



Full wwPDB X-ray Structure Validation Report ⓘ

Sep 4, 2017 – 06:45 AM EDT

PDB ID : 5N1U
Title : Structure of xEco2 acetyltransferase domain
Authors : Chao, W.C.H.; Wade, B.O.; Singleton, M.R.
Deposited on : unknown
Resolution : 2.98 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 : 1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20029824
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
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-20029824

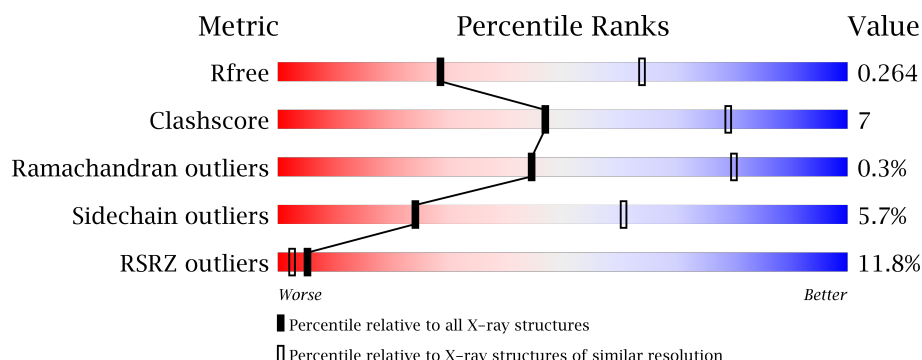
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.98 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}	100719	2168 (3.00-2.96)
Clashscore	112137	2535 (3.00-2.96)
Ramachandran outliers	110173	2451 (3.00-2.96)
Sidechain outliers	110143	2454 (3.00-2.96)
RSRZ outliers	101464	2192 (3.00-2.96)

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	183	
1	B	183	

2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 5260 atoms, of which 2535 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.

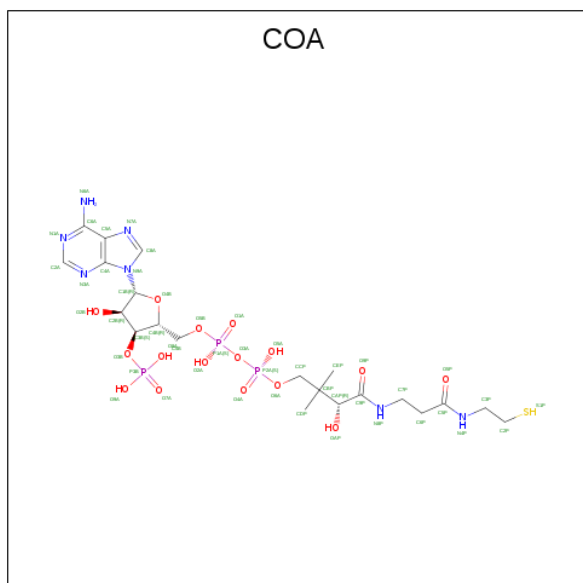
- Molecule 1 is a protein called XEco2.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	170	Total	C	H	N	O	S	0	0	0
			2566	850	1243	223	240	10			
1	B	170	Total	C	H	N	O	S	0	0	0
			2534	838	1228	220	237	11			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	520	GLY	-	expression tag	UNP A8QZK6
A	521	SER	-	expression tag	UNP A8QZK6
A	522	MET	-	expression tag	UNP A8QZK6
B	520	GLY	-	expression tag	UNP A8QZK6
B	521	SER	-	expression tag	UNP A8QZK6
B	522	MET	-	expression tag	UNP A8QZK6

- Molecule 2 is COENZYME A (three-letter code: COA) (formula: $C_{21}H_{36}N_7O_{16}P_3S$).

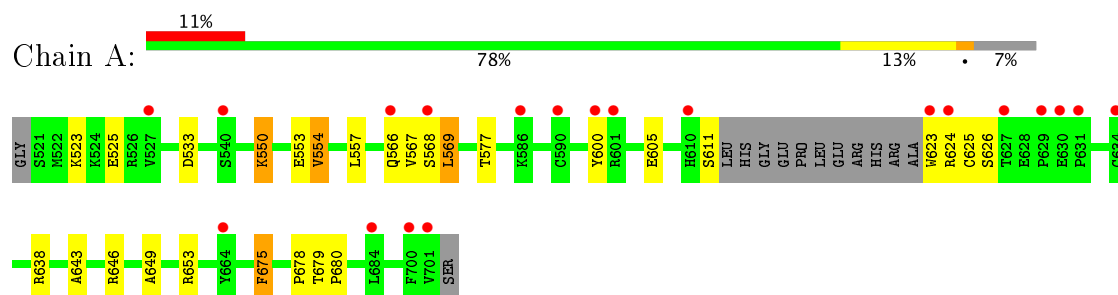


Mol	Chain	Residues	Atoms							ZeroOcc	AltConf
2	A	1	Total	C	H	N	O	P	S	0	0
			80	21	32	7	16	3	1		
2	B	1	Total	C	H	N	O	P	S	0	0
			80	21	32	7	16	3	1		

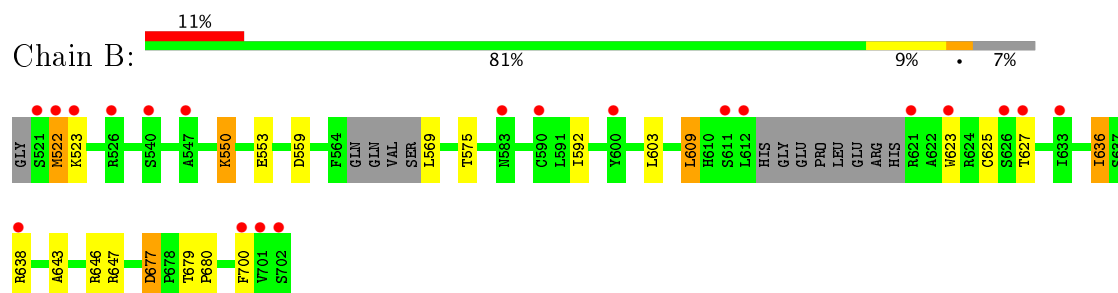
3 Residue-property plots [i](#)

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: XEco2



• Molecule 1: XEco2



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	58.48 Å 66.29 Å 109.76 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	28.83 – 2.98 28.83 – 2.98	Depositor EDS
% Data completeness (in resolution range)	98.8 (28.83-2.98) 98.8 (28.83-2.98)	Depositor EDS
R_{merge}	0.30	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.21 (at 3.00 Å)	Xtriage
Refinement program	PHENIX 1.10.1_2155	Depositor
R, R_{free}	0.223 , 0.268 0.223 , 0.264	Depositor DCC
R_{free} test set	907 reflections (9.98%)	DCC
Wilson B-factor (Å ²)	83.5	Xtriage
Anisotropy	0.079	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 52.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	5260	wwPDB-VP
Average B, all atoms (Å ²)	78.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 14.52% 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, 0.333 respectively for untwinned datasets, and 0.375, 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	0.41	0/1354	0.54	0/1837
1	B	0.34	0/1334	0.53	1/1807 (0.1%)
All	All	0.37	0/2688	0.54	1/3644 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	677	ASP	CB-CG-OD2	6.39	124.05	118.30

There are no chirality outliers.

There are no planarity outliers.

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	1323	1243	1301	21	3
1	B	1306	1228	1290	15	3
2	A	48	32	32	2	0
2	B	48	32	32	5	0
All	All	2725	2535	2655	38	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 7.

All (38) 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)
2:B:801:COA:C1B	2:B:801:COA:O4B	1.65	1.28
2:A:801:COA:O4B	2:A:801:COA:C1B	1.65	1.12
1:A:567:VAL:HG21	1:A:638:ARG:HH22	1.25	0.99
1:A:567:VAL:HG21	1:A:638:ARG:NH2	1.89	0.87
1:A:567:VAL:HG23	1:A:568:SER:N	1.96	0.78
1:A:550:LYS:HD3	1:B:609:LEU:HA	1.75	0.68
1:A:567:VAL:CG2	1:A:568:SER:N	2.56	0.68
1:A:533:ASP:OD1	1:A:653:ARG:NH2	2.29	0.65
1:B:623:TRP:HZ2	1:B:677:ASP:OD1	1.80	0.64
1:A:569:LEU:HD22	1:A:569:LEU:H	1.63	0.62
1:B:559:ASP:OD1	1:B:638:ARG:NH1	2.33	0.61
1:B:592:ILE:HD12	1:B:638:ARG:HD2	1.81	0.61
1:B:550:LYS:NZ	1:B:553:GLU:OE1	2.26	0.59
1:A:569:LEU:HD22	1:A:569:LEU:N	2.18	0.57
2:B:801:COA:H52A	2:B:801:COA:H8A	1.86	0.56
1:A:600:TYR:HB2	1:A:625:CYS:SG	2.50	0.52
1:B:647:ARG:N	2:B:801:COA:O4A	2.38	0.52
1:A:550:LYS:O	1:A:553:GLU:HG2	2.11	0.51
1:A:567:VAL:CG2	1:A:568:SER:H	2.24	0.50
1:B:623:TRP:CZ2	1:B:677:ASP:OD1	2.64	0.49
1:A:623:TRP:CZ2	1:A:678:PRO:HD2	2.48	0.49
1:B:636:ILE:HD12	1:B:636:ILE:N	2.28	0.49
1:A:550:LYS:O	1:A:554:VAL:HG12	2.13	0.48
1:A:643:ALA:HA	1:A:646:ARG:NE	2.28	0.48
1:A:605:GLU:HB2	1:A:623:TRP:HA	1.95	0.48
1:A:649:ALA:O	1:A:653:ARG:HD3	2.13	0.48
2:A:801:COA:O5B	2:A:801:COA:H8A	2.13	0.47
1:A:675:PHE:N	1:A:675:PHE:CD1	2.82	0.47
1:B:679:THR:HB	1:B:680:PRO:HD2	1.97	0.47
1:B:647:ARG:NH2	2:B:801:COA:O7A	2.45	0.46
1:A:679:THR:HB	1:A:680:PRO:HD2	1.97	0.45
1:B:522:MET:HE2	1:B:522:MET:HB3	1.84	0.44
1:A:605:GLU:OE1	1:A:624:ARG:HG2	2.19	0.42
1:B:575:THR:CG2	1:B:592:ILE:HG23	2.50	0.41
1:B:646:ARG:HD3	2:B:801:COA:CEP	2.50	0.41
1:A:643:ALA:HB3	1:B:603:LEU:HD13	2.02	0.40
1:B:643:ALA:HA	1:B:646:ARG:NE	2.36	0.40
1:A:569:LEU:HD13	1:A:638:ARG:NH2	2.36	0.40

All (3) symmetry-related close contacts are listed below. The label for Atom-2 includes the sym-

metry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:566:GLN:NE2	1:B:627:THR:HA[3_545]	0.94	0.66
1:A:566:GLN:NE2	1:B:627:THR:CA[3_545]	1.77	0.43
1:A:567:VAL:O	1:B:700:PHE:O[3_545]	2.02	0.18

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	166/183 (91%)	156 (94%)	10 (6%)	0	100	100
1	B	164/183 (90%)	153 (93%)	10 (6%)	1 (1%)	28	68
All	All	330/366 (90%)	309 (94%)	20 (6%)	1 (0%)	44	80

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	625	CYS

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	142/162 (88%)	132 (93%)	10 (7%)	18	51
1	B	140/162 (86%)	134 (96%)	6 (4%)	33	71
All	All	282/324 (87%)	266 (94%)	16 (6%)	24	61

All (16) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	523	LYS
1	A	525	GLU
1	A	550	LYS
1	A	554	VAL
1	A	557	LEU
1	A	569	LEU
1	A	577	THR
1	A	611	SER
1	A	626	SER
1	A	675	PHE
1	B	522	MET
1	B	523	LYS
1	B	550	LYS
1	B	569	LEU
1	B	609	LEU
1	B	636	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	610	HIS

5.3.3 RNA ⓘ

There are no RNA molecules 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 ⓘ

2 ligands are 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	A	801	-	43,50,50	4.10	12 (27%)	48,75,75	1.95	9 (18%)
2	COA	B	801	-	43,50,50	4.11	12 (27%)	48,75,75	2.01	7 (14%)

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	A	801	-	-	0/44/64/64	0/3/3/3
2	COA	B	801	-	-	0/44/64/64	0/3/3/3

All (24) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	801	COA	C2B-C1B	-13.50	1.32	1.53
2	A	801	COA	C2B-C1B	-13.28	1.32	1.53
2	B	801	COA	O4B-C4B	-6.04	1.31	1.45
2	A	801	COA	O4B-C4B	-5.96	1.31	1.45
2	B	801	COA	C5A-C4A	-2.91	1.33	1.40
2	A	801	COA	C5A-C4A	-2.82	1.34	1.40
2	A	801	COA	OAP-CAP	-2.61	1.37	1.42
2	A	801	COA	O3B-C3B	-2.52	1.34	1.44
2	B	801	COA	O3B-C3B	-2.51	1.34	1.44
2	B	801	COA	OAP-CAP	-2.45	1.37	1.42
2	B	801	COA	C6A-N6A	2.16	1.42	1.34
2	A	801	COA	C6A-N6A	2.26	1.43	1.34
2	A	801	COA	C7P-N8P	2.79	1.52	1.46
2	B	801	COA	C7P-N8P	2.82	1.52	1.46
2	B	801	COA	C2A-N3A	3.09	1.37	1.32
2	A	801	COA	C2A-N3A	3.13	1.37	1.32
2	A	801	COA	C5P-N4P	3.37	1.41	1.33
2	B	801	COA	C5P-N4P	3.73	1.42	1.33
2	A	801	COA	P3B-O3B	5.97	1.70	1.59

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	801	COA	P3B-O3B	6.01	1.70	1.59
2	A	801	COA	C9P-N8P	9.23	1.52	1.33
2	B	801	COA	C9P-N8P	9.25	1.52	1.33
2	B	801	COA	O4B-C1B	17.18	1.65	1.41
2	A	801	COA	O4B-C1B	17.35	1.65	1.41

All (16) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	801	COA	N3A-C2A-N1A	-9.57	120.52	128.86
2	A	801	COA	N3A-C2A-N1A	-9.28	120.77	128.86
2	A	801	COA	N6A-C6A-N1A	-4.13	110.57	118.77
2	B	801	COA	N6A-C6A-N1A	-4.12	110.61	118.77
2	B	801	COA	C7P-C6P-C5P	-3.91	105.93	112.22
2	A	801	COA	C7P-C6P-C5P	-2.90	107.55	112.22
2	A	801	COA	C3P-N4P-C5P	-2.37	118.29	122.84
2	B	801	COA	C4B-O4B-C1B	-2.29	107.33	109.77
2	A	801	COA	O5P-C5P-N4P	-2.27	118.64	122.97
2	B	801	COA	C6P-C7P-N8P	-2.26	107.19	111.87
2	A	801	COA	C4B-O4B-C1B	-2.08	107.55	109.77
2	A	801	COA	C6P-C7P-N8P	-2.06	107.61	111.87
2	A	801	COA	C6P-C5P-N4P	2.52	120.84	116.49
2	B	801	COA	C6P-C5P-N4P	2.87	121.45	116.49
2	B	801	COA	C5A-C6A-N6A	4.43	129.50	120.47
2	A	801	COA	C5A-C6A-N6A	4.56	129.75	120.47

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	801	COA	2	0
2	B	801	COA	5	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 ⓘ

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	170/183 (92%)	0.63	20 (11%) 5 2	45, 68, 101, 248	0
1	B	170/183 (92%)	0.70	20 (11%) 5 2	49, 70, 100, 173	0
All	All	340/366 (92%)	0.67	40 (11%) 5 2	45, 69, 102, 248	0

All (40) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	701	VAL	5.3
1	A	623	TRP	4.6
1	B	702	SER	4.5
1	B	612	LEU	4.5
1	B	626	SER	4.2
1	A	700	PHE	4.0
1	B	522	MET	3.7
1	B	701	VAL	3.7
1	A	664	TYR	3.3
1	A	627	THR	3.2
1	B	633	ILE	3.1
1	A	566	GLN	2.9
1	B	621	ARG	2.9
1	A	568	SER	2.8
1	B	523	LYS	2.8
1	B	590	CYS	2.7
1	B	638	ARG	2.7
1	A	586	LYS	2.7
1	B	700	PHE	2.6
1	A	630	GLU	2.5
1	B	600	TYR	2.5
1	B	623	TRP	2.5
1	A	631	PRO	2.4
1	B	627	THR	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	601	ARG	2.4
1	A	527	VAL	2.4
1	B	547	ALA	2.3
1	A	629	PRO	2.3
1	A	540	SER	2.2
1	B	611	SER	2.2
1	B	583	ASN	2.1
1	A	600	TYR	2.1
1	A	624	ARG	2.1
1	B	540	SER	2.1
1	A	590	CYS	2.1
1	B	521	SER	2.1
1	B	526	ARG	2.1
1	A	634	CYS	2.1
1	A	684	LEU	2.0
1	A	610	HIS	2.0

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, 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	B	801	48/48	0.92	0.20	-0.52	44,65,80,83	0
2	COA	A	801	48/48	0.92	0.20	-0.55	43,63,83,88	0

6.5 Other polymers ⓘ

There are no such residues in this entry.