



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2026 – 12:42 PM UTC

PDB ID : 3TPC / pdb\_00003tpc  
Title : Crystal structure of a hypothetical protein SMa1452 from Sinorhizobium meliloti 1021  
Authors : Agarwal, R.; Chamala, S.; Evans, B.; Foti, R.; Gizzi, A.; Hillerich, B.; Kar, A.; LaFleur, J.; Seidel, R.; Villigas, G.; Zencheck, W.; Almo, S.C.; Swaminathan, S.; New York Structural Genomics Research Consortium (NYSGRG)  
Deposited on : 2011-09-07  
Resolution : 2.34 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	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

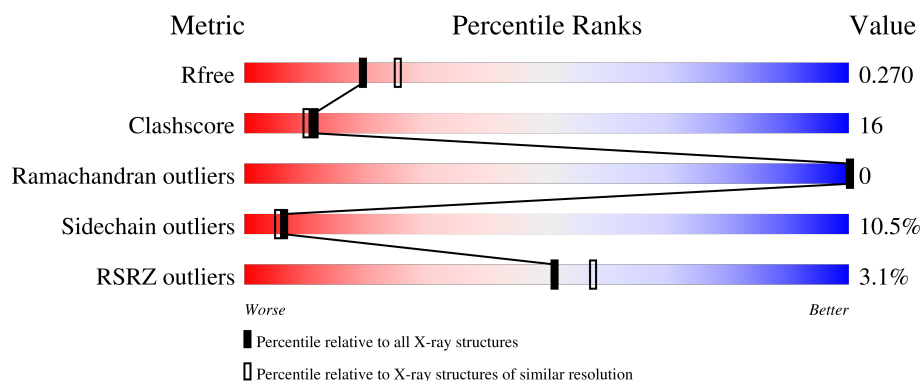
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.34 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))
$R_{free}$	180053	3031 (2.36-2.32)
Clashscore	190562	3127 (2.36-2.32)
Ramachandran outliers	187476	3095 (2.36-2.32)
Sidechain outliers	187428	3095 (2.36-2.32)
RSRZ outliers	180081	3033 (2.36-2.32)

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	257	<div> <div>4%</div> <div>65%18%12%</div> </div>
1	B	257	<div> <div>4%</div> <div>56%24%5%14%</div> </div>
1	C	257	<div> <div>2%</div> <div>61%21%13%</div> </div>
1	D	257	<div> <div>3%</div> <div>61%23%12%</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	E	257	<div><div><div></div><div></div><div></div><div></div></div><div>4%</div><div>64%</div><div>19%</div><div>•</div><div>12%</div></div>
1	F	257	<div><div><div></div><div></div><div></div><div></div></div><div>2%</div><div>61%</div><div>21%</div><div>• •</div><div>13%</div></div>
1	G	257	<div><div><div></div><div></div><div></div><div></div></div><div>2%</div><div>64%</div><div>19%</div><div>•</div><div>14%</div></div>
1	H	257	<div><div><div></div><div></div><div></div><div></div></div><div>5%</div><div>66%</div><div>16%</div><div>•</div><div>14%</div></div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 13015 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 Short chain alcohol dehydrogenase-related dehydrogenase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	225	Total	C	N	O	S	Se	0	0	0
			1627	1022	295	303	2	5			
1	B	221	Total	C	N	O	S	Se	0	0	0
			1573	989	280	297	2	5			
1	C	224	Total	C	N	O	S	Se	0	0	0
			1604	1007	288	302	2	5			
1	D	226	Total	C	N	O	S	Se	0	0	0
			1597	1007	282	301	2	5			
1	E	225	Total	C	N	O	S	Se	0	0	0
			1624	1021	295	301	2	5			
1	F	223	Total	C	N	O	S	Se	0	0	0
			1593	1000	289	297	2	5			
1	G	222	Total	C	N	O	S	Se	0	0	0
			1542	968	275	292	2	5			
1	H	222	Total	C	N	O	S	Se	0	0	0
			1592	997	290	298	2	5			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	-	expression tag	UNP Q92YS1
A	2	VAL	-	expression tag	UNP Q92YS1
B	1	MSE	-	expression tag	UNP Q92YS1
B	2	VAL	-	expression tag	UNP Q92YS1
C	1	MSE	-	expression tag	UNP Q92YS1
C	2	VAL	-	expression tag	UNP Q92YS1
D	1	MSE	-	expression tag	UNP Q92YS1
D	2	VAL	-	expression tag	UNP Q92YS1
E	1	MSE	-	expression tag	UNP Q92YS1
E	2	VAL	-	expression tag	UNP Q92YS1
F	1	MSE	-	expression tag	UNP Q92YS1
F	2	VAL	-	expression tag	UNP Q92YS1
G	1	MSE	-	expression tag	UNP Q92YS1

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
G	2	VAL	-	expression tag	UNP Q92YS1
H	1	MSE	-	expression tag	UNP Q92YS1
H	2	VAL	-	expression tag	UNP Q92YS1

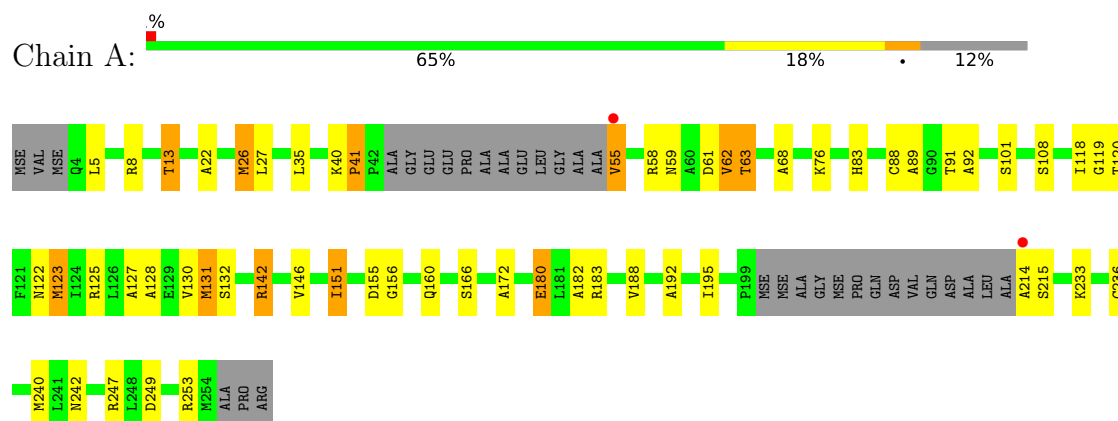
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	32	Total 32	O 32	0	0
2	B	36	Total 36	O 36	0	0
2	C	42	Total 42	O 42	0	0
2	D	39	Total 39	O 39	0	0
2	E	41	Total 41	O 41	0	0
2	F	30	Total 30	O 30	0	0
2	G	24	Total 24	O 24	0	0
2	H	19	Total 19	O 19	0	0

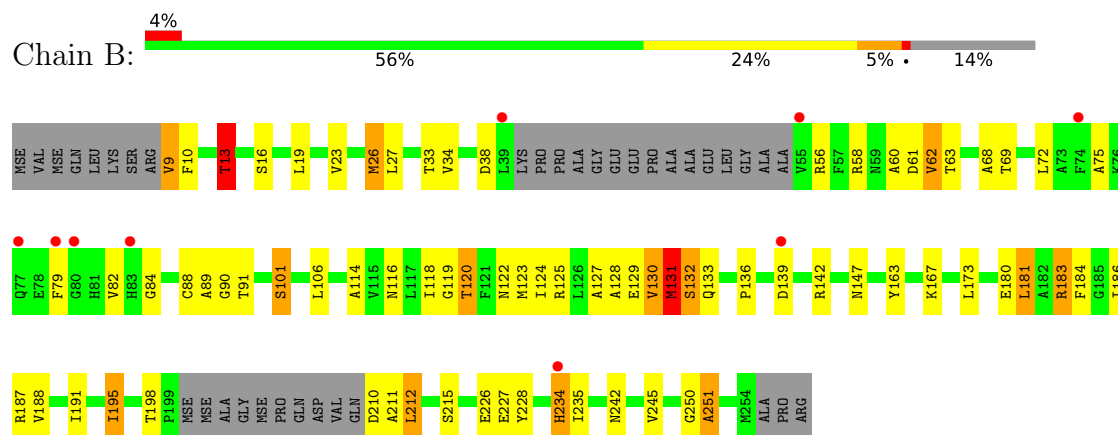
### 3 Residue-property plots [i](#)

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

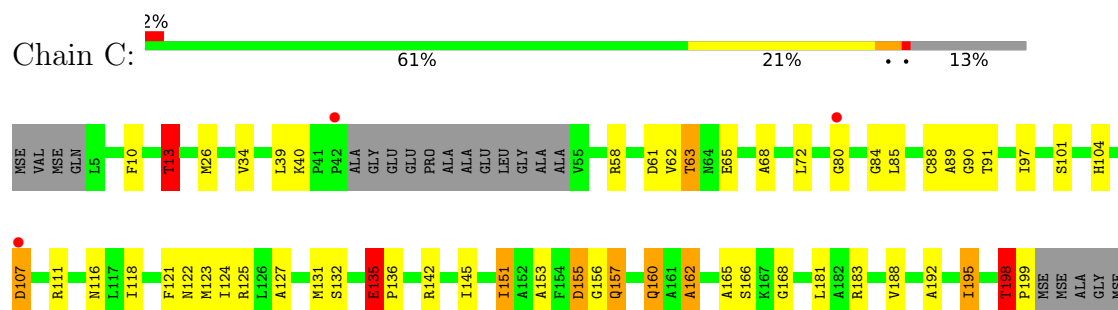
- Molecule 1: Short chain alcohol dehydrogenase-related dehydrogenase



- Molecule 1: Short chain alcohol dehydrogenase-related dehydrogenase



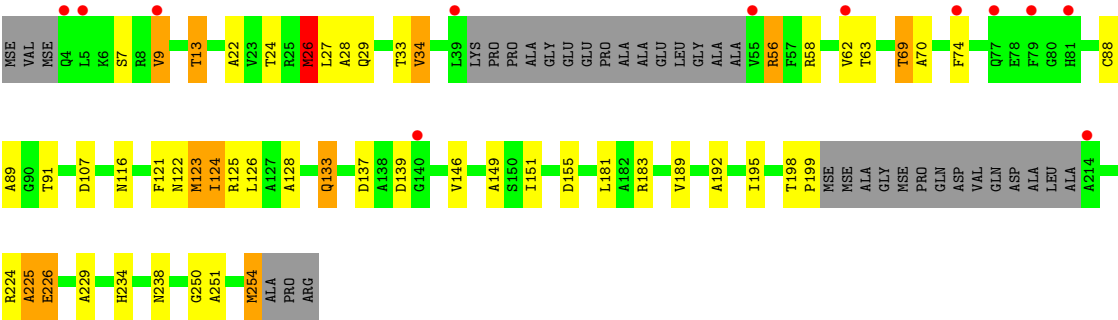
- Molecule 1: Short chain alcohol dehydrogenase-related dehydrogenase







● Molecule 1: Short chain alcohol dehydrogenase-related dehydrogenase





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	69.15Å 69.38Å 121.85Å 98.36° 92.72° 113.01°	Depositor
Resolution (Å)	46.20 – 2.34 46.20 – 2.34	Depositor EDS
% Data completeness (in resolution range)	97.4 (46.20-2.34) 97.4 (46.20-2.34)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.59 (at 2.34Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, $R_{free}$	0.212 , 0.270 0.212 , 0.270	Depositor DCC
$R_{free}$ test set	4193 reflections (4.88%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	35.5	Xtriage
Anisotropy	0.063	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 36.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	13015	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.20% 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	1.33	6/1646 (0.4%)	1.30	15/2223 (0.7%)
1	B	1.23	6/1589 (0.4%)	1.29	12/2149 (0.6%)
1	C	1.32	6/1622 (0.4%)	1.26	11/2193 (0.5%)
1	D	1.22	3/1616 (0.2%)	1.27	9/2189 (0.4%)
1	E	1.26	4/1643 (0.2%)	1.22	9/2219 (0.4%)
1	F	1.19	2/1610 (0.1%)	1.23	8/2174 (0.4%)
1	G	1.16	2/1558 (0.1%)	1.22	5/2112 (0.2%)
1	H	1.19	3/1609 (0.2%)	1.19	13/2173 (0.6%)
All	All	1.24	32/12893 (0.2%)	1.25	82/17432 (0.5%)

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	B	0	1
1	C	0	1
All	All	0	2

The worst 5 of 32 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	131	MSE	SE-CE	-7.36	1.73	1.95
1	C	121	PHE	C-O	-6.72	1.16	1.24
1	B	251	ALA	CA-CB	6.49	1.62	1.53
1	C	118	ILE	C-O	-6.36	1.16	1.24
1	B	198	THR	C-O	-6.21	1.18	1.24

The worst 5 of 82 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	130	VAL	N-CA-C	10.23	121.06	110.72

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	H	34	VAL	N-CA-C	8.77	120.75	107.77
1	E	149	ALA	N-CA-C	-8.68	101.66	108.78
1	A	40	LYS	CA-C-N	8.44	125.73	119.66
1	A	40	LYS	C-N-CA	8.44	125.73	119.66

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	211	ALA	Peptide
1	C	198	THR	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	1627	0	1636	44	0
1	B	1573	0	1566	84	0
1	C	1604	0	1598	44	0
1	D	1597	0	1578	67	0
1	E	1624	0	1634	47	0
1	F	1593	0	1586	49	0
1	G	1542	0	1498	48	0
1	H	1592	0	1581	48	0
2	A	32	0	0	4	0
2	B	36	0	0	3	0
2	C	42	0	0	2	0
2	D	39	0	0	1	0
2	E	41	0	0	2	0
2	F	30	0	0	2	0
2	G	24	0	0	0	0
2	H	19	0	0	0	0
All	All	13015	0	12677	408	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 16.

The worst 5 of 408 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:234:HIS:ND1	1:D:234:HIS:HB3	1.37	1.37
1:B:234:HIS:CE1	1:D:234:HIS:HB3	1.74	1.20
1:G:68:ALA:CB	1:G:123:MSE:HE1	1.74	1.18
1:A:123:MSE:HA	1:A:123:MSE:CE	1.74	1.17
1:G:68:ALA:HB1	1:G:123:MSE:HE1	1.28	1.14

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	219/257 (85%)	213 (97%)	6 (3%)	0	100	100
1	B	215/257 (84%)	206 (96%)	9 (4%)	0	100	100
1	C	218/257 (85%)	208 (95%)	10 (5%)	0	100	100
1	D	220/257 (86%)	211 (96%)	9 (4%)	0	100	100
1	E	219/257 (85%)	206 (94%)	13 (6%)	0	100	100
1	F	217/257 (84%)	204 (94%)	13 (6%)	0	100	100
1	G	216/257 (84%)	202 (94%)	14 (6%)	0	100	100
1	H	216/257 (84%)	204 (94%)	12 (6%)	0	100	100
All	All	1740/2056 (85%)	1654 (95%)	86 (5%)	0	100	100

There are no Ramachandran outliers to report.

#### 5.3.2 Protein sidechains [i](#)

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	160/173 (92%)	144 (90%)	16 (10%)	7	6
1	B	151/173 (87%)	135 (89%)	16 (11%)	6	5
1	C	156/173 (90%)	139 (89%)	17 (11%)	6	5
1	D	153/173 (88%)	135 (88%)	18 (12%)	5	4
1	E	159/173 (92%)	143 (90%)	16 (10%)	7	6
1	F	153/173 (88%)	139 (91%)	14 (9%)	8	8
1	G	143/173 (83%)	128 (90%)	15 (10%)	6	5
1	H	153/173 (88%)	136 (89%)	17 (11%)	6	4
All	All	1228/1384 (89%)	1099 (90%)	129 (10%)	6	5

5 of 129 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	H	56	ARG
1	H	91	THR
1	C	217	PRO
1	C	195	ILE
1	H	123	MSE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 37 such sidechains are listed below:

Mol	Chain	Res	Type
1	G	234	HIS
1	H	238	ASN
1	G	242	ASN
1	H	122	ASN
1	D	122	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules 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 [i](#)

There are no oligosaccharides 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	220/257 (85%)	-0.11	2 (0%) 81 84	17, 29, 45, 58	0
1	B	216/257 (84%)	0.12	9 (4%) 40 46	18, 32, 50, 58	0
1	C	219/257 (85%)	0.03	4 (1%) 67 72	17, 30, 49, 60	0
1	D	221/257 (85%)	0.11	7 (3%) 50 57	19, 32, 52, 61	0
1	E	220/257 (85%)	0.27	10 (4%) 38 44	22, 35, 59, 69	0
1	F	218/257 (84%)	0.10	6 (2%) 55 61	21, 33, 48, 64	0
1	G	217/257 (84%)	0.29	5 (2%) 61 66	21, 38, 55, 61	0
1	H	217/257 (84%)	0.44	12 (5%) 30 36	23, 39, 63, 68	0
All	All	1748/2056 (85%)	0.16	55 (3%) 51 58	17, 33, 54, 69	0

The worst 5 of 55 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	42	PRO	3.9
1	H	4	GLN	3.9
1	H	214	ALA	3.6
1	B	74	PHE	3.6
1	F	214	ALA	3.5

### 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 oligosaccharides 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.