



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 13, 2024 – 02:06 AM EDT

PDB ID : 4CBH  
Title : Pestivirus NS3 helicase  
Authors : Tortorici, M.A.; Duquerroy, S.; Kwok, J.; Vonnrhein, C.; Perez, J.; Lamp, B.;  
Bricogne, G.; Rumenapf, T.; Vachette, P.; Rey, F.A.  
Deposited on : 2013-10-14  
Resolution : 2.51 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

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.20.1
EDS	:	2.36.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

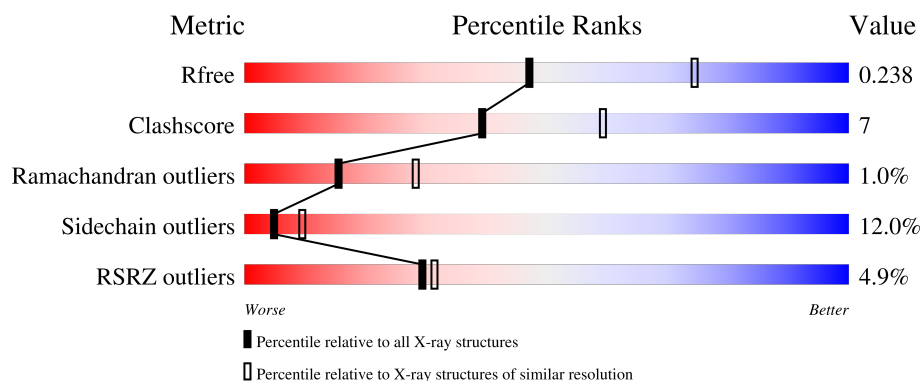
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.51 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))
$R_{free}$	130704	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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 of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. 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	516	<div> <div>3%</div> <div>49%12%•35%</div> </div>
1	B	516	<div> <div>49%12%•36%</div> </div>
1	C	516	<div> <div>3%</div> <div>48%13%•36%</div> </div>
1	D	516	<div> <div>7%</div> <div>56%15%•25%</div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 11813 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 NS3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	333	Total	C	N	O	S	0	0	0
			2664	1693	451	502	18			
1	B	330	Total	C	N	O	S	0	0	0
			2652	1688	448	498	18			
1	C	330	Total	C	N	O	S	0	1	1
			2654	1687	452	497	18			
1	D	385	Total	C	N	O	S	0	1	1
			3075	1948	530	579	18			

There are 68 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	176	MET	-	expression tag	UNP P19712
A	177	ALA	-	expression tag	UNP P19712
A	178	SER	-	expression tag	UNP P19712
A	179	HIS	-	expression tag	UNP P19712
A	180	HIS	-	expression tag	UNP P19712
A	181	HIS	-	expression tag	UNP P19712
A	182	HIS	-	expression tag	UNP P19712
A	183	HIS	-	expression tag	UNP P19712
A	184	HIS	-	expression tag	UNP P19712
A	185	HIS	-	expression tag	UNP P19712
A	186	GLU	-	expression tag	UNP P19712
A	187	ASN	-	expression tag	UNP P19712
A	188	LEU	-	expression tag	UNP P19712
A	189	TYR	-	expression tag	UNP P19712
A	190	PHE	-	expression tag	UNP P19712
A	191	GLN	-	expression tag	UNP P19712
A	192	GLY	-	expression tag	UNP P19712
B	176	MET	-	expression tag	UNP P19712
B	177	ALA	-	expression tag	UNP P19712
B	178	SER	-	expression tag	UNP P19712
B	179	HIS	-	expression tag	UNP P19712

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	180	HIS	-	expression tag	UNP P19712
B	181	HIS	-	expression tag	UNP P19712
B	182	HIS	-	expression tag	UNP P19712
B	183	HIS	-	expression tag	UNP P19712
B	184	HIS	-	expression tag	UNP P19712
B	185	HIS	-	expression tag	UNP P19712
B	186	GLU	-	expression tag	UNP P19712
B	187	ASN	-	expression tag	UNP P19712
B	188	LEU	-	expression tag	UNP P19712
B	189	TYR	-	expression tag	UNP P19712
B	190	PHE	-	expression tag	UNP P19712
B	191	GLN	-	expression tag	UNP P19712
B	192	GLY	-	expression tag	UNP P19712
C	176	MET	-	expression tag	UNP P19712
C	177	ALA	-	expression tag	UNP P19712
C	178	SER	-	expression tag	UNP P19712
C	179	HIS	-	expression tag	UNP P19712
C	180	HIS	-	expression tag	UNP P19712
C	181	HIS	-	expression tag	UNP P19712
C	182	HIS	-	expression tag	UNP P19712
C	183	HIS	-	expression tag	UNP P19712
C	184	HIS	-	expression tag	UNP P19712
C	185	HIS	-	expression tag	UNP P19712
C	186	GLU	-	expression tag	UNP P19712
C	187	ASN	-	expression tag	UNP P19712
C	188	LEU	-	expression tag	UNP P19712
C	189	TYR	-	expression tag	UNP P19712
C	190	PHE	-	expression tag	UNP P19712
C	191	GLN	-	expression tag	UNP P19712
C	192	GLY	-	expression tag	UNP P19712
D	176	MET	-	expression tag	UNP P19712
D	177	ALA	-	expression tag	UNP P19712
D	178	SER	-	expression tag	UNP P19712
D	179	HIS	-	expression tag	UNP P19712
D	180	HIS	-	expression tag	UNP P19712
D	181	HIS	-	expression tag	UNP P19712
D	182	HIS	-	expression tag	UNP P19712
D	183	HIS	-	expression tag	UNP P19712
D	184	HIS	-	expression tag	UNP P19712
D	185	HIS	-	expression tag	UNP P19712
D	186	GLU	-	expression tag	UNP P19712
D	187	ASN	-	expression tag	UNP P19712

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	188	LEU	-	expression tag	UNP P19712
D	189	TYR	-	expression tag	UNP P19712
D	190	PHE	-	expression tag	UNP P19712
D	191	GLN	-	expression tag	UNP P19712
D	192	GLY	-	expression tag	UNP P19712

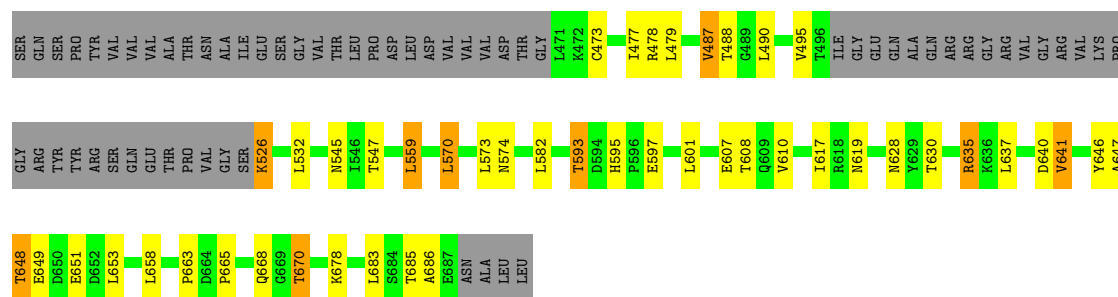
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	197	Total	O	0	0
			197	197		
2	B	187	Total	O	0	0
			187	187		
2	C	154	Total	O	0	0
			154	154		
2	D	230	Total	O	0	0
			230	230		

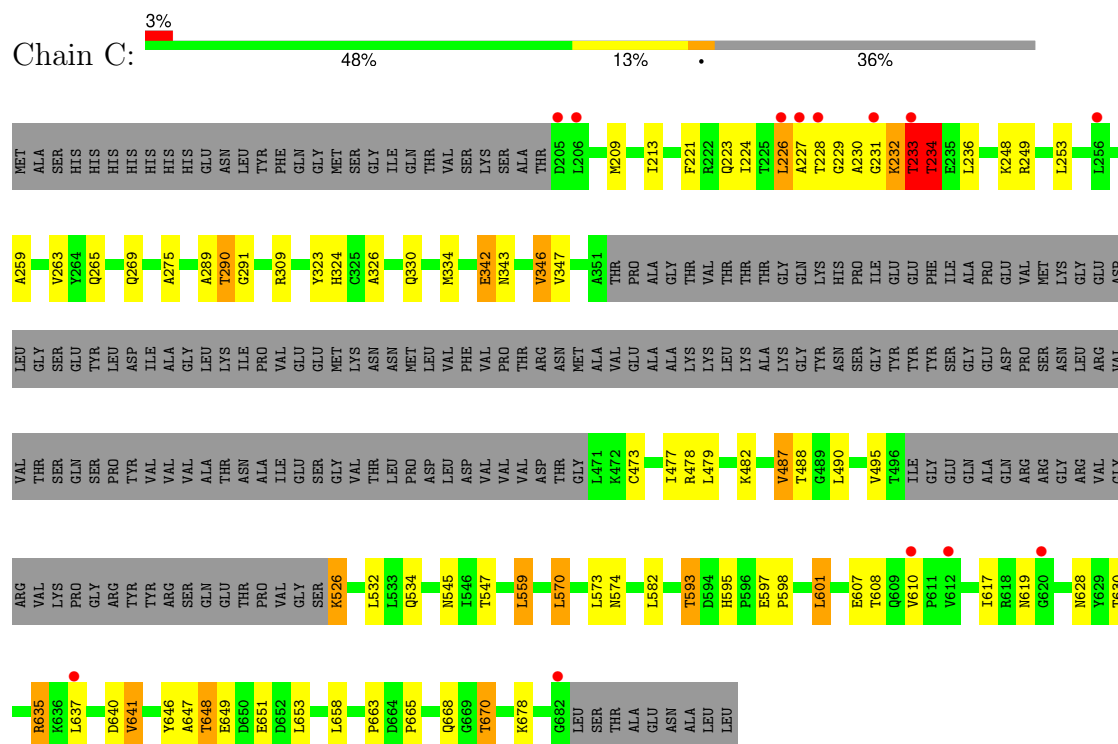


● Molecule 1: SERINE PROTEASE NS3

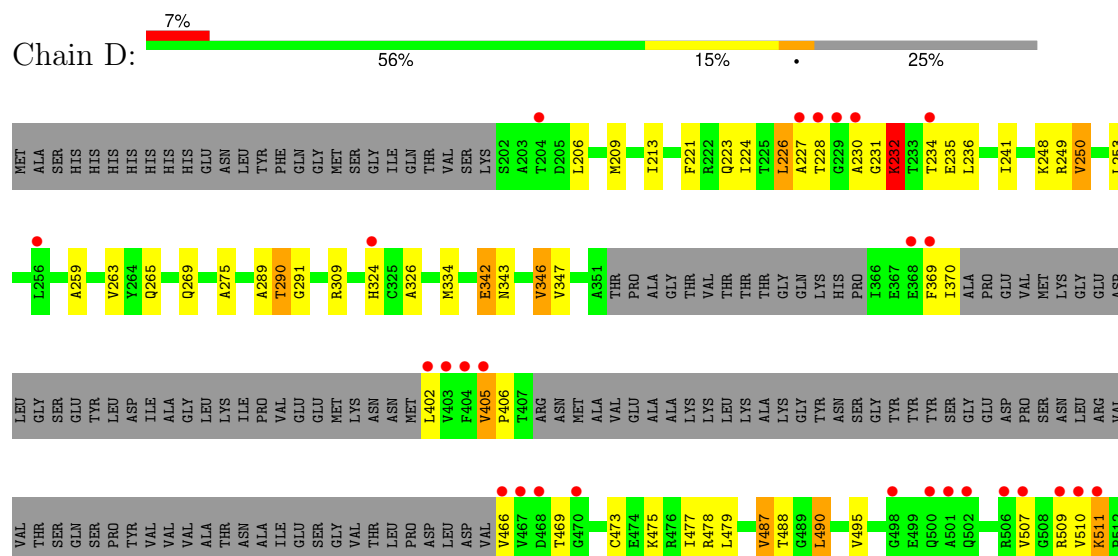


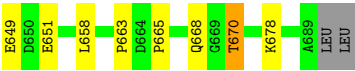
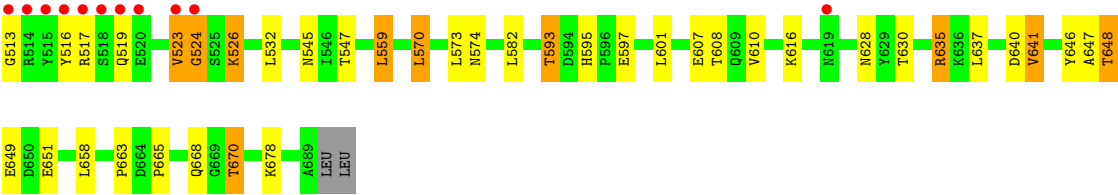


• Molecule 1: SERINE PROTEASE NS3



• Molecule 1: SERINE PROTEASE NS3







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	70.07Å 144.44Å 118.69Å 90.00° 92.94° 90.00°	Depositor
Resolution (Å)	72.22 – 2.51 72.22 – 2.51	Depositor EDS
% Data completeness (in resolution range)	88.6 (72.22-2.51) 88.7 (72.22-2.51)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.37 (at 2.51Å)	Xtriage
Refinement program	BUSTER 2.13.0	Depositor
R, $R_{free}$	0.203 , 0.232 0.215 , 0.238	Depositor DCC
$R_{free}$ test set	3572 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	49.1	Xtriage
Anisotropy	0.174	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 64.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.048 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	11813	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	60.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.18% 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, 0.333 respectively for untwinned datasets, and 0.375, 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	0.53	0/2715	0.81	3/3671 (0.1%)
1	B	0.52	0/2702	0.77	2/3651 (0.1%)
1	C	0.54	0/2705	0.81	4/3657 (0.1%)
1	D	0.55	0/3132	0.78	1/4234 (0.0%)
All	All	0.53	0/11254	0.79	10/15213 (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	2
1	C	0	2
All	All	0	4

There are no bond length outliers.

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	233	THR	C-N-CA	9.71	145.98	121.70
1	C	233	THR	C-N-CA	9.48	145.40	121.70
1	A	231	GLY	C-N-CA	6.41	137.72	121.70
1	C	229	GLY	C-N-CA	5.94	136.56	121.70
1	A	290	THR	N-CA-C	5.81	126.67	111.00

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	233	THR	Peptide,Mainchain
1	C	233	THR	Peptide,Mainchain

## 5.2 Too-close contacts [i](#)

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	2664	0	2666	43	0
1	B	2652	0	2658	41	0
1	C	2654	0	2658	46	0
1	D	3075	0	3078	46	0
2	A	197	0	0	3	0
2	B	187	0	0	1	0
2	C	154	0	0	3	0
2	D	230	0	0	2	0
All	All	11813	0	11060	164	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 7.

The worst 5 of 164 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:A:682:GLY:O	1:B:685:THR:CG2	1.75	1.35
1:A:324:HIS:CD2	1:A:350:THR:OG1	2.02	1.11
1:A:682:GLY:O	1:B:685:THR:HG22	0.93	1.09
1:A:324:HIS:HD2	1:A:350:THR:OG1	1.53	0.91
1:B:234:THR:HG21	1:B:263:VAL:HG11	1.55	0.88

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	327/516 (63%)	315 (96%)	7 (2%)	5 (2%)	10	18
1	B	322/516 (62%)	316 (98%)	4 (1%)	2 (1%)	25	43
1	C	325/516 (63%)	314 (97%)	7 (2%)	4 (1%)	13	24
1	D	378/516 (73%)	364 (96%)	11 (3%)	3 (1%)	19	35
All	All	1352/2064 (66%)	1309 (97%)	29 (2%)	14 (1%)	15	28

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	232	LYS
1	C	495	VAL
1	D	523	VAL
1	C	234	THR
1	D	524	GLY

### 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	289/442 (65%)	254 (88%)	35 (12%)	5	9
1	B	289/442 (65%)	256 (89%)	33 (11%)	5	11
1	C	288/442 (65%)	256 (89%)	32 (11%)	6	11
1	D	333/442 (75%)	289 (87%)	44 (13%)	4	7
All	All	1199/1768 (68%)	1055 (88%)	144 (12%)	5	9

5 of 144 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	469	THR
1	D	678	LYS
1	D	495	VAL
1	D	593	THR
1	B	532	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 26 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	534	GLN
1	C	600	GLN
1	D	600	GLN
1	C	595	HIS
1	C	628	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 monosaccharides 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	333/516 (64%)	0.36	14 (4%) 36 39	33, 53, 112, 149	0
1	B	330/516 (63%)	0.24	2 (0%) 89 90	30, 54, 105, 125	0
1	C	330/516 (63%)	0.38	13 (3%) 39 42	33, 56, 101, 129	0
1	D	385/516 (74%)	0.58	38 (9%) 7 7	31, 54, 117, 146	0
All	All	1378/2064 (66%)	0.40	67 (4%) 29 31	30, 54, 111, 149	0

The worst 5 of 67 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	228	THR	9.4
1	D	516	TYR	8.1
1	D	515	TYR	8.0
1	D	523	VAL	7.5
1	D	227	ALA	7.5

### 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 monosaccharides 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.