



# Full wwPDB X-ray Structure Validation Report

Feb 28, 2014 – 10:02 PM GMT

PDB ID : 2RSP  
Title : STRUCTURE OF THE ASPARTIC PROTEASE FROM ROUS SARCOMA  
RETROVIRUS REFINED AT 2 ANGSTROMS RESOLUTION  
Authors : Wlodawer, A.; Miller, M.; Jaskolski, M.  
Deposited on : 1989-10-17  
Resolution : 2.00 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>

---

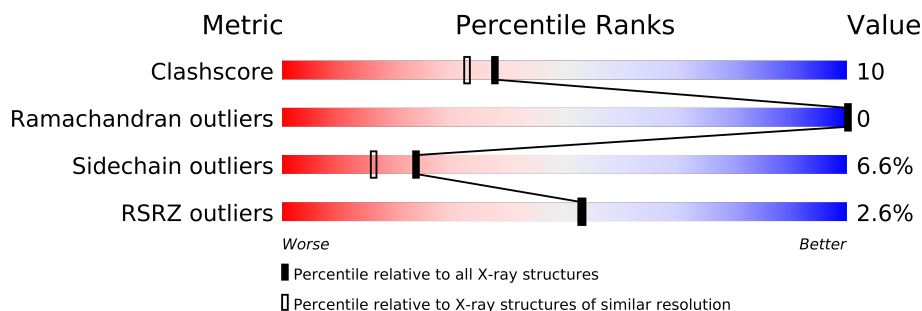
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 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 2.00 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))
Clashscore	79885	6188 (2.00-2.00)
Ramachandran outliers	78287	6102 (2.00-2.00)
Sidechain outliers	78261	6100 (2.00-2.00)
RSRZ outliers	66119	4890 (2.00-2.00)

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  $\geq 3$ , 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	124	
1	B	124	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2024 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 RSV PROTEASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	115	Total	C	N	O	S	0	1	0
			891	560	159	164	8			
1	B	113	Total	C	N	O	S	0	1	0
			881	554	157	162	8			

- Molecule 2 is water.

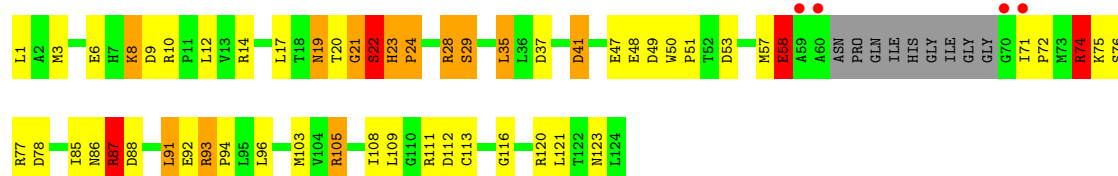
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	114	Total	O	0	0
			114	114		
2	B	138	Total	O	0	0
			138	138		

### 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.

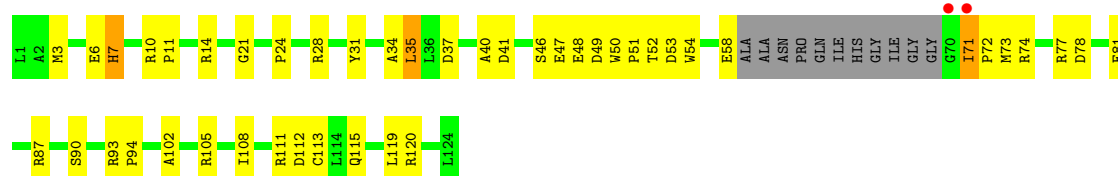
#### • Molecule 1: RSV PROTEASE

Chain A: 



#### • Molecule 1: RSV PROTEASE

Chain B: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	88.95Å 88.95Å 78.90Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	10.00 – 2.00 9.97 – 2.06	Depositor EDS
% Data completeness (in resolution range)	(Not available) (10.00-2.00) 77.9 (9.97-2.06)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$	-	Xtriage
Refinement program	PROFFT	Depositor
R, $R_{free}$	0.144 , (Not available) 0.134 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	DCC
Wilson B-factor (Å <sup>2</sup> )	17.5	Xtriage
Anisotropy	0.184	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 76.4	EDS
Estimated twinning fraction	0.053 for -h,-k,l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.44$ , $\langle L^2 \rangle = 0.26$	Xtriage
Outliers	0 of 17609 reflections	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	2024	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	22.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.30% of the height of the origin peak. No significant pseudotranslation is detected.*

## 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	1.25	1/909 (0.1%)	2.77	55/1229 (4.5%)
1	B	1.22	2/899 (0.2%)	2.62	55/1215 (4.5%)
All	All	1.23	3/1808 (0.2%)	2.69	110/2444 (4.5%)

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

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	92	GLU	CD-OE1	-7.49	1.17	1.25
1	B	90	SER	CB-OG	6.20	1.50	1.42
1	B	111	ARG	CZ-NH1	5.16	1.39	1.33

All (110) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	87	ARG	CD-NE-CZ	25.95	159.93	123.60
1	A	74	ARG	CD-NE-CZ	24.03	157.24	123.60
1	A	93	ARG	NE-CZ-NH2	-22.58	109.01	120.30
1	A	28	ARG	CG-CD-NE	18.86	151.41	111.80
1	A	93	ARG	NE-CZ-NH1	16.41	128.51	120.30
1	A	105	ARG	NE-CZ-NH2	-15.99	112.31	120.30
1	A	74	ARG	NE-CZ-NH1	15.22	127.91	120.30
1	A	111	ARG	NE-CZ-NH1	14.85	127.72	120.30
1	B	111	ARG	NE-CZ-NH2	14.72	127.66	120.30
1	B	93	ARG	NE-CZ-NH1	13.85	127.22	120.30
1	A	120	ARG	NE-CZ-NH2	-13.23	113.68	120.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	6	GLU	OE1-CD-OE2	12.73	138.57	123.30
1	A	111	ARG	CD-NE-CZ	12.14	140.60	123.60
1	B	37	ASP	CB-CG-OD2	-12.11	107.40	118.30
1	B	78	ASP	CB-CG-OD1	11.81	128.93	118.30
1	B	74	ARG	NE-CZ-NH2	-11.76	114.42	120.30
1	A	111	ARG	NE-CZ-NH2	-11.66	114.47	120.30
1	A	105	ARG	CD-NE-CZ	11.26	139.36	123.60
1	A	105	ARG	NE-CZ-NH1	11.25	125.93	120.30
1	B	28	ARG	NE-CZ-NH2	10.94	125.77	120.30
1	B	93	ARG	NE-CZ-NH2	-10.40	115.10	120.30
1	B	37	ASP	CB-CG-OD1	10.28	127.55	118.30
1	A	88	ASP	CB-CG-OD2	-10.16	109.15	118.30
1	B	78	ASP	CB-CG-OD2	-9.34	109.89	118.30
1	B	93	ARG	CD-NE-CZ	9.20	136.48	123.60
1	B	47	GLU	OE1-CD-OE2	9.04	134.15	123.30
1	A	74	ARG	NE-CZ-NH2	-8.97	115.82	120.30
1	B	87	ARG	NE-CZ-NH2	8.92	124.76	120.30
1	B	41	ASP	CB-CG-OD1	8.76	126.18	118.30
1	B	49	ASP	CB-CG-OD2	8.70	126.12	118.30
1	A	6	GLU	CA-CB-CG	-8.16	95.45	113.40
1	A	58	GLU	CB-CA-C	7.95	126.31	110.40
1	B	74	ARG	NH1-CZ-NH2	7.69	127.86	119.40
1	B	14	ARG	CD-NE-CZ	-7.67	112.86	123.60
1	A	21	GLY	O-C-N	7.52	134.74	122.70
1	A	87	ARG	CD-NE-CZ	-7.47	113.14	123.60
1	B	31	TYR	CG-CD2-CE2	-7.20	115.54	121.30
1	A	77	ARG	NE-CZ-NH1	7.18	123.89	120.30
1	B	77	ARG	NE-CZ-NH1	7.13	123.86	120.30
1	B	112	ASP	CB-CG-OD1	7.13	124.72	118.30
1	A	9	ASP	CB-CG-OD2	-7.11	111.90	118.30
1	B	10	ARG	CG-CD-NE	-7.09	96.91	111.80
1	A	8	LYS	CB-CG-CD	7.07	129.98	111.60
1	A	41	ASP	CB-CG-OD2	7.04	124.64	118.30
1	A	35	LEU	N-CA-CB	-7.01	96.38	110.40
1	B	52	THR	CA-CB-CG2	-7.00	102.59	112.40
1	B	28	ARG	NH1-CZ-NH2	-6.93	111.78	119.40
1	B	50	TRP	CE3-CZ3-CH2	6.89	128.78	121.20
1	B	24	PRO	O-C-N	-6.85	111.74	122.70
1	A	20	THR	C-N-CA	6.80	136.58	122.30
1	B	81	GLU	OE1-CD-OE2	-6.74	115.22	123.30
1	A	96	LEU	CB-CG-CD2	-6.70	99.61	111.00
1	B	77	ARG	CD-NE-CZ	-6.62	114.34	123.60

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	23	HIS	CA-CB-CG	-6.57	102.44	113.60
1	B	14	ARG	NE-CZ-NH1	6.55	123.58	120.30
1	B	105	ARG	CD-NE-CZ	6.54	132.75	123.60
1	B	48	GLU	CG-CD-OE2	-6.53	105.23	118.30
1	B	74	ARG	CD-NE-CZ	-6.37	114.68	123.60
1	A	113[A]	CYS	N-CA-CB	-6.37	99.14	110.60
1	A	113[B]	CYS	N-CA-CB	-6.37	99.14	110.60
1	B	50	TRP	CZ3-CH2-CZ2	-6.30	114.04	121.60
1	B	111	ARG	NH1-CZ-NH2	-6.17	112.61	119.40
1	A	91	LEU	CA-CB-CG	6.17	129.49	115.30
1	B	112	ASP	O-C-N	-6.09	112.96	122.70
1	B	47	GLU	CA-CB-CG	-6.08	100.01	113.40
1	A	6	GLU	OE1-CD-OE2	6.06	130.58	123.30
1	A	105	ARG	O-C-N	-6.00	113.00	123.20
1	B	74	ARG	CA-CB-CG	-5.96	100.28	113.40
1	B	108	ILE	CB-CG1-CD1	-5.95	97.24	113.90
1	A	28	ARG	NE-CZ-NH2	5.90	123.25	120.30
1	A	120	ARG	N-CA-CB	5.87	121.16	110.60
1	B	115	GLN	CG-CD-NE2	-5.83	102.70	116.70
1	B	74	ARG	CG-CD-NE	-5.82	99.58	111.80
1	A	12	LEU	O-C-N	5.80	131.98	122.70
1	A	49	ASP	C-N-CA	5.79	136.18	121.70
1	B	53	ASP	CB-CG-OD2	5.79	123.51	118.30
1	B	34	ALA	N-CA-CB	-5.68	102.15	110.10
1	B	53	ASP	CA-C-O	-5.64	108.25	120.10
1	B	87	ARG	CA-CB-CG	-5.55	101.19	113.40
1	B	21	GLY	CA-C-O	5.52	130.54	120.60
1	A	49	ASP	O-C-N	-5.51	113.89	122.70
1	A	37	ASP	N-CA-CB	-5.49	100.71	110.60
1	A	116	GLY	CA-C-N	5.43	129.14	117.20
1	A	14	ARG	CD-NE-CZ	-5.41	116.02	123.60
1	B	46	SER	N-CA-CB	5.41	118.61	110.50
1	A	77	ARG	CA-CB-CG	5.39	125.25	113.40
1	A	29	SER	CB-CA-C	-5.38	99.88	110.10
1	B	35	LEU	N-CA-CB	-5.38	99.65	110.40
1	A	86	ASN	CB-CG-ND2	5.36	129.55	116.70
1	B	87	ARG	NH1-CZ-NH2	-5.35	113.51	119.40
1	A	109	LEU	C-N-CA	5.28	133.39	122.30
1	B	74	ARG	CB-CA-C	5.28	120.95	110.40
1	A	112	ASP	CB-CG-OD1	5.25	123.03	118.30
1	B	6	GLU	CA-CB-CG	-5.25	101.85	113.40
1	A	76	SER	CB-CA-C	-5.22	100.18	110.10

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	49	ASP	CB-CG-OD1	-5.22	113.61	118.30
1	A	121	LEU	CB-CA-C	5.20	120.07	110.20
1	A	28	ARG	NE-CZ-NH1	-5.19	117.71	120.30
1	B	40	ALA	CB-CA-C	5.18	117.87	110.10
1	A	78	ASP	CB-CG-OD2	-5.18	113.64	118.30
1	B	74	ARG	NE-CZ-NH1	-5.17	117.71	120.30
1	A	22	SER	CA-C-N	5.17	128.56	117.20
1	A	47	GLU	CG-CD-OE1	5.15	128.59	118.30
1	B	120	ARG	CD-NE-CZ	-5.09	116.47	123.60
1	B	7	HIS	O-C-N	-5.07	114.59	122.70
1	A	1	LEU	CD1-CG-CD2	-5.03	95.40	110.50
1	A	14	ARG	N-CA-CB	5.03	119.65	110.60
1	A	14	ARG	NE-CZ-NH1	5.02	122.81	120.30
1	A	48	GLU	OE1-CD-OE2	-5.01	117.28	123.30
1	A	91	LEU	CB-CG-CD1	5.01	119.51	111.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	87	ARG	Sidechain

## 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	891	0	926	24	0
1	B	881	0	916	14	0
2	A	114	0	0	5	0
2	B	138	0	0	1	0
All	All	2024	0	1842	36	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 10.

All (36) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:3:MET:HE1	1:B:11:PRO:CB	2.17	0.74
1:B:113[A]:CYS:SG	2:B:377:HOH:O	2.50	0.70
1:A:23:HIS:HB2	1:A:24:PRO:HD2	1.74	0.69
1:A:71:ILE:HG13	1:A:72:PRO:HD2	1.78	0.65
1:B:3:MET:HE1	1:B:11:PRO:CA	2.27	0.64
1:A:10:ARG:NH1	2:A:289:HOH:O	2.32	0.63
1:A:58:GLU:CD	1:A:58:GLU:H	2.05	0.60
1:A:91:LEU:HD12	2:A:305:HOH:O	2.03	0.59
1:A:35:LEU:HD23	1:A:108:ILE:CD1	2.34	0.57
1:B:71:ILE:HB	1:B:73:MET:HE1	1.86	0.56
1:A:71:ILE:HD11	1:A:103:MET:O	2.06	0.56
1:A:41:ASP:O	1:B:7:HIS:HD2	1.89	0.56
1:A:19:ASN:C	1:A:19:ASN:HD22	2.09	0.55
1:B:71:ILE:HB	1:B:73:MET:CE	2.39	0.53
1:B:3:MET:HE1	1:B:11:PRO:HA	1.91	0.52
1:A:23:HIS:HB2	1:A:24:PRO:CD	2.39	0.51
1:B:11:PRO:HD2	1:B:35:LEU:HD11	1.93	0.49
1:A:53:ASP:N	1:A:53:ASP:OD2	2.45	0.49
1:A:19:ASN:ND2	1:A:21:GLY:H	2.11	0.48
1:B:73:MET:HE2	1:B:102:ALA:HB2	1.96	0.48
1:A:19:ASN:HD22	1:A:21:GLY:H	1.62	0.47
1:A:85:ILE:HG12	1:A:91:LEU:HD13	1.96	0.46
1:B:3:MET:HB2	1:B:3:MET:HE2	1.72	0.46
1:B:51:PRO:HB2	1:B:54:TRP:CD1	2.52	0.45
1:A:91:LEU:HB3	2:A:305:HOH:O	2.15	0.45
1:B:71:ILE:HA	1:B:72:PRO:HD3	1.59	0.44
1:A:19:ASN:HD21	1:A:21:GLY:HA3	1.82	0.44
1:A:57:MET:HG3	1:A:58:GLU:O	2.19	0.43
1:A:93:ARG:HG3	1:A:94:PRO:HD2	2.00	0.43
1:A:74:ARG:NH1	2:A:363:HOH:O	2.51	0.43
1:A:87:ARG:HD2	1:A:87:ARG:HH11	1.56	0.42
1:A:17:LEU:O	1:A:29:SER:HA	2.21	0.41
1:A:22:SER:N	2:A:364:HOH:O	2.54	0.41
1:A:123:ASN:HB3	1:B:119:LEU:HD23	2.03	0.41
1:A:50:TRP:HA	1:A:51:PRO:HD3	1.96	0.40
1:B:3:MET:HE1	1:B:11:PRO:HB2	1.97	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	112/124 (90%)	110 (98%)	2 (2%)	0	100	100
1	B	110/124 (89%)	107 (97%)	3 (3%)	0	100	100
All	All	222/248 (90%)	217 (98%)	5 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	100/105 (95%)	90 (90%)	10 (10%)	11	6
1	B	100/105 (95%)	97 (97%)	3 (3%)	53	50
All	All	200/210 (95%)	187 (94%)	13 (6%)	24	17

All (13) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	3	MET
1	A	8	LYS
1	A	19	ASN
1	A	22	SER
1	A	24	PRO
1	A	28	ARG
1	A	58	GLU
1	A	74	ARG
1	A	75	LYS
1	A	105	ARG
1	B	58	GLU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	71	ILE
1	B	94	PRO

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	19	ASN
1	B	7	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 ⓘ

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, 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	115/124 (92%)	-0.76	4 (3%) 42 41	5, 16, 50, 66	0
1	B	113/124 (91%)	-0.93	2 (1%) 65 66	7, 14, 44, 65	0
All	All	228/248 (91%)	-0.85	6 (2%) 53 53	5, 15, 48, 66	0

All (6) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	70	GLY	5.8
1	A	59	ALA	5.4
1	A	70	GLY	4.3
1	A	60	ALA	3.9
1	B	71	ILE	3.0
1	A	71	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 ⓘ

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.