



Full wwPDB X-ray Structure Validation Report

Feb 26, 2014 – 10:56 PM GMT

PDB ID : 3EE1
Title : Novel fold of VirA, a type III secretion system effector protein from *Shigella flexneri*
Authors : Davis, J.S.
Deposited on : 2008-09-03
Resolution : 3.01 Å(reported)

This is a full wwPDB 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/ValidationPDFNotes.html>

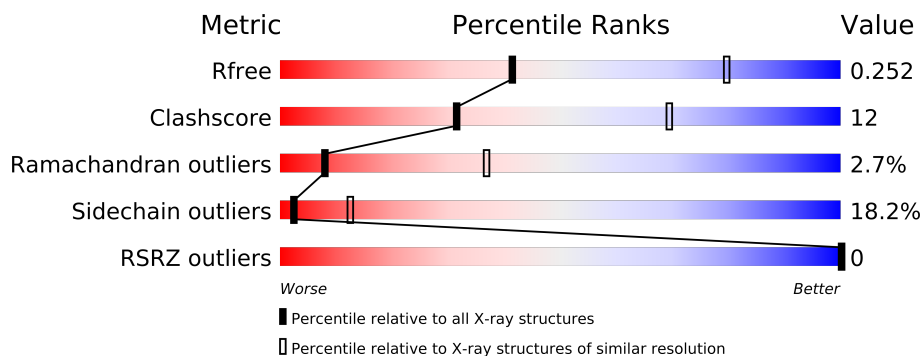
The following versions of software and data (see [references](#)) were used in the production of this report:

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 3.01 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	1332 (3.04-3.00)
Clashscore	79885	1732 (3.04-3.00)
Ramachandran outliers	78287	1669 (3.04-3.00)
Sidechain outliers	78261	1672 (3.04-3.00)
RSRZ outliers	66119	1333 (3.04-3.00)

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. 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	400	
1	B	400	

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 5244 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 Effector protein virA.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	341	Total	C	N	O	S	Se	0	0	0
			2622	1660	439	511	7	5			
1	B	341	Total	C	N	O	S	Se	0	0	0
			2622	1660	439	511	7	5			

There are 34 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLU	-	EXPRESSION TAG	UNP Q7BU69
A	2	VAL	-	EXPRESSION TAG	UNP Q7BU69
A	3	ALA	-	EXPRESSION TAG	UNP Q7BU69
A	4	SER	-	EXPRESSION TAG	UNP Q7BU69
A	5	TYR	-	EXPRESSION TAG	UNP Q7BU69
A	6	CYS	-	EXPRESSION TAG	UNP Q7BU69
A	7	ASP	-	EXPRESSION TAG	UNP Q7BU69
A	8	ARG	-	EXPRESSION TAG	UNP Q7BU69
A	9	VAL	-	EXPRESSION TAG	UNP Q7BU69
A	10	VAL	-	EXPRESSION TAG	UNP Q7BU69
A	11	ALA	-	EXPRESSION TAG	UNP Q7BU69
A	12	ALA	-	EXPRESSION TAG	UNP Q7BU69
A	13	VAL	-	EXPRESSION TAG	UNP Q7BU69
A	36	VAL	ILE	VARIANT	UNP Q7BU69
A	56	TYR	HIS	VARIANT	UNP Q7BU69
A	173	PRO	HIS	VARIANT	UNP Q7BU69
A	239	LEU	VAL	VARIANT	UNP Q7BU69
B	1	GLU	-	EXPRESSION TAG	UNP Q7BU69
B	2	VAL	-	EXPRESSION TAG	UNP Q7BU69
B	3	ALA	-	EXPRESSION TAG	UNP Q7BU69
B	4	SER	-	EXPRESSION TAG	UNP Q7BU69
B	5	TYR	-	EXPRESSION TAG	UNP Q7BU69
B	6	CYS	-	EXPRESSION TAG	UNP Q7BU69
B	7	ASP	-	EXPRESSION TAG	UNP Q7BU69
B	8	ARG	-	EXPRESSION TAG	UNP Q7BU69

Continued on next page...

Continued from previous page...

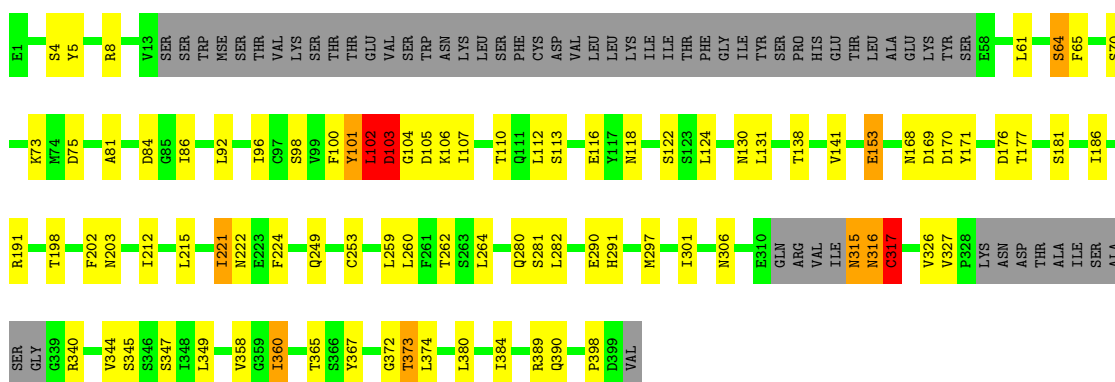
Chain	Residue	Modelled	Actual	Comment	Reference
B	9	VAL	-	EXPRESSION TAG	UNP Q7BU69
B	10	VAL	-	EXPRESSION TAG	UNP Q7BU69
B	11	ALA	-	EXPRESSION TAG	UNP Q7BU69
B	12	ALA	-	EXPRESSION TAG	UNP Q7BU69
B	13	VAL	-	EXPRESSION TAG	UNP Q7BU69
B	36	VAL	ILE	VARIANT	UNP Q7BU69
B	56	TYR	HIS	VARIANT	UNP Q7BU69
B	173	PRO	HIS	VARIANT	UNP Q7BU69
B	239	LEU	VAL	VARIANT	UNP Q7BU69

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 ($\text{RSRZ} > 2$). 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.

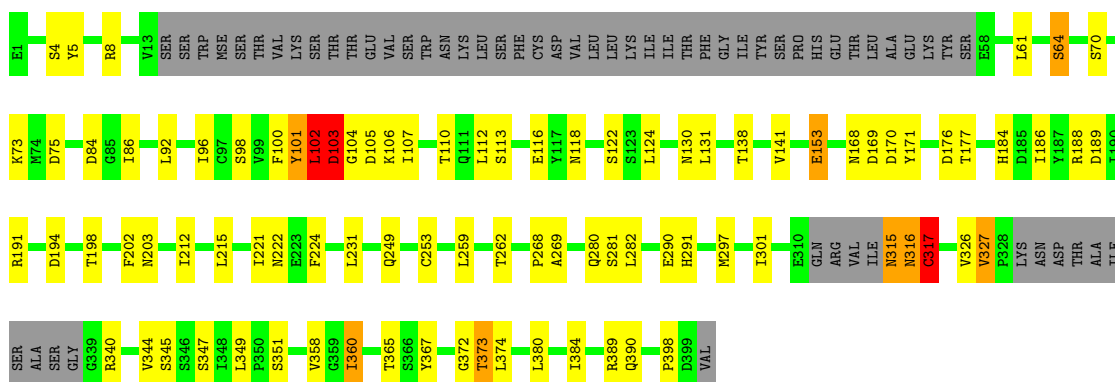
- Molecule 1: Effector protein virA

Chain A:



- Molecule 1: Effector protein virA

Chain B:



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	150.25Å 170.91Å 46.16Å 90.00° 104.89° 90.00°	Depositor
Resolution (Å)	28.63 – 3.01 28.63 – 3.01	Depositor EDS
% Data completeness (in resolution range)	96.2 (28.63-3.01) 96.3 (28.63-3.01)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.77 (at 3.00Å)	Xtriage
Refinement program	PHENIX (phenix.refine)	Depositor
R, R_{free}	0.214 , 0.256 0.209 , 0.252	Depositor DCC
R_{free} test set	1096 reflections (5.11%)	DCC
Wilson B-factor (Å ²)	98.9	Xtriage
Anisotropy	0.038	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 83.7	EDS
Estimated twinning fraction	0.032 for -h-2*1,-k,l	Xtriage
L-test for twinning	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 21470 reflections	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	5244	wwPDB-VP
Average B, all atoms (Å ²)	120.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.70% of the height of the origin peak. No significant pseudotranslation is detected.*

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	0.56	0/2670	0.74	2/3618 (0.1%)
1	B	0.53	0/2670	0.72	2/3618 (0.1%)
All	All	0.54	0/5340	0.73	4/7236 (0.1%)

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	1
1	B	0	1
All	All	0	2

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	104	GLY	N-CA-C	-6.59	96.63	113.10
1	B	104	GLY	N-CA-C	-6.23	97.52	113.10
1	A	102	LEU	CA-CB-CG	5.45	127.83	115.30
1	B	102	LEU	CA-CB-CG	5.19	127.23	115.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	103	ASP	Peptide
1	B	103	ASP	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	2622	0	0	34	0
1	B	2622	0	0	31	0
All	All	5244	0	0	61	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 12.

All (61) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:101:TYR:O	1:A:103:ASP:N	2.28	0.66
1:A:101:TYR:O	1:A:103:ASP:OD1	2.14	0.66
1:A:102:LEU:C	1:A:103:ASP:OD1	2.34	0.65
1:B:101:TYR:O	1:B:103:ASP:N	2.30	0.64
1:A:61:LEU:O	1:A:64:SER:OG	2.17	0.63
1:A:102:LEU:CD1	1:A:103:ASP:OD1	2.48	0.62
1:B:102:LEU:CD1	1:B:103:ASP:OD1	2.49	0.60
1:B:61:LEU:O	1:B:64:SER:OG	2.21	0.59
1:B:101:TYR:O	1:B:103:ASP:OD1	2.22	0.57
1:A:102:LEU:C	1:A:103:ASP:CG	2.63	0.57
1:A:315:ASN:O	1:A:317:CYS:N	2.39	0.56
1:B:316:ASN:O	1:B:317:CYS:O	2.23	0.56
1:A:316:ASN:O	1:A:317:CYS:O	2.25	0.55
1:A:291:HIS:CD2	1:B:290:GLU:OE1	2.60	0.55
1:B:315:ASN:O	1:B:317:CYS:N	2.41	0.54
1:A:306:ASN:ND2	1:B:188:ARG:NE	2.57	0.52
1:B:326:VAL:CG2	1:B:327:VAL:N	2.73	0.52
1:B:5:TYR:O	1:B:8:ARG:N	2.44	0.51
1:B:102:LEU:C	1:B:103:ASP:OD1	2.48	0.51
1:A:297:MSE:O	1:A:301:ILE:CD1	2.57	0.51
1:B:153:GLU:N	1:B:153:GLU:OE2	2.43	0.51
1:B:102:LEU:C	1:B:103:ASP:CG	2.69	0.51
1:A:84:ASP:O	1:A:86:ILE:N	2.44	0.51
1:A:102:LEU:O	1:A:103:ASP:CG	2.49	0.51
1:A:103:ASP:N	1:A:103:ASP:OD1	2.44	0.50

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:326:VAL:CG2	1:A:327:VAL:N	2.75	0.50
1:B:153:GLU:CD	1:B:153:GLU:N	2.65	0.49
1:A:102:LEU:C	1:A:107:ILE:O	2.50	0.49
1:A:100:PHE:O	1:A:102:LEU:CD1	2.61	0.49
1:B:102:LEU:C	1:B:107:ILE:O	2.51	0.48
1:A:153:GLU:N	1:A:153:GLU:OE2	2.46	0.48
1:B:297:MSE:O	1:B:301:ILE:CD1	2.63	0.47
1:A:5:TYR:O	1:A:8:ARG:N	2.47	0.47
1:B:84:ASP:O	1:B:86:ILE:N	2.48	0.46
1:B:102:LEU:O	1:B:103:ASP:CG	2.54	0.46
1:A:372:GLY:O	1:A:373:THR:O	2.35	0.45
1:A:344:VAL:CG1	1:A:367:TYR:CE2	3.00	0.45
1:A:153:GLU:N	1:A:153:GLU:CD	2.71	0.45
1:A:349:LEU:CB	1:A:360:ILE:CD1	2.95	0.44
1:B:349:LEU:CB	1:B:360:ILE:CD1	2.95	0.44
1:A:290:GLU:OE1	1:B:291:HIS:CD2	2.71	0.44
1:A:344:VAL:CG1	1:A:367:TYR:CD2	3.01	0.44
1:A:84:ASP:C	1:A:86:ILE:N	2.72	0.44
1:A:384:ILE:O	1:A:389:ARG:NH1	2.51	0.43
1:A:202:PHE:O	1:A:203:ASN:C	2.57	0.43
1:B:202:PHE:O	1:B:203:ASN:C	2.56	0.43
1:A:221:ILE:CG1	1:A:221:ILE:O	2.66	0.43
1:A:65:PHE:CE2	1:A:81:ALA:CB	3.02	0.43
1:B:316:ASN:O	1:B:317:CYS:C	2.58	0.42
1:A:306:ASN:OD1	1:B:184:HIS:NE2	2.52	0.42
1:B:100:PHE:O	1:B:102:LEU:CD1	2.67	0.42
1:A:316:ASN:O	1:A:317:CYS:C	2.58	0.42
1:B:344:VAL:CG1	1:B:367:TYR:CE2	3.03	0.42
1:B:384:ILE:O	1:B:389:ARG:NH1	2.53	0.42
1:A:168:ASN:CB	1:A:171:TYR:CZ	3.03	0.41
1:A:260:LEU:O	1:A:264:LEU:N	2.53	0.41
1:B:268:PRO:O	1:B:269:ALA:C	2.58	0.41
1:B:168:ASN:CB	1:B:171:TYR:CZ	3.04	0.41
1:B:372:GLY:O	1:B:373:THR:O	2.38	0.41
1:B:188:ARG:NH1	1:B:189:ASP:OD1	2.53	0.41
1:B:84:ASP:C	1:B:86:ILE:N	2.74	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	333/400 (83%)	283 (85%)	41 (12%)	9 (3%)	8	37
1	B	333/400 (83%)	285 (86%)	39 (12%)	9 (3%)	8	37
All	All	666/800 (83%)	568 (85%)	80 (12%)	18 (3%)	8	37

All (18) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	102	LEU
1	A	316	ASN
1	A	317	CYS
1	A	373	THR
1	B	4	SER
1	B	102	LEU
1	B	316	ASN
1	B	317	CYS
1	B	373	THR
1	A	4	SER
1	A	70	SER
1	A	105	ASP
1	A	141	VAL
1	B	70	SER
1	B	105	ASP
1	B	141	VAL
1	A	398	PRO
1	B	398	PRO

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	294/357 (82%)	242 (82%)	52 (18%)	3	13
1	B	294/357 (82%)	239 (81%)	55 (19%)	2	12
All	All	588/714 (82%)	481 (82%)	107 (18%)	2	12

All (107) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	64	SER
1	A	73	LYS
1	A	75	ASP
1	A	92	LEU
1	A	96	ILE
1	A	98	SER
1	A	101	TYR
1	A	102	LEU
1	A	103	ASP
1	A	106	LYS
1	A	110	THR
1	A	112	LEU
1	A	113	SER
1	A	116	GLU
1	A	118	ASN
1	A	122	SER
1	A	124	LEU
1	A	130	ASN
1	A	131	LEU
1	A	138	THR
1	A	153	GLU
1	A	169	ASP
1	A	170	ASP
1	A	176	ASP
1	A	177	THR
1	A	181	SER
1	A	186	ILE
1	A	191	ARG
1	A	198	THR
1	A	212	ILE
1	A	215	LEU
1	A	221	ILE
1	A	222	ASN
1	A	224	PHE
1	A	249	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	253	CYS
1	A	259	LEU
1	A	262	THR
1	A	280	GLN
1	A	281	SER
1	A	282	LEU
1	A	315	ASN
1	A	317	CYS
1	A	340	ARG
1	A	345	SER
1	A	347	SER
1	A	358	VAL
1	A	360	ILE
1	A	365	THR
1	A	374	LEU
1	A	380	LEU
1	A	390	GLN
1	B	64	SER
1	B	73	LYS
1	B	75	ASP
1	B	92	LEU
1	B	96	ILE
1	B	98	SER
1	B	101	TYR
1	B	102	LEU
1	B	103	ASP
1	B	106	LYS
1	B	110	THR
1	B	112	LEU
1	B	113	SER
1	B	116	GLU
1	B	118	ASN
1	B	122	SER
1	B	124	LEU
1	B	130	ASN
1	B	131	LEU
1	B	138	THR
1	B	153	GLU
1	B	169	ASP
1	B	170	ASP
1	B	176	ASP
1	B	177	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	186	ILE
1	B	191	ARG
1	B	194	ASP
1	B	198	THR
1	B	212	ILE
1	B	215	LEU
1	B	221	ILE
1	B	222	ASN
1	B	224	PHE
1	B	231	LEU
1	B	249	GLN
1	B	253	CYS
1	B	259	LEU
1	B	262	THR
1	B	280	GLN
1	B	281	SER
1	B	282	LEU
1	B	315	ASN
1	B	317	CYS
1	B	327	VAL
1	B	340	ARG
1	B	345	SER
1	B	347	SER
1	B	351	SER
1	B	358	VAL
1	B	360	ILE
1	B	365	THR
1	B	374	LEU
1	B	380	LEU
1	B	390	GLN

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 ⓘ

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	341/400 (85%)	-0.21	0 100 100	68, 106, 171, 489	0
1	B	341/400 (85%)	-0.17	0 100 100	71, 116, 209, 465	0
All	All	682/800 (85%)	-0.19	0 100 100	68, 111, 190, 489	0

There are no RSRZ outliers to report.

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.