



wwPDB X-ray Structure Validation Summary Report

Feb 27, 2014 – 02:45 PM GMT

PDB ID : 4APL
Title : Crystal Structure of AMA1 from Neospora caninum
Authors : Tonkin, M.L.; Crawford, J.; Lebrun, M.L.; Boulanger, M.J.
Deposited on : 2012-04-04
Resolution : 2.90 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

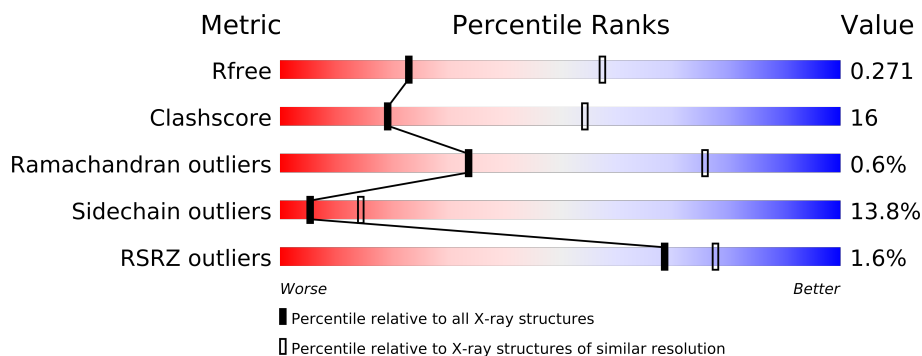
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
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)	:	stable22683

1 Overall quality at a glance

The reported resolution of this entry is 2.90 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	1053 (2.90-2.90)
Clashscore	79885	1326 (2.90-2.90)
Ramachandran outliers	78287	1290 (2.90-2.90)
Sidechain outliers	78261	1292 (2.90-2.90)
RSRZ outliers	66119	1054 (2.90-2.90)

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. 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	431	
1	B	431	
1	D	431	
1	E	431	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 12647 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 APICAL MEMBRANE ANTIGEN 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	398	Total	C	N	O	S	0	1	0
			3155	1975	541	618	21			
1	B	394	Total	C	N	O	S	0	1	0
			3123	1958	533	611	21			
1	D	395	Total	C	N	O	S	0	1	0
			3133	1963	537	612	21			
1	E	390	Total	C	N	O	S	0	1	0
			3094	1940	529	604	21			

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	54	GLY	-	EXPRESSION TAG	UNP A2A114
A	55	SER	-	EXPRESSION TAG	UNP A2A114
A	56	ALA	-	EXPRESSION TAG	UNP A2A114
A	57	MET	-	EXPRESSION TAG	UNP A2A114
A	58	GLY	-	EXPRESSION TAG	UNP A2A114
A	479	ALA	-	EXPRESSION TAG	UNP A2A114
A	480	ALA	-	EXPRESSION TAG	UNP A2A114
A	481	LEU	-	EXPRESSION TAG	UNP A2A114
A	482	VAL	-	EXPRESSION TAG	UNP A2A114
A	483	PRO	-	EXPRESSION TAG	UNP A2A114
A	484	ARG	-	EXPRESSION TAG	UNP A2A114
B	54	GLY	-	EXPRESSION TAG	UNP A2A114
B	55	SER	-	EXPRESSION TAG	UNP A2A114
B	56	ALA	-	EXPRESSION TAG	UNP A2A114
B	57	MET	-	EXPRESSION TAG	UNP A2A114
B	58	GLY	-	EXPRESSION TAG	UNP A2A114
B	479	ALA	-	EXPRESSION TAG	UNP A2A114
B	480	ALA	-	EXPRESSION TAG	UNP A2A114
B	481	LEU	-	EXPRESSION TAG	UNP A2A114
B	482	VAL	-	EXPRESSION TAG	UNP A2A114
B	483	PRO	-	EXPRESSION TAG	UNP A2A114

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	484	ARG	-	EXPRESSION TAG	UNP A2A114
D	54	GLY	-	EXPRESSION TAG	UNP A2A114
D	55	SER	-	EXPRESSION TAG	UNP A2A114
D	56	ALA	-	EXPRESSION TAG	UNP A2A114
D	57	MET	-	EXPRESSION TAG	UNP A2A114
D	58	GLY	-	EXPRESSION TAG	UNP A2A114
D	479	ALA	-	EXPRESSION TAG	UNP A2A114
D	480	ALA	-	EXPRESSION TAG	UNP A2A114
D	481	LEU	-	EXPRESSION TAG	UNP A2A114
D	482	VAL	-	EXPRESSION TAG	UNP A2A114
D	483	PRO	-	EXPRESSION TAG	UNP A2A114
D	484	ARG	-	EXPRESSION TAG	UNP A2A114
E	54	GLY	-	EXPRESSION TAG	UNP A2A114
E	55	SER	-	EXPRESSION TAG	UNP A2A114
E	56	ALA	-	EXPRESSION TAG	UNP A2A114
E	57	MET	-	EXPRESSION TAG	UNP A2A114
E	58	GLY	-	EXPRESSION TAG	UNP A2A114
E	479	ALA	-	EXPRESSION TAG	UNP A2A114
E	480	ALA	-	EXPRESSION TAG	UNP A2A114
E	481	LEU	-	EXPRESSION TAG	UNP A2A114
E	482	VAL	-	EXPRESSION TAG	UNP A2A114
E	483	PRO	-	EXPRESSION TAG	UNP A2A114
E	484	ARG	-	EXPRESSION TAG	UNP A2A114

- Molecule 2 is water.

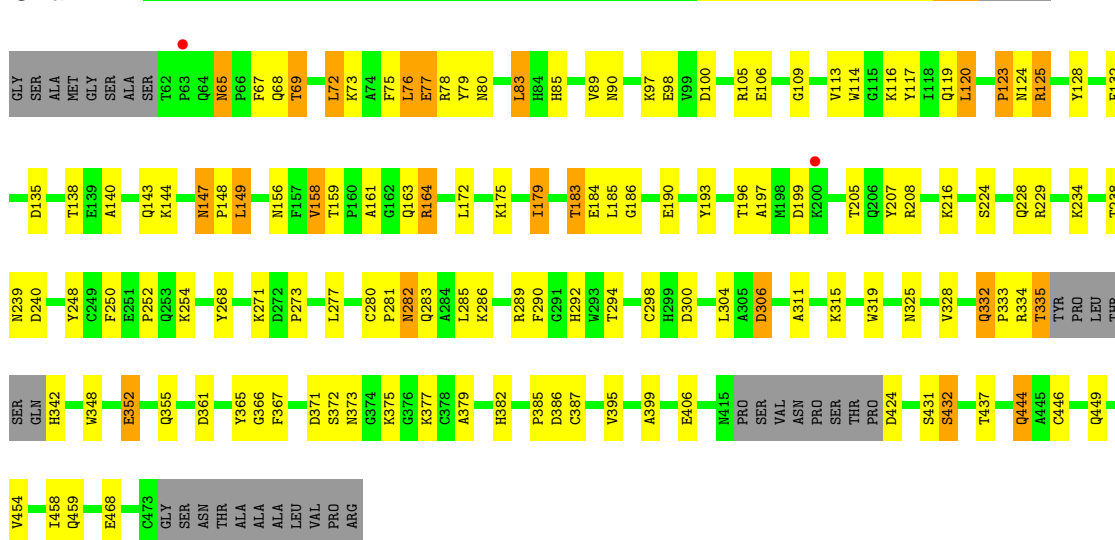
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	45	Total O 45 45	0	0
2	B	33	Total O 33 33	0	0
2	D	45	Total O 45 45	0	0
2	E	19	Total O 19 19	0	0

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.

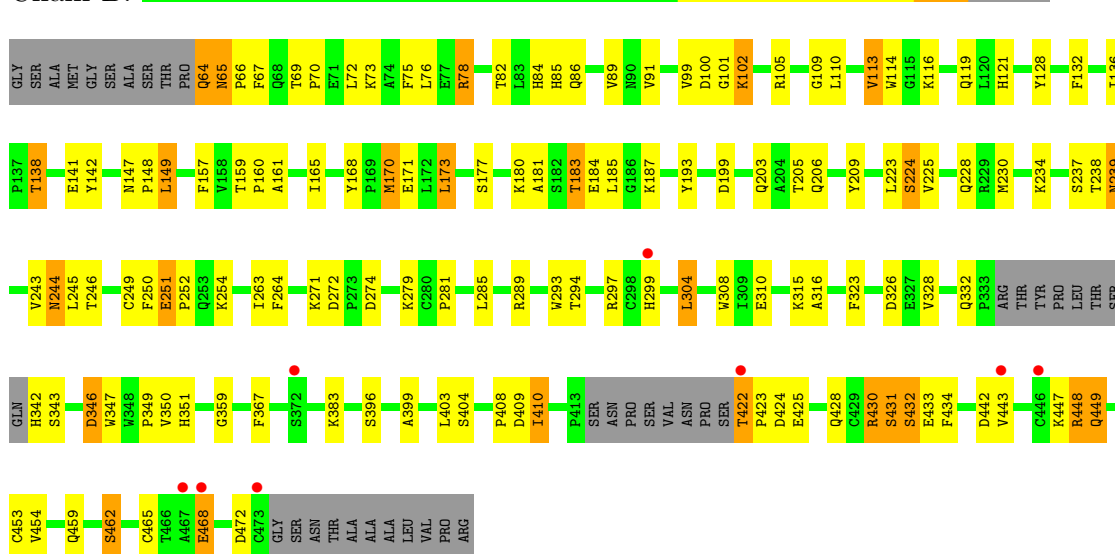
• Molecule 1: APICAL MEMBRANE ANTIGEN 1

Chain A:

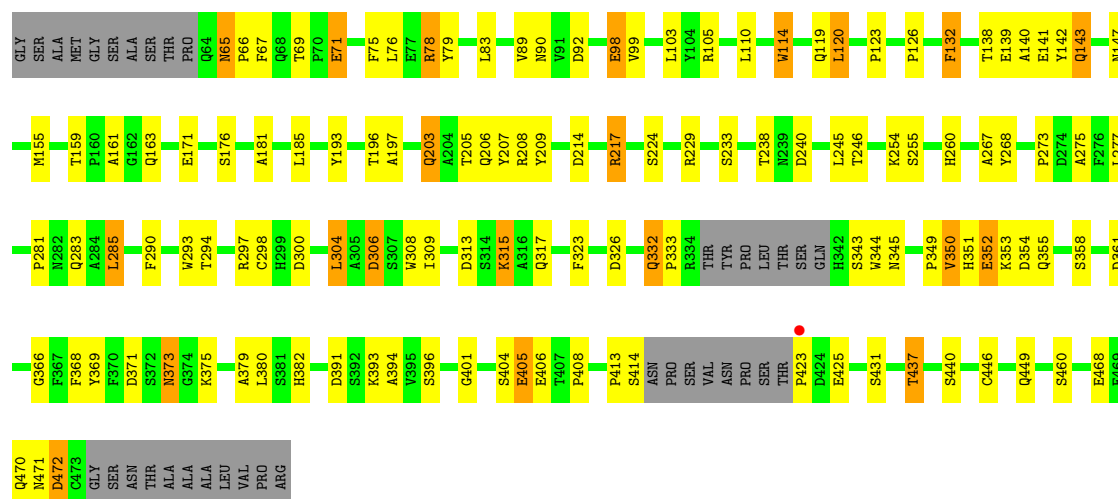


• Molecule 1: APICAL MEMBRANE ANTIGEN 1

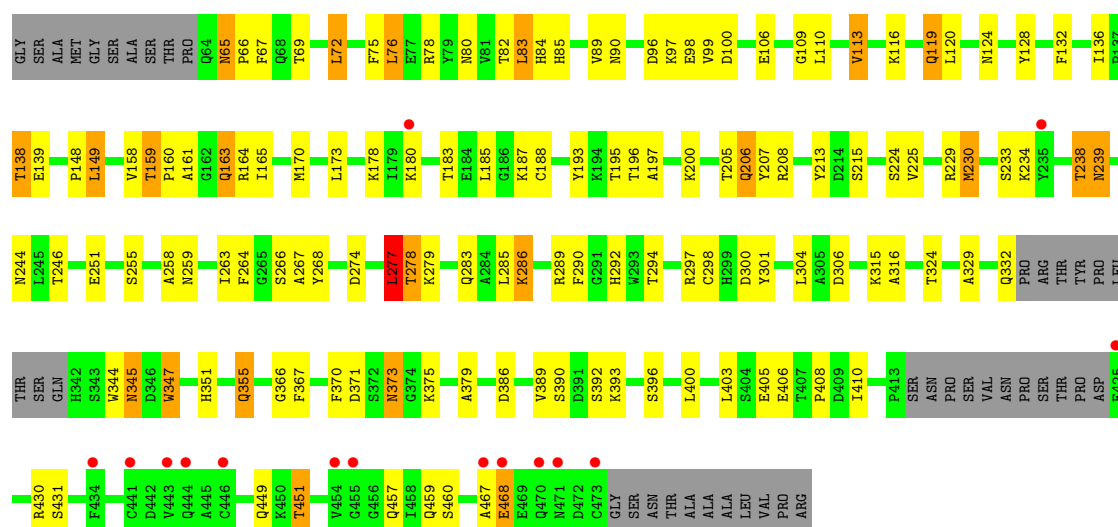
Chain B:



• Molecule 1: APICAL MEMBRANE ANTIGEN 1

Chain D: 

● Molecule 1: APICAL MEMBRANE ANTIGEN 1

Chain E: 

4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	251.56Å 51.00Å 145.34Å 90.00° 90.93° 90.00°	Depositor
Resolution (Å)	48.44 – 2.90 47.31 – 2.90	Depositor EDS
% Data completeness (in resolution range)	98.8 (48.44-2.90) 98.8 (47.31-2.90)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.74 (at 2.91Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, R_{free}	0.193 , 0.271 0.192 , 0.271	Depositor DCC
R_{free} test set	2078 reflections (5.31%)	DCC
Wilson B-factor (Å ²)	43.3	Xtriage
Anisotropy	0.051	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 11.6	EDS
Estimated twinning fraction	0.011 for -h,-k,l	Xtriage
L-test for twinning	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	1 of 41234 reflections (0.002%)	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	12647	wwPDB-VP
Average B, all atoms (Å ²)	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 21.42 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 7.0365e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹Intensities estimated from amplitudes.

5 Model quality

5.1 Standard geometry

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	0.50	1/3244 (0.0%)	0.55	0/4401
1	B	0.50	2/3212 (0.1%)	0.53	0/4358
1	D	0.52	1/3222 (0.0%)	0.56	0/4369
1	E	0.49	1/3181 (0.0%)	0.53	0/4313
All	All	0.50	5/12859 (0.0%)	0.54	0/17441

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	D	0	1
1	E	0	1
All	All	0	2

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	348	TRP	CD2-CE2	5.21	1.47	1.41
1	B	293	TRP	CD2-CE2	5.11	1.47	1.41
1	E	347	TRP	CD2-CE2	5.10	1.47	1.41
1	B	308	TRP	CD2-CE2	5.04	1.47	1.41
1	D	114	TRP	CD2-CE2	5.00	1.47	1.41

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	D	353	LYS	Peptide
1	E	277	LEU	Peptide

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	3155	0	2950	96	0
1	B	3123	0	2919	109	0
1	D	3133	0	2931	94	0
1	E	3094	0	2894	100	0
2	A	45	0	0	1	0
2	B	33	0	0	0	0
2	D	45	0	0	5	0
2	E	19	0	0	1	0
All	All	12647	0	11694	396	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 16.

The worst 5 of 396 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:69:THR:HG22	1:A:72:LEU:HB2	1.26	1.18
1:B:422:THR:HB	1:B:423:PRO:CD	1.76	1.13
1:A:183:THR:HG22	1:A:186:GLY:H	1.10	1.12
1:D:205:THR:HG22	1:D:207:TYR:H	1.15	1.11
1:D:217:ARG:HH11	1:D:217:ARG:HG2	1.11	1.08

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	393/431 (91%)	371 (94%)	22 (6%)	0	100	100
1	B	389/431 (90%)	353 (91%)	35 (9%)	1 (0%)	50	85
1	D	390/431 (90%)	358 (92%)	26 (7%)	6 (2%)	15	50
1	E	385/431 (89%)	345 (90%)	38 (10%)	2 (0%)	38	79
All	All	1557/1724 (90%)	1427 (92%)	121 (8%)	9 (1%)	33	76

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	326	ASP
1	D	471	ASN
1	D	472	ASP
1	E	467	ALA
1	D	233	SER

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	348/372 (94%)	301 (86%)	47 (14%)	6	15
1	B	344/372 (92%)	294 (86%)	50 (14%)	5	13
1	D	345/372 (93%)	302 (88%)	43 (12%)	7	19
1	E	340/372 (91%)	290 (85%)	50 (15%)	4	13
All	All	1377/1488 (92%)	1187 (86%)	190 (14%)	5	14

5 of 190 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	409	ASP
1	D	120	LEU
1	E	294	THR
1	B	430	ARG
1	B	472	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 59 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	373	ASN
1	D	124	ASN
1	E	332	GLN
1	B	428	GLN
1	D	65	ASN

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 ⓘ

There are no ligands in this entry.

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	398/431 (92%)	-0.23	2 (0%) 88 93	19, 33, 59, 79	0
1	B	394/431 (91%)	-0.06	8 (2%) 62 71	17, 35, 71, 97	0
1	D	395/431 (91%)	-0.24	1 (0%) 91 95	18, 32, 56, 80	0
1	E	390/431 (90%)	0.05	15 (3%) 38 45	21, 42, 79, 95	0
All	All	1577/1724 (91%)	-0.12	26 (1%) 68 78	17, 35, 69, 97	0

The worst 5 of 26 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	467	ALA	4.0
1	B	468	GLU	3.5
1	E	425	GLU	3.5
1	B	473	CYS	3.3
1	B	443	VAL	3.2

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

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.