514	4/4	0.18	3.12	55,55,57,61	0
8	EDO	B	505	4/4	0.19	3.03	78,78,79,79	0
7	SO4	B	504	5/5	0.15	2.96	97,99,100,100	0
8	EDO	C	509	4/4	0.13	2.55	67,68,68,69	0
7	SO4	A	509[B]	5/5	0.16	2.30	31,34,35,35	5
8	EDO	A	515	4/4	0.14	2.10	35,40,40,40	0
4	COD	A	503[B]	44/44	0.17	1.78	24,36,50,55	44
4	COD	B	503[B]	44/44	0.22	1.26	30,64,92,92	44
4	COD	D	503[B]	44/44	0.17	1.17	21,36,53,56	44
8	EDO	D	506	4/4	0.13	0.62	52,52,54,58	0
7	SO4	D	504	5/5	0.10	-0.02	81,81,83,85	0
7	SO4	C	508	5/5	0.13	-0.21	113,113,113,113	0
3	AMP	B	502[A]	23/23	0.11	-0.37	21,26,31,34	23
3	AMP	C	502	23/23	0.12	-0.53	39,44,46,51	0
3	AMP	D	502[A]	23/23	0.09	-0.83	51,61,63,65	23
3	AMP	A	502[A]	23/23	0.09	-0.87	15,29,31,32	23
7	SO4	A	510	5/5	0.07	-0.98	47,51,54,55	0
2	ZN	A	501	1/1	0.08	-1.17	43,43,43,43	0
6	K	A	505	1/1	0.07	-1.24	39,39,39,39	0
7	SO4	A	511	5/5	0.10	-1.26	126,127,127,127	0
2	ZN	D	501	1/1	0.09	-1.39	48,48,48,48	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
6	K	C	505	1/1	0.08	-1.54	38,38,38,38	0
2	ZN	C	501	1/1	0.06	-2.09	56,56,56,56	0
6	K	A	506	1/1	0.07	-2.11	44,44,44,44	0
6	K	C	506	1/1	0.05	-2.31	38,38,38,38	0
2	ZN	B	501	1/1	0.07	-2.45	49,49,49,49	0
6	K	A	508	1/1	0.04	-2.57	73,73,73,73	0
6	K	A	507	1/1	0.04	-2.64	31,31,31,31	0
6	K	C	504	1/1	0.07	-7.92	45,45,45,45	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.