



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 11:54 PM GMT

PDB ID : 1Z0G  
Title : Crystal Structure of *A. fulgidus* Lon proteolytic domain  
Authors : Botos, I.; Melnikov, E.E.; Cherry, S.; Kozlov, S.; Makhovskaya, O.V.; Tropea, J.E.; Gustchina, A.; Rotanova, T.V.; Wlodawer, A.  
Deposited on : 2005-03-01  
Resolution : 2.27 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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

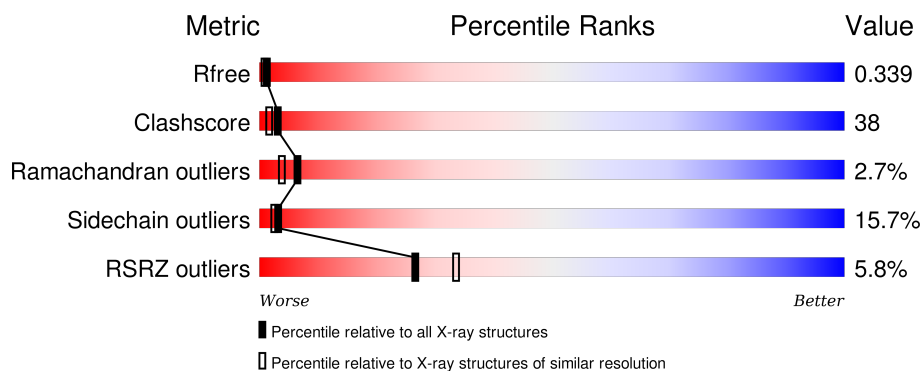
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.27 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	5193 (2.30-2.26)
Clashscore	102246	5929 (2.30-2.26)
Ramachandran outliers	100387	5851 (2.30-2.26)
Sidechain outliers	100360	5850 (2.30-2.26)
RSRZ outliers	91569	5204 (2.30-2.26)

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	205	<div> <div>2%</div> <div>50% 30% 11% 5%</div> </div>
1	B	205	<div> <div>9%</div> <div>39% 37% 16%</div> </div>
1	C	205	<div> <div>3%</div> <div>36% 39% 15% 5%</div> </div>
1	D	205	<div> <div>2%</div> <div>46% 32% 15%</div> </div>
1	E	205	<div> <div>14%</div> <div>40% 39% 13% 5%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	205	 A horizontal bar chart showing the quality of chain F. The bar is divided into five segments: 3% (red), 45% (green), 30% (yellow), 15% (orange), and 6% (red). A small black dot is at the end of the bar.

## 2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 9873 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 Putative protease La homolog type.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	195	Total	C	N	O	S	0	0	0
			1457	921	247	284	5			
1	B	196	Total	C	N	O	S	0	0	0
			1466	926	248	287	5			
1	C	196	Total	C	N	O	S	0	0	0
			1466	926	248	287	5			
1	D	196	Total	C	N	O	S	0	0	0
			1466	926	248	287	5			
1	E	195	Total	C	N	O	S	0	0	0
			1457	921	247	284	5			
1	F	196	Total	C	N	O	S	0	0	0
			1466	926	248	287	5			

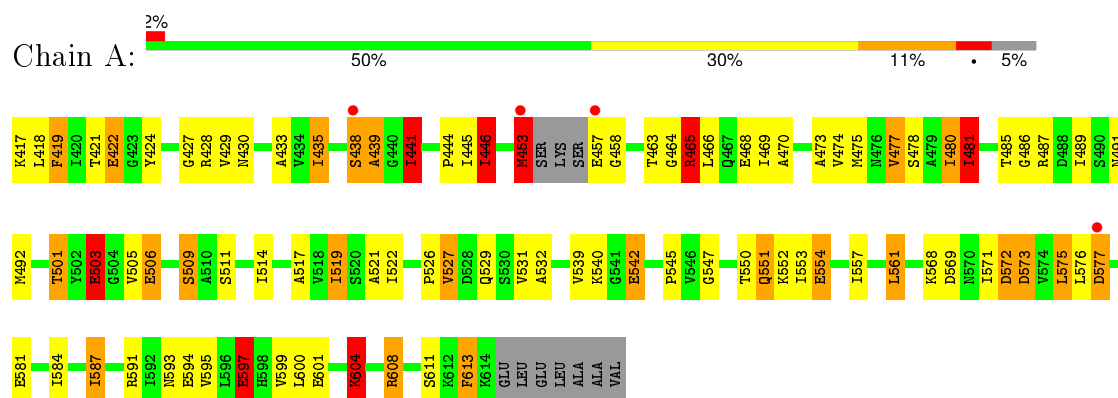
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	163	Total	O	0	0
			163	163		
2	B	165	Total	O	0	0
			165	165		
2	C	198	Total	O	0	0
			198	198		
2	D	197	Total	O	0	0
			197	197		
2	E	147	Total	O	0	0
			147	147		
2	F	225	Total	O	0	0
			225	225		

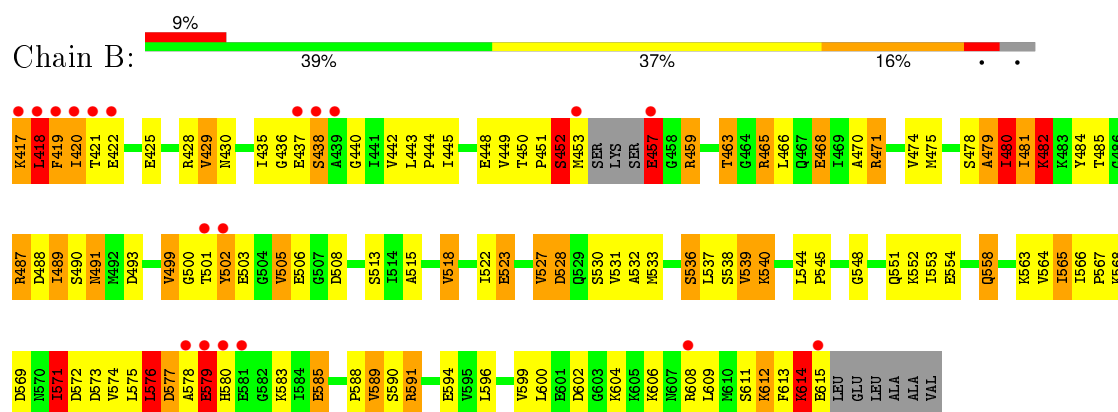
### 3 Residue-property plots [i](#)

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

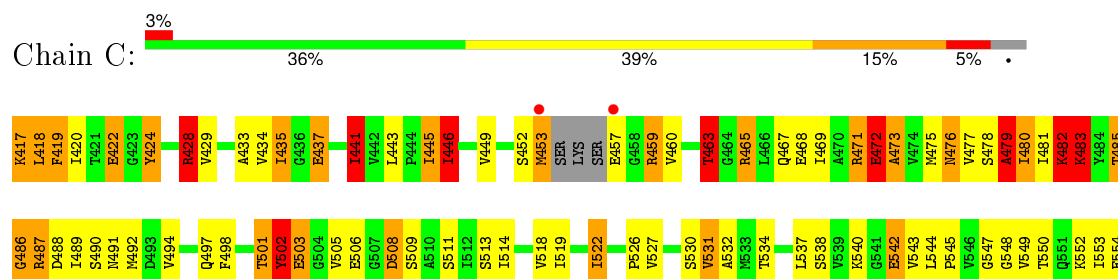
#### • Molecule 1: Putative protease La homolog type



#### • Molecule 1: Putative protease La homolog type

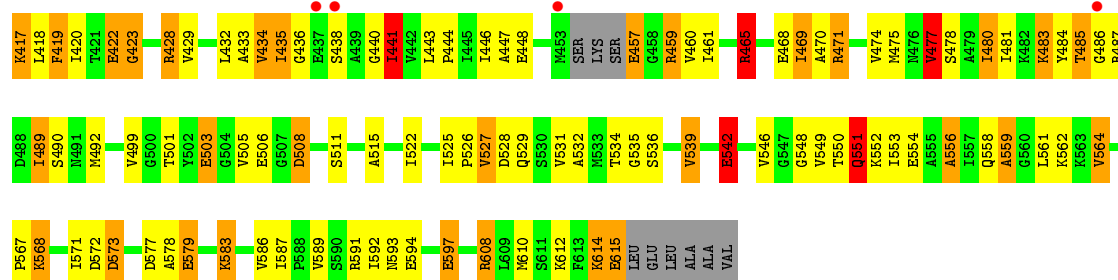


#### • Molecule 1: Putative protease La homolog type

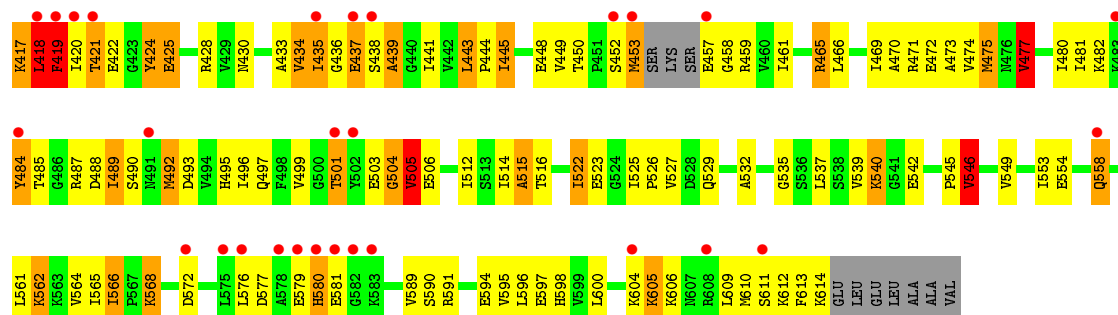




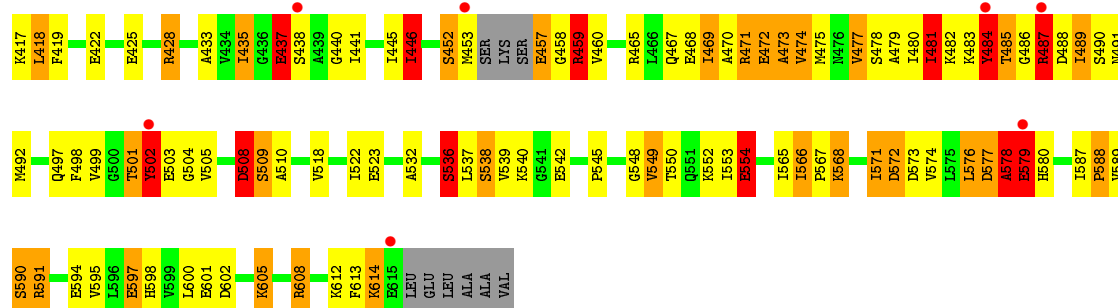
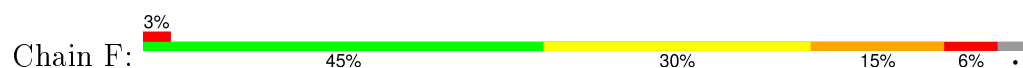
• Molecule 1: Putative protease La homolog type



• Molecule 1: Putative protease La homolog type



• Molecule 1: Putative protease La homolog type



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	85.65Å 88.69Å 147.24Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	15.00 – 2.27 14.97 – 2.27	Depositor EDS
% Data completeness (in resolution range)	100.0 (15.00-2.27) 96.2 (14.97-2.27)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.42 (at 2.27Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.195 , 0.329 0.212 , 0.339	Depositor DCC
$R_{free}$ test set	1038 reflections (2.11%)	DCC
Wilson B-factor (Å <sup>2</sup> )	24.8	Xtriage
Anisotropy	0.582	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 64.1	EDS
Estimated twinning fraction	0.000 for k,h,l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.57$ , $\langle L^2 \rangle = 0.42$	Xtriage
Outliers	5 of 50312 reflections (0.010%)	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	9873	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	14.0	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 57.46 % 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 2.3605e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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	2.00	35/1470 (2.4%)	1.51	19/1983 (1.0%)
1	B	2.11	26/1479 (1.8%)	1.54	24/1995 (1.2%)
1	C	2.23	44/1479 (3.0%)	1.69	30/1995 (1.5%)
1	D	2.13	45/1479 (3.0%)	1.65	25/1995 (1.3%)
1	E	1.83	21/1470 (1.4%)	1.54	17/1983 (0.9%)
1	F	2.08	45/1479 (3.0%)	1.70	26/1995 (1.3%)
All	All	2.07	216/8856 (2.4%)	1.61	141/11946 (1.2%)

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	B	0	1
1	C	0	2
1	D	1	1
1	E	0	3
1	F	1	3
All	All	2	10

The worst 5 of 216 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	579	GLU	CD-OE1	26.73	1.55	1.25
1	B	579	GLU	CD-OE2	20.38	1.48	1.25
1	D	597	GLU	CD-OE1	17.72	1.45	1.25
1	C	597	GLU	CG-CD	16.45	1.76	1.51
1	A	597	GLU	CG-CD	15.68	1.75	1.51

The worst 5 of 141 bond angle outliers are listed below:



Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	471	ARG	NE-CZ-NH2	-14.75	112.92	120.30
1	F	471	ARG	NE-CZ-NH2	-12.66	113.97	120.30
1	F	594	GLU	OE1-CD-OE2	-11.58	109.41	123.30
1	A	600	LEU	CB-CG-CD1	-10.88	92.51	111.00
1	D	577	ASP	CB-CG-OD1	-10.53	108.83	118.30

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	D	615	GLU	CA
1	F	502	TYR	CA

5 of 10 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	435	ILE	Peptide
1	C	562	LYS	Mainchain
1	C	613	PHE	Peptide
1	D	614	LYS	Peptide
1	E	435	ILE	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	1457	0	1523	95	0
1	B	1466	0	1529	136	0
1	C	1466	0	1529	158	0
1	D	1466	0	1529	84	0
1	E	1457	0	1523	113	0
1	F	1466	0	1529	141	0
2	A	163	0	0	21	0
2	B	165	0	0	41	0
2	C	198	0	0	70	2
2	D	197	0	0	34	2
2	E	147	0	0	38	0
2	F	225	0	0	50	0
All	All	9873	0	9162	686	2

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 38.

The worst 5 of 686 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:E:461:ILE:CD1	1:E:461:ILE:CG1	1.75	1.60
1:E:445:ILE:CG1	1:E:445:ILE:CD1	1.76	1.60
1:E:512:ILE:CD1	1:E:512:ILE:CG1	1.74	1.59
1:C:479:ALA:CB	1:C:479:ALA:CA	1.81	1.59
1:F:474:VAL:CB	1:F:474:VAL:CG1	1.76	1.55

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:C:625:HOH:O	2:D:802:HOH:O[4_465]	1.89	0.31
2:C:803:HOH:O	2:D:716:HOH:O[4_465]	2.10	0.10

## 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	191/205 (93%)	176 (92%)	10 (5%)	5 (3%)	7	4
1	B	192/205 (94%)	166 (86%)	15 (8%)	11 (6%)	2	1
1	C	192/205 (94%)	181 (94%)	8 (4%)	3 (2%)	12	10
1	D	192/205 (94%)	177 (92%)	14 (7%)	1 (0%)	34	39
1	E	191/205 (93%)	167 (87%)	16 (8%)	8 (4%)	3	1
1	F	192/205 (94%)	175 (91%)	14 (7%)	3 (2%)	12	10
All	All	1150/1230 (94%)	1042 (91%)	77 (7%)	31 (3%)	6	4

5 of 31 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	438	SER
1	A	439	ALA
1	B	438	SER
1	B	505	VAL
1	B	577	ASP

### 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	160/168 (95%)	136 (85%)	24 (15%)	3	3
1	B	161/168 (96%)	134 (83%)	27 (17%)	2	2
1	C	161/168 (96%)	132 (82%)	29 (18%)	2	1
1	D	161/168 (96%)	145 (90%)	16 (10%)	10	10
1	E	160/168 (95%)	133 (83%)	27 (17%)	2	2
1	F	161/168 (96%)	133 (83%)	28 (17%)	2	2
All	All	964/1008 (96%)	813 (84%)	151 (16%)	3	2

5 of 151 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	502	TYR
1	D	465	ARG
1	F	502	TYR
1	C	509	SER
1	C	597	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	476	ASN
1	D	558	GLN
1	F	476	ASN
1	B	558	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	E	598	HIS

### 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 [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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	195/205 (95%)	-0.25	4 (2%) 67 73	2, 7, 27, 38	0
1	B	196/205 (95%)	0.47	19 (9%) 10 14	3, 13, 35, 54	0
1	C	196/205 (95%)	-0.13	6 (3%) 52 61	2, 6, 27, 41	0
1	D	196/205 (95%)	-0.21	4 (2%) 68 75	2, 7, 25, 40	0
1	E	195/205 (95%)	0.44	28 (14%) 3 5	3, 14, 34, 49	0
1	F	196/205 (95%)	-0.16	7 (3%) 46 54	2, 7, 30, 52	0
All	All	1174/1230 (95%)	0.03	68 (5%) 26 33	2, 10, 32, 54	0

The worst 5 of 68 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	418	LEU	7.9
1	B	578	ALA	6.5
1	B	453	MET	5.9
1	E	418	LEU	5.4
1	E	453	MET	5.3

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

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.