



# Full wwPDB X-ray Structure Validation Report

Mar 12, 2014 – 05:20 PM GMT

PDB ID : 4N4B  
Title : Crystal Structure of the alpha-L-arabinofuranosidasePaAbf62A from *Podospira anserina*  
Authors : Siguier, B.; Dumon, C.; Mourey, L.; Tranier, S.  
Deposited on : 2013-10-08  
Resolution : 1.44 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>

---

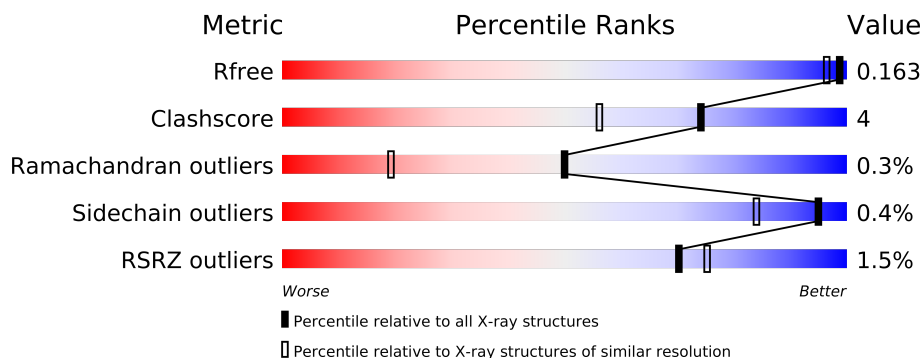
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.16 November 2013  
Xtriage (Phenix) : dev-1323  
EDS : trunk22714  
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) : trunk22714

# 1 Overall quality at a glance


The reported resolution of this entry is 1.44 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	1819 (1.48-1.40)
Clashscore	79885	2050 (1.48-1.40)
Ramachandran outliers	78287	2002 (1.48-1.40)
Sidechain outliers	78261	2001 (1.48-1.40)
RSRZ outliers	66119	1819 (1.48-1.40)

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	360	

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	1PE	A	404	-	X

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 3022 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 GH62 arabinofuranosidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	319	Total	C	N	O	S	0	9	0
			2570	1644	433	481	12			

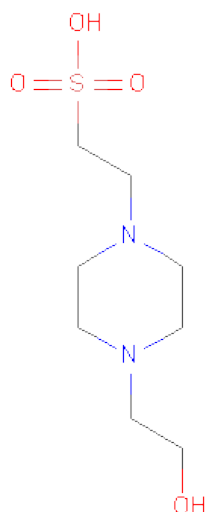
There are 23 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	338	GLY	-	EXPRESSION TAG	UNP E2GHW5
A	339	LEU	-	EXPRESSION TAG	UNP E2GHW5
A	340	GLU	-	EXPRESSION TAG	UNP E2GHW5
A	341	GLN	-	EXPRESSION TAG	UNP E2GHW5
A	342	LYS	-	EXPRESSION TAG	UNP E2GHW5
A	343	LEU	-	EXPRESSION TAG	UNP E2GHW5
A	344	ILE	-	EXPRESSION TAG	UNP E2GHW5
A	345	SER	-	EXPRESSION TAG	UNP E2GHW5
A	346	GLU	-	EXPRESSION TAG	UNP E2GHW5
A	347	GLU	-	EXPRESSION TAG	UNP E2GHW5
A	348	ASP	-	EXPRESSION TAG	UNP E2GHW5
A	349	LEU	-	EXPRESSION TAG	UNP E2GHW5
A	350	ASN	-	EXPRESSION TAG	UNP E2GHW5
A	351	SER	-	EXPRESSION TAG	UNP E2GHW5
A	352	ALA	-	EXPRESSION TAG	UNP E2GHW5
A	353	VAL	-	EXPRESSION TAG	UNP E2GHW5
A	354	ASP	-	EXPRESSION TAG	UNP E2GHW5
A	355	HIS	-	EXPRESSION TAG	UNP E2GHW5
A	356	HIS	-	EXPRESSION TAG	UNP E2GHW5
A	357	HIS	-	EXPRESSION TAG	UNP E2GHW5
A	358	HIS	-	EXPRESSION TAG	UNP E2GHW5
A	359	HIS	-	EXPRESSION TAG	UNP E2GHW5
A	360	HIS	-	EXPRESSION TAG	UNP E2GHW5

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

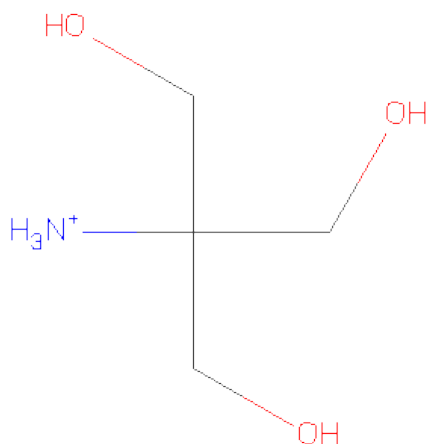
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Ca	0	0
			1	1		

- Molecule 3 is 4-(2-HYDROXYETHYL)-1-PIPERAZINEETHANESULFONIC ACID (three-letter code: EPE) (formula:  $C_8H_{18}N_2O_4S$ ).



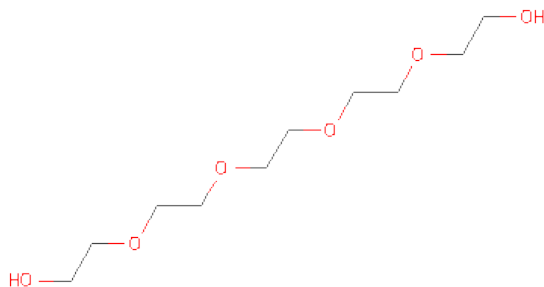
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	S	0	0
			15	8	2	4	1		

- Molecule 4 is 2-AMINO-2-HYDROXYMETHYL-PROPANE-1,3-DIOL (three-letter code: TRS) (formula:  $C_4H_{12}NO_3$ ).



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

- Molecule 5 is PENTAETHYLENE GLYCOL (three-letter code: 1PE) (formula:  $C_{10}H_{22}O_6$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			13	8	5		

- Molecule 6 is water.

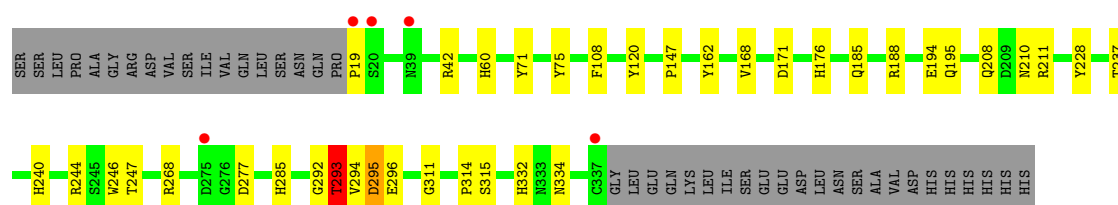
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	415	Total	O	0	0
			415	415		

### 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: GH62 arabinofuranosidase

Chain A: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	102.41Å 66.83Å 60.57Å 90.00° 117.39° 90.00°	Depositor
Resolution (Å)	44.61 – 1.44 44.61 – 1.44	Depositor EDS
% Data completeness (in resolution range)	92.5 (44.61-1.44) 92.5 (44.61-1.44)	Depositor EDS
$R_{merge}$	0.04	Depositor
$R_{sym}$	0.04	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	8.04 (at 1.44Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, $R_{free}$	0.128 , 0.152 0.142 , 0.163	Depositor DCC
$R_{free}$ test set	3039 reflections (5.27%)	DCC
Wilson B-factor (Å <sup>2</sup> )	12.5	Xtriage
Anisotropy	0.078	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.40 , 39.8	EDS
Estimated twinning fraction	0.011 for -1/2*h+1/2*k+l,1/2*h-1/2*k+l,1/2*h+1/2*k 0.017 for -1/2*h-1/2*k+l,-1/2*h-1/2*k-l,1/2*h-1/2*k	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 60690 reflections	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	3022	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	15.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.62% 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: CA, EPE, TRS, 1PE

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	1.37	12/2680 (0.4%)	1.26	15/3666 (0.4%)

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	2

All (12) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	293[A]	THR	N-CA	7.94	1.62	1.46
1	A	293[B]	THR	N-CA	7.94	1.62	1.46
1	A	311	GLY	C-O	-5.98	1.14	1.23
1	A	315[A]	SER	N-CA	5.79	1.57	1.46
1	A	315[B]	SER	N-CA	5.79	1.57	1.46
1	A	296	GLU	CD-OE1	-5.71	1.19	1.25
1	A	314	PRO	N-CD	5.62	1.55	1.47
1	A	294	VAL	C-O	-5.50	1.12	1.23
1	A	194	GLU	CG-CD	5.20	1.59	1.51
1	A	71	TYR	CE2-CZ	-5.19	1.31	1.38
1	A	246	TRP	CD1-NE1	-5.07	1.29	1.38
1	A	194	GLU	CD-OE1	5.03	1.31	1.25

All (15) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	171	ASP	CB-CG-OD1	5.94	123.65	118.30

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	295	ASP	CB-CG-OD1	5.74	123.47	118.30
1	A	268	ARG	NE-CZ-NH2	-5.66	117.47	120.30
1	A	211[A]	ARG	NE-CZ-NH1	5.60	123.10	120.30
1	A	211[B]	ARG	NE-CZ-NH1	5.60	123.10	120.30
1	A	162	TYR	CB-CG-CD1	-5.56	117.66	121.00
1	A	120	TYR	CG-CD2-CE2	5.33	125.57	121.30
1	A	228	TYR	CB-CG-CD2	-5.32	117.81	121.00
1	A	108	PHE	CB-CG-CD1	5.16	124.41	120.80
1	A	237	THR	CA-CB-CG2	-5.14	105.20	112.40
1	A	244	ARG	NE-CZ-NH2	-5.13	117.73	120.30
1	A	42	ARG	NE-CZ-NH1	-5.05	117.77	120.30
1	A	171	ASP	CB-CG-OD2	-5.04	113.76	118.30
1	A	120	TYR	CD1-CE1-CZ	5.01	124.31	119.80
1	A	292	GLY	C-N-CA	-5.00	109.20	121.70

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	293[A]	THR	Mainchain
1	A	293[B]	THR	Mainchain

## 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	2570	0	2414	20	0
2	A	1	0	0	0	0
3	A	15	0	18	0	0
4	A	8	0	12	0	0
5	A	13	0	17	1	0
6	A	415	0	0	8	0
All	All	3022	0	2461	21	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 4.

All (21) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:176:HIS:HD2	1:A:188:ARG:HH11	1.33	0.73
1:A:240:HIS:HD2	6:A:963:HOH:O	1.81	0.63
1:A:60:HIS:HD2	1:A:75:TYR:OH	1.81	0.62
1:A:147:PRO:O	5:A:404:1PE:H141	2.01	0.61
1:A:185:GLN:HE21	1:A:208:GLN:HE21	1.51	0.58
1:A:176:HIS:CD2	1:A:188:ARG:HH11	2.19	0.55
1:A:185:GLN:NE2	1:A:208:GLN:HE21	2.05	0.55
1:A:332:HIS:HD2	6:A:842:HOH:O	1.91	0.53
1:A:195[A]:GLN:NE2	6:A:1156:HOH:O	2.27	0.50
1:A:60:HIS:CD2	1:A:75:TYR:OH	2.65	0.49
1:A:60:HIS:HE1	6:A:857:HOH:O	1.97	0.46
1:A:210:ASN:ND2	6:A:826:HOH:O	2.49	0.45
1:A:176:HIS:HE1	6:A:928:HOH:O	1.99	0.45
1:A:240:HIS:HE1	6:A:1027:HOH:O	2.00	0.44
1:A:293[A]:THR:CG2	1:A:295:ASP:CG	2.87	0.43
1:A:19:PRO:HG2	1:A:247:THR:HG21	2.01	0.42
1:A:168:VAL:HA	1:A:176:HIS:O	2.19	0.42
1:A:332:HIS:CE1	1:A:334:ASN:OD1	2.73	0.41
1:A:293[A]:THR:HG22	1:A:295:ASP:H	1.86	0.41
1:A:293[B]:THR:H	1:A:293[B]:THR:HG23	1.75	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	326/360 (91%)	313 (96%)	12 (4%)	1 (0%)	50 19

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	285	HIS

### 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	273/301 (91%)	272 (100%)	1 (0%)	95 84

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	277	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	60	HIS
1	A	134	ASN
1	A	157	ASN
1	A	176	HIS
1	A	185	GLN
1	A	210	ASN
1	A	212	ASN
1	A	240	HIS
1	A	332	HIS

### 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 4 ligands modelled in this entry, 1 is monoatomic - leaving 3 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	EPE	A	402	-	15,15,15	2.04	4 (26%)	20,20,20	3.11	10 (50%)
4	TRS	A	403	-	7,7,7	1.59	1 (14%)	9,9,9	2.08	4 (44%)
5	1PE	A	404	-	12,12,15	0.83	0	11,11,14	0.90	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	EPE	A	402	-	-	0/9/19/19	0/1/1/1
4	TRS	A	403	-	-	0/9/9/9	0/0/0/0
5	1PE	A	404	-	-	0/10/10/13	0/0/0/0

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	402	EPE	C10-S	-5.31	1.69	1.77
3	A	402	EPE	O2S-S	3.86	1.57	1.45
4	A	403	TRS	O1-C1	2.96	1.52	1.42
3	A	402	EPE	O1S-S	2.53	1.53	1.45
3	A	402	EPE	O3S-S	2.35	1.52	1.46

All (14) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	402	EPE	C6-N1-C2	5.83	122.94	108.88
3	A	402	EPE	C7-N4-C5	5.08	124.01	111.28
3	A	402	EPE	C9-N1-C6	4.99	123.77	111.28
3	A	402	EPE	C5-N4-C3	4.80	120.46	108.88

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	402	EPE	O1S-S-C10	4.12	110.46	106.79
3	A	402	EPE	O3S-S-O1S	3.88	120.32	111.69
3	A	402	EPE	C6-C5-N4	3.83	118.07	110.61
3	A	402	EPE	C5-C6-N1	3.54	117.51	110.61
4	A	403	TRS	O3-C3-C	-3.47	104.25	111.28
4	A	403	TRS	C2-C-N	-3.37	102.00	108.10
3	A	402	EPE	O2S-S-O1S	-2.93	103.26	113.45
3	A	402	EPE	O3S-S-C10	-2.73	102.64	105.93
4	A	403	TRS	C1-C-N	2.37	112.40	108.10
4	A	403	TRS	C2-C-C1	2.35	115.87	110.78

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	319/360 (88%)	-0.23	5 (1%) 68 73	7, 12, 26, 44	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	20	SER	3.7
1	A	275	ASP	2.9
1	A	337	CYS	2.8
1	A	39	ASN	2.7
1	A	19	PRO	2.2

### 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
5	1PE	A	404	13/16	0.11	6.07	25,29,36,38	13

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	EPE	A	402	15/15	0.15	1.56	25,29,42,52	0
4	TRS	A	403	8/8	0.09	0.20	11,11,12,13	0
2	CA	A	401	1/1	0.02	-5.63	11,11,11,11	0

## 6.5 Other polymers

There are no such residues in this entry.