



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 06:11 AM GMT

PDB ID : 2W7U  
Title : SPLA SERINE PROTEASE OF STAPHYLOCOCCUS AUREUS (2.4A)  
Authors : Stec-Niemczyka, J.; Pustelny, K.; Kisielewska, M.; Bista, M.; Boulware, K.T.;  
Stennicke, H.R.; Thogersen, I.B.; Daugherty, P.S.; Enghild, J.J.; Popowicz,  
G.M.; Dubin, A.; Potempa, J.; Dubin, G.  
Deposited on : 2008-12-30  
Resolution : 2.43 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](mailto: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 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
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) : trunk26865

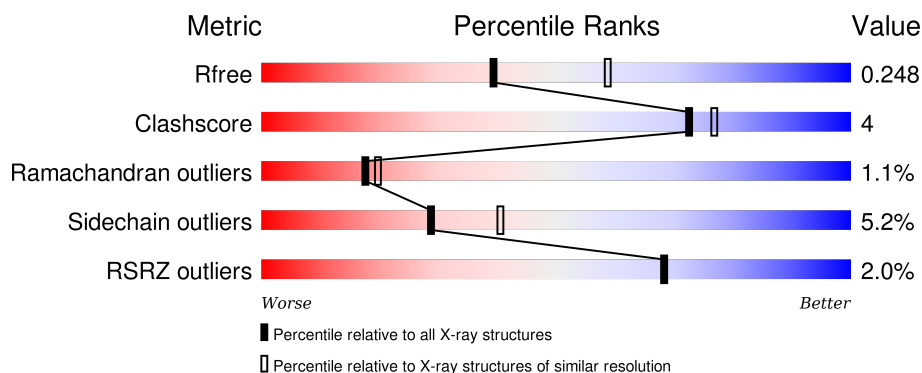
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.43 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))
$R_{free}$	91344	1003 (2.46-2.42)
Clashscore	102246	1071 (2.46-2.42)
Ramachandran outliers	100387	1065 (2.46-2.42)
Sidechain outliers	100360	1065 (2.46-2.42)
RSRZ outliers	91569	1005 (2.46-2.42)

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. 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	200	<div> <div>2%</div> <div>85%</div> <div>12%</div> <div>..</div> </div>
1	B	200	<div> <div>%</div> <div>84%</div> <div>12%</div> <div>..</div> </div>
1	C	200	<div> <div>2%</div> <div>89%</div> <div>11%</div> <div>.</div> </div>
1	D	200	<div> <div>4%</div> <div>84%</div> <div>14%</div> <div>..</div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 5995 atoms, of which 0 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 SERINE PROTEASE SPLA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	198	Total	C	N	O	S	0	0	0
			1487	938	250	297	2			
1	B	200	Total	C	N	O	S	0	0	0
			1500	946	250	302	2			
1	C	200	Total	C	N	O	S	0	0	0
			1482	934	247	299	2			
1	D	200	Total	C	N	O	S	0	0	0
			1491	940	249	300	2			

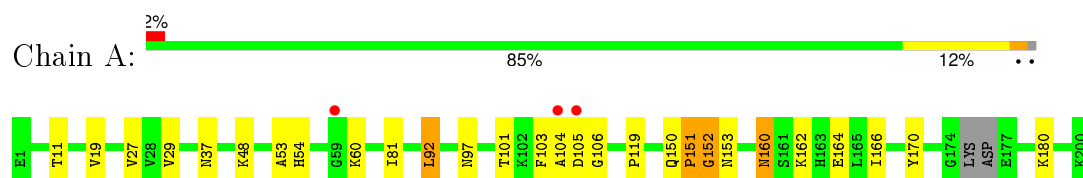
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	9	Total	O	0	0
			9	9		
2	B	17	Total	O	0	0
			17	17		
2	C	5	Total	O	0	0
			5	5		
2	D	4	Total	O	0	0
			4	4		

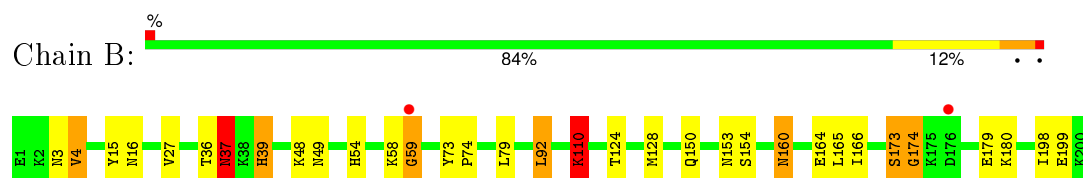
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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.

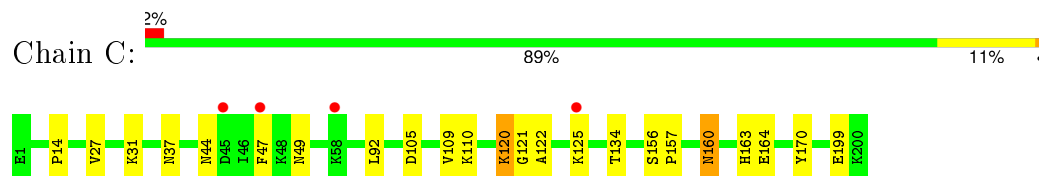
#### • Molecule 1: SERINE PROTEASE SPLA



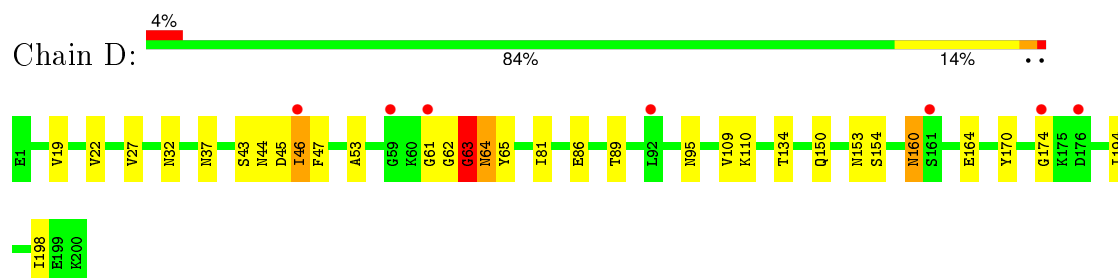
#### • Molecule 1: SERINE PROTEASE SPLA



#### • Molecule 1: SERINE PROTEASE SPLA



#### • Molecule 1: SERINE PROTEASE SPLA



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	73.99Å 82.22Å 134.19Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.87 – 2.43 19.87 – 2.43	Depositor EDS
% Data completeness (in resolution range)	97.9 (19.87-2.43) 97.9 (19.87-2.43)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.62 (at 2.44Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.209 , 0.248 0.207 , 0.248	Depositor DCC
$R_{free}$ test set	1559 reflections (5.32%)	DCC
Wilson B-factor (Å <sup>2</sup> )	33.7	Xtriage
Anisotropy	0.059	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 30.5	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Outliers	0 of 30841 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	5995	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.78% of the height of the origin peak. No significant pseudotranslation is detected.*

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

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

## 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	0.69	6/1517 (0.4%)	0.60	3/2052 (0.1%)
1	B	0.75	4/1531 (0.3%)	0.49	2/2073 (0.1%)
1	C	0.33	1/1513 (0.1%)	0.41	2/2053 (0.1%)
1	D	0.22	0/1522	0.35	0/2062
All	All	0.55	11/6083 (0.2%)	0.47	7/8240 (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	4
1	B	0	7
1	C	0	1
1	D	0	8
All	All	0	20

All (11) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	110	LYS	C-O	23.79	1.68	1.23
1	A	152	GLY	CA-C	-14.68	1.28	1.51
1	A	151	PRO	C-N	14.68	1.59	1.33
1	A	151	PRO	C-O	10.14	1.43	1.23
1	B	110	LYS	C-N	-7.75	1.16	1.34
1	B	39	HIS	C-O	7.58	1.37	1.23
1	A	152	GLY	N-CA	-7.49	1.34	1.46
1	C	120	LYS	CA-CB	7.31	1.70	1.53
1	A	101	THR	CB-OG1	6.09	1.55	1.43
1	B	39	HIS	N-CA	5.37	1.57	1.46
1	A	29	VAL	CB-CG2	-5.34	1.41	1.52

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	151	PRO	CA-C-N	-15.02	86.16	116.20
1	A	101	THR	OG1-CB-CG2	-13.17	79.72	110.00
1	B	37	ASN	CB-CA-C	9.23	128.87	110.40
1	C	120	LYS	N-CA-C	7.07	130.08	111.00
1	B	110	LYS	CA-C-O	-6.67	106.10	120.10
1	A	151	PRO	C-N-CA	-5.72	110.28	122.30
1	C	120	LYS	CA-CB-CG	-5.36	101.60	113.40

There are no chirality outliers.

All (20) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	103	PHE	Peptide
1	A	105	ASP	Peptide
1	A	151	PRO	Mainchain,Peptide
1	B	110	LYS	Mainchain
1	B	173	SER	Peptide
1	B	179	GLU	Peptide
1	B	198	ILE	Peptide
1	B	3	ASN	Peptide
1	B	58	LYS	Peptide
1	B	59	GLY	Peptide
1	C	125	LYS	Peptide
1	D	174	GLY	Peptide
1	D	43	SER	Peptide
1	D	44	ASN	Peptide
1	D	45	ASP	Peptide
1	D	46	ILE	Peptide
1	D	61	GLY	Peptide
1	D	62	GLY	Peptide
1	D	63	GLY	Peptide

## 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	1487	0	1403	11	0
1	B	1500	0	1404	18	0
1	C	1482	0	1357	9	0
1	D	1491	0	1381	10	0
2	A	9	0	0	0	0
2	B	17	0	0	0	0
2	C	5	0	0	0	0
2	D	4	0	0	0	0
All	All	5995	0	5545	48	0

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 4.

All (48) 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)
1:B:110:LYS:C	1:B:110:LYS:O	1.68	1.31
1:B:173:SER:H	1:B:174:GLY:HA3	1.35	0.91
1:B:173:SER:N	1:B:174:GLY:HA3	1.94	0.81
1:B:160:ASN:HD22	1:B:160:ASN:C	2.04	0.60
1:B:150:GLN:H	1:B:153:ASN:HD22	1.48	0.59
1:C:31:LYS:O	1:C:199:GLU:HB2	2.04	0.57
1:A:37:ASN:HD21	1:A:170:TYR:HA	1.71	0.56
1:D:150:GLN:H	1:D:153:ASN:HD22	1.54	0.55
1:B:160:ASN:HB3	1:B:166:ILE:HD11	1.87	0.55
1:D:63:GLY:HA2	1:D:65:TYR:CE1	2.41	0.55
1:A:160:ASN:C	1:A:160:ASN:HD22	2.10	0.55
1:A:160:ASN:ND2	1:A:164:GLU:H	2.06	0.54
1:C:109:VAL:O	1:C:110:LYS:HB2	2.08	0.52
1:B:37:ASN:HD21	1:B:154:SER:HB3	1.73	0.52
1:B:37:ASN:ND2	1:B:154:SER:HB3	2.25	0.51
1:D:32:ASN:HD21	1:D:86:GLU:HG3	1.76	0.51
1:C:44:ASN:HA	1:C:49:ASN:HD22	1.75	0.51
1:B:4:VAL:O	1:B:4:VAL:HG12	2.11	0.50
1:C:160:ASN:C	1:C:160:ASN:HD22	2.15	0.50
1:D:19:VAL:HG12	1:D:53:ALA:HA	1.95	0.49
1:B:15:TYR:HB3	1:B:128:MET:HE1	1.95	0.49
1:A:119:PRO:HA	1:A:152:GLY:HA3	1.96	0.47
1:A:19:VAL:HG12	1:A:53:ALA:HA	1.95	0.47
1:B:110:LYS:O	1:B:110:LYS:CA	2.59	0.47
1:B:160:ASN:ND2	1:B:164:GLU:H	2.12	0.47
1:A:160:ASN:HB3	1:A:166:ILE:HD11	1.96	0.47

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:173:SER:N	1:B:174:GLY:CA	2.71	0.46
1:D:37:ASN:HD21	1:D:170:TYR:HA	1.81	0.46
1:B:36:THR:OG1	1:B:37:ASN:N	2.49	0.44
1:C:37:ASN:HD21	1:C:170:TYR:HA	1.83	0.44
1:A:160:ASN:ND2	1:A:162:LYS:H	2.15	0.43
1:C:160:ASN:ND2	1:C:164:GLU:H	2.16	0.43
1:C:120:LYS:O	1:C:122:ALA:N	2.51	0.43
1:B:173:SER:H	1:B:174:GLY:CA	2.19	0.43
1:D:109:VAL:O	1:D:110:LYS:HB2	2.19	0.42
1:A:150:GLN:H	1:A:153:ASN:HD22	1.68	0.42
1:D:160:ASN:ND2	1:D:164:GLU:H	2.17	0.41
1:B:16:ASN:HA	1:B:54:HIS:HB3	2.01	0.41
1:C:14:PRO:HG3	1:C:163:HIS:CG	2.56	0.41
1:A:92:LEU:HD12	1:A:97:ASN:HD21	1.85	0.41
1:A:104:ALA:O	1:A:106:GLY:HA2	2.20	0.41
1:B:92:LEU:HA	1:B:92:LEU:HD12	1.89	0.41
1:D:194:ILE:O	1:D:198:ILE:HG13	2.21	0.41
1:B:73:TYR:HA	1:B:74:PRO:HD2	1.95	0.41
1:A:54:HIS:ND1	1:A:54:HIS:O	2.54	0.40
1:D:32:ASN:H	1:D:95:ASN:HD21	1.67	0.40
1:D:63:GLY:O	1:D:64:ASN:ND2	2.54	0.40
1:C:156:SER:HA	1:C:157:PRO:HD2	1.91	0.40

There are no symmetry-related clashes.

## 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	194/200 (97%)	182 (94%)	11 (6%)	1 (0%)	34	41
1	B	198/200 (99%)	185 (93%)	9 (4%)	4 (2%)	9	7
1	C	198/200 (99%)	185 (93%)	12 (6%)	1 (0%)	34	41

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	D	198/200 (99%)	181 (91%)	14 (7%)	3 (2%)	13	12
All	All	788/800 (98%)	733 (93%)	46 (6%)	9 (1%)	17	20

All (9) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	46	ILE
1	B	4	VAL
1	B	59	GLY
1	C	121	GLY
1	D	47	PHE
1	B	39	HIS
1	B	174	GLY
1	D	63	GLY
1	A	60	LYS

### 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	155/168 (92%)	148 (96%)	7 (4%)	34	47
1	B	155/168 (92%)	144 (93%)	11 (7%)	18	24
1	C	149/168 (89%)	143 (96%)	6 (4%)	38	53
1	D	151/168 (90%)	143 (95%)	8 (5%)	28	40
All	All	610/672 (91%)	578 (95%)	32 (5%)	29	40

All (32) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	11	THR
1	A	27	VAL
1	A	48	LYS
1	A	81	ILE
1	A	92	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	160	ASN
1	A	180	LYS
1	B	27	VAL
1	B	37	ASN
1	B	48	LYS
1	B	49	ASN
1	B	79	LEU
1	B	92	LEU
1	B	124	THR
1	B	160	ASN
1	B	165	LEU
1	B	180	LYS
1	B	199	GLU
1	C	27	VAL
1	C	47	PHE
1	C	92	LEU
1	C	105	ASP
1	C	134	THR
1	C	160	ASN
1	D	22	VAL
1	D	27	VAL
1	D	64	ASN
1	D	81	ILE
1	D	89	THR
1	D	134	THR
1	D	154	SER
1	D	160	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (28) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	32	ASN
1	A	37	ASN
1	A	44	ASN
1	A	95	ASN
1	A	153	ASN
1	A	160	ASN
1	A	196	ASN
1	B	32	ASN
1	B	37	ASN
1	B	49	ASN
1	B	64	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	95	ASN
1	B	153	ASN
1	B	160	ASN
1	B	196	ASN
1	C	32	ASN
1	C	37	ASN
1	C	49	ASN
1	C	95	ASN
1	C	160	ASN
1	C	196	ASN
1	D	32	ASN
1	D	37	ASN
1	D	64	ASN
1	D	95	ASN
1	D	153	ASN
1	D	160	ASN
1	D	196	ASN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

### 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

### 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 5.6 Ligand geometry [i](#)

There are no ligands in this entry.

### 5.7 Other polymers [i](#)

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	198/200 (99%)	-0.29	3 (1%) 76 77	14, 28, 44, 52	0
1	B	200/200 (100%)	-0.23	2 (1%) 84 85	14, 27, 45, 55	0
1	C	200/200 (100%)	-0.11	4 (2%) 68 68	19, 31, 54, 68	0
1	D	200/200 (100%)	-0.00	7 (3%) 48 46	20, 36, 55, 60	0
All	All	798/800 (99%)	-0.16	16 (2%) 68 68	14, 30, 51, 68	0

All (16) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	176	ASP	3.8
1	D	59	GLY	3.8
1	D	176	ASP	3.6
1	D	174	GLY	3.2
1	A	104	ALA	3.2
1	A	105	ASP	2.6
1	C	45	ASP	2.6
1	D	161	SER	2.4
1	B	59	GLY	2.3
1	C	47	PHE	2.3
1	D	61	GLY	2.2
1	D	92	LEU	2.2
1	A	59	GLY	2.2
1	C	125	LYS	2.1
1	C	58	LYS	2.0
1	D	46	ILE	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 [i](#)

There are no carbohydrates in this entry.

### 6.4 Ligands [i](#)

There are no ligands in this entry.

### 6.5 Other polymers [i](#)

There are no such residues in this entry.