



# Full wwPDB X-ray Structure Validation Report ⓘ

May 23, 2020 – 02:42 pm BST

PDB ID : 4LCL  
Title : Simvastatin Synthase (LOVD), from *Aspergillus Terreus*, LovD6 mutant (simh6208)  
Authors : Gao, X.; Sawaya, M.R.; Yeates, T.O.; Tang, Y.  
Deposited on : 2013-06-21  
Resolution : 1.80 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

<https://www.wwpdb.org/validation/2017/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 : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.11  
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) : 2.11

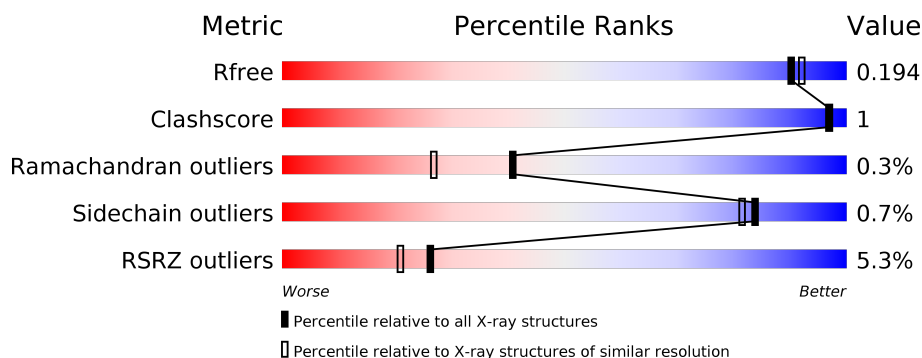
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.80 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	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)
RSRZ outliers	127900	5850 (1.80-1.80)

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 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	433	<div> <div>4%</div> <div>89%</div> <div>8%</div> </div>
1	B	433	<div> <div>6%</div> <div>92%</div> <div>6%</div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 6786 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 Transesterase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	400	Total	C	N	O	S	0	5	0
			3181	2001	571	586	23			
1	B	409	Total	C	N	O	S	0	2	0
			3214	2020	576	595	23			

There are 84 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	EXPRESSION TAG	UNP Q9Y7D1
A	-18	GLY	-	EXPRESSION TAG	UNP Q9Y7D1
A	-17	SER	-	EXPRESSION TAG	UNP Q9Y7D1
A	-16	SER	-	EXPRESSION TAG	UNP Q9Y7D1
A	-15	HIS	-	EXPRESSION TAG	UNP Q9Y7D1
A	-14	HIS	-	EXPRESSION TAG	UNP Q9Y7D1
A	-13	HIS	-	EXPRESSION TAG	UNP Q9Y7D1
A	-12	HIS	-	EXPRESSION TAG	UNP Q9Y7D1
A	-11	HIS	-	EXPRESSION TAG	UNP Q9Y7D1
A	-10	HIS	-	EXPRESSION TAG	UNP Q9Y7D1
A	-9	SER	-	EXPRESSION TAG	UNP Q9Y7D1
A	-8	SER	-	EXPRESSION TAG	UNP Q9Y7D1
A	-7	GLY	-	EXPRESSION TAG	UNP Q9Y7D1
A	-6	LEU	-	EXPRESSION TAG	UNP Q9Y7D1
A	-5	VAL	-	EXPRESSION TAG	UNP Q9Y7D1
A	-4	PRO	-	EXPRESSION TAG	UNP Q9Y7D1
A	-3	ARG	-	EXPRESSION TAG	UNP Q9Y7D1
A	-2	GLY	-	EXPRESSION TAG	UNP Q9Y7D1
A	-1	SER	-	EXPRESSION TAG	UNP Q9Y7D1
A	0	HIS	-	EXPRESSION TAG	UNP Q9Y7D1
A	9	VAL	ALA	ENGINEERED MUTATION	UNP Q9Y7D1
A	26	GLU	LYS	ENGINEERED MUTATION	UNP Q9Y7D1
A	40	ALA	CYS	ENGINEERED MUTATION	UNP Q9Y7D1
A	43	ARG	ASN	ENGINEERED MUTATION	UNP Q9Y7D1
A	60	ASN	CYS	ENGINEERED MUTATION	UNP Q9Y7D1

*Continued on next page...*

*Continued from previous page...*

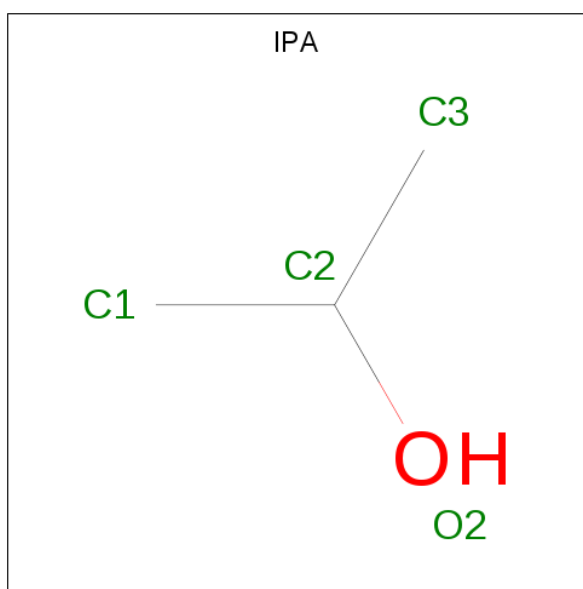
Chain	Residue	Modelled	Actual	Comment	Reference
A	123	PRO	ALA	ENGINEERED MUTATION	UNP Q9Y7D1
A	157	VAL	MET	ENGINEERED MUTATION	UNP Q9Y7D1
A	164	GLY	SER	ENGINEERED MUTATION	UNP Q9Y7D1
A	172	ASN	SER	ENGINEERED MUTATION	UNP Q9Y7D1
A	174	PHE	LEU	ENGINEERED MUTATION	UNP Q9Y7D1
A	178	LEU	ALA	ENGINEERED MUTATION	UNP Q9Y7D1
A	191	GLY	ASN	ENGINEERED MUTATION	UNP Q9Y7D1
A	192	ILE	LEU	ENGINEERED MUTATION	UNP Q9Y7D1
A	241	MET	GLN	ENGINEERED MUTATION	UNP Q9Y7D1
A	247	SER	ALA	ENGINEERED MUTATION	UNP Q9Y7D1
A	250	LYS	ARG	ENGINEERED MUTATION	UNP Q9Y7D1
A	275	SER	GLY	ENGINEERED MUTATION	UNP Q9Y7D1
A	297	GLU	GLN	ENGINEERED MUTATION	UNP Q9Y7D1
A	361	MET	LEU	ENGINEERED MUTATION	UNP Q9Y7D1
A	370	ILE	VAL	ENGINEERED MUTATION	UNP Q9Y7D1
A	383	VAL	ALA	ENGINEERED MUTATION	UNP Q9Y7D1
A	404	LYS	HIS	ENGINEERED MUTATION	UNP Q9Y7D1
B	-19	MET	-	EXPRESSION TAG	UNP Q9Y7D1
B	-18	GLY	-	EXPRESSION TAG	UNP Q9Y7D1
B	-17	SER	-	EXPRESSION TAG	UNP Q9Y7D1
B	-16	SER	-	EXPRESSION TAG	UNP Q9Y7D1
B	-15	HIS	-	EXPRESSION TAG	UNP Q9Y7D1
B	-14	HIS	-	EXPRESSION TAG	UNP Q9Y7D1
B	-13	HIS	-	EXPRESSION TAG	UNP Q9Y7D1
B	-12	HIS	-	EXPRESSION TAG	UNP Q9Y7D1
B	-11	HIS	-	EXPRESSION TAG	UNP Q9Y7D1
B	-10	HIS	-	EXPRESSION TAG	UNP Q9Y7D1
B	-9	SER	-	EXPRESSION TAG	UNP Q9Y7D1
B	-8	SER	-	EXPRESSION TAG	UNP Q9Y7D1
B	-7	GLY	-	EXPRESSION TAG	UNP Q9Y7D1
B	-6	LEU	-	EXPRESSION TAG	UNP Q9Y7D1
B	-5	VAL	-	EXPRESSION TAG	UNP Q9Y7D1
B	-4	PRO	-	EXPRESSION TAG	UNP Q9Y7D1
B	-3	ARG	-	EXPRESSION TAG	UNP Q9Y7D1
B	-2	GLY	-	EXPRESSION TAG	UNP Q9Y7D1
B	-1	SER	-	EXPRESSION TAG	UNP Q9Y7D1
B	0	HIS	-	EXPRESSION TAG	UNP Q9Y7D1
B	9	VAL	ALA	ENGINEERED MUTATION	UNP Q9Y7D1
B	26	GLU	LYS	ENGINEERED MUTATION	UNP Q9Y7D1
B	40	ALA	CYS	ENGINEERED MUTATION	UNP Q9Y7D1
B	43	ARG	ASN	ENGINEERED MUTATION	UNP Q9Y7D1
B	60	ASN	CYS	ENGINEERED MUTATION	UNP Q9Y7D1

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	123	PRO	ALA	ENGINEERED MUTATION	UNP Q9Y7D1
B	157	VAL	MET	ENGINEERED MUTATION	UNP Q9Y7D1
B	164	GLY	SER	ENGINEERED MUTATION	UNP Q9Y7D1
B	172	ASN	SER	ENGINEERED MUTATION	UNP Q9Y7D1
B	174	PHE	LEU	ENGINEERED MUTATION	UNP Q9Y7D1
B	178	LEU	ALA	ENGINEERED MUTATION	UNP Q9Y7D1
B	191	GLY	ASN	ENGINEERED MUTATION	UNP Q9Y7D1
B	192	ILE	LEU	ENGINEERED MUTATION	UNP Q9Y7D1
B	241	MET	GLN	ENGINEERED MUTATION	UNP Q9Y7D1
B	247	SER	ALA	ENGINEERED MUTATION	UNP Q9Y7D1
B	250	LYS	ARG	ENGINEERED MUTATION	UNP Q9Y7D1
B	275	SER	GLY	ENGINEERED MUTATION	UNP Q9Y7D1
B	297	GLU	GLN	ENGINEERED MUTATION	UNP Q9Y7D1
B	361	MET	LEU	ENGINEERED MUTATION	UNP Q9Y7D1
B	370	ILE	VAL	ENGINEERED MUTATION	UNP Q9Y7D1
B	383	VAL	ALA	ENGINEERED MUTATION	UNP Q9Y7D1
B	404	LYS	HIS	ENGINEERED MUTATION	UNP Q9Y7D1

- Molecule 2 is ISOPROPYL ALCOHOL (three-letter code: IPA) (formula: C<sub>3</sub>H<sub>8</sub>O).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			4	3	1		
2	B	1	Total	C	O	0	0
			4	3	1		

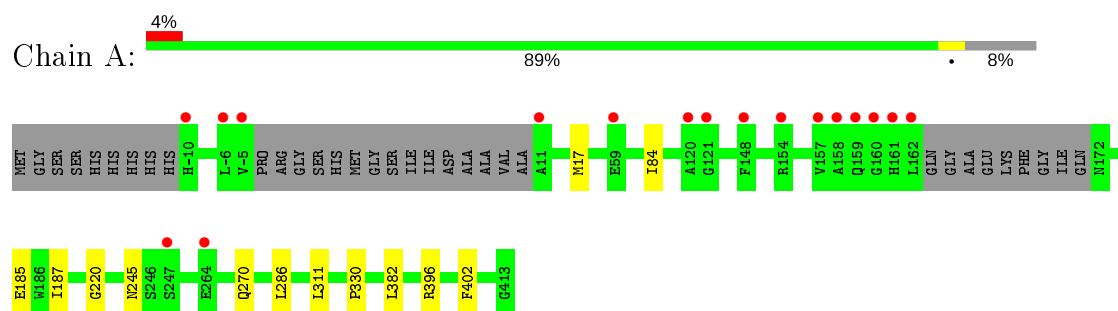
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	219	Total 219	O 219	0	0
3	B	164	Total 164	O 164	0	0

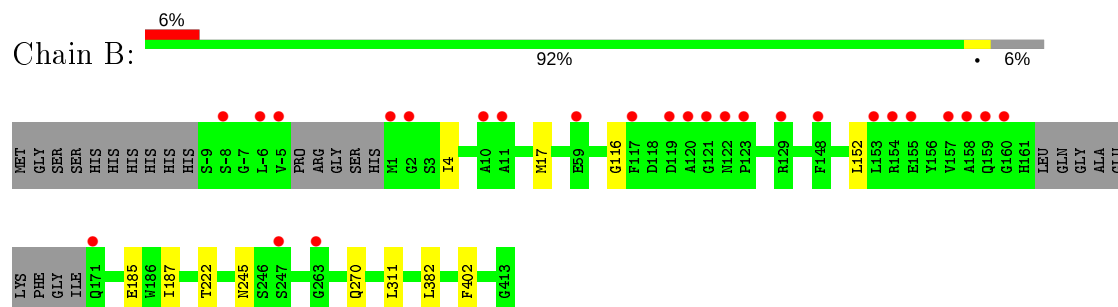
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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 ( $\text{RSRZ} > 2$ ). 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: Transesterase



- Molecule 1: Transesterase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	43.80Å 76.39Å 113.76Å 90.00° 93.91° 90.00°	Depositor
Resolution (Å)	18.10 – 1.80 18.10 – 1.80	Depositor EDS
% Data completeness (in resolution range)	99.5 (18.10-1.80) 99.5 (18.10-1.80)	Depositor EDS
$R_{merge}$	0.04	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.18 (at 1.80Å)	Xtriage
Refinement program	BUSTER-TNT, BUSTER 2.10.0	Depositor
R, $R_{free}$	0.167 , 0.191 0.169 , 0.194	Depositor DCC
$R_{free}$ test set	3448 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	19.6	Xtriage
Anisotropy	0.202	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.42 , 55.1	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	0.96	EDS
Total number of atoms	6786	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	28.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.04% 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

Bond lengths and bond angles in the following residue types are not validated in this section: IPA

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	0.52	0/3263	0.63	0/4412
1	B	0.49	0/3286	0.64	0/4445
All	All	0.50	0/6549	0.63	0/8857

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 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	3181	0	3162	7	0
1	B	3214	0	3192	6	0
2	A	4	0	8	0	0
2	B	4	0	8	0	0
3	A	219	0	0	0	0
3	B	164	0	0	0	0
All	All	6786	0	6370	12	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 1.

All (12) 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:220:GLY:O	1:B:222:THR:HG21	1.98	0.63
1:B:17:MET:HG3	1:B:402:PHE:HA	1.91	0.53
1:A:17:MET:HG3	1:A:402:PHE:HA	1.92	0.51
1:A:185:GLU:HA	1:A:311:LEU:CD2	2.41	0.50
1:B:185:GLU:HA	1:B:311:LEU:CD2	2.41	0.49
1:A:330:PRO:HA	1:A:396[B]:ARG:HG3	1.97	0.46
1:B:116:GLY:HA2	1:B:152:LEU:HD22	1.96	0.46
1:A:185:GLU:HA	1:A:311:LEU:HD22	1.99	0.44
1:B:185:GLU:HG3	1:B:187:ILE:HG13	2.00	0.42
1:B:185:GLU:HA	1:B:311:LEU:HD22	2.00	0.42
1:A:84:ILE:HD11	1:A:286:LEU:HD11	2.02	0.41
1:A:185:GLU:HG3	1:A:187:ILE:HG13	2.02	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	399/433 (92%)	387 (97%)	11 (3%)	1 (0%)	41	27
1	B	405/433 (94%)	391 (96%)	13 (3%)	1 (0%)	47	33
All	All	804/866 (93%)	778 (97%)	24 (3%)	2 (0%)	41	33

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	270	GLN
1	B	270	GLN

### 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	342/361 (95%)	340 (99%)	2 (1%)	86	84
1	B	343/361 (95%)	340 (99%)	3 (1%)	78	75
All	All	685/722 (95%)	680 (99%)	5 (1%)	84	81

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	245	ASN
1	A	382	LEU
1	B	4	ILE
1	B	245	ASN
1	B	382	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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](#)

2 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
2	IPA	B	501	-	3,3,3	0.69	0	3,3,3	0.35	0
2	IPA	A	501	-	3,3,3	0.68	0	3,3,3	0.30	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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

### 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	400/433 (92%)	-0.04	17 (4%) 35 29	11, 21, 44, 92	0
1	B	409/433 (94%)	0.28	26 (6%) 19 15	13, 27, 58, 94	0
All	All	809/866 (93%)	0.12	43 (5%) 26 21	11, 24, 54, 94	0

All (43) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	121	GLY	9.3
1	B	158	ALA	7.8
1	B	1	MET	7.7
1	B	120	ALA	7.6
1	B	157	VAL	6.6
1	B	122	ASN	5.9
1	A	158	ALA	5.7
1	B	119	ASP	5.5
1	A	-10	HIS	5.3
1	B	2	GLY	5.1
1	A	162	LEU	4.4
1	B	159	GLN	4.3
1	A	264	GLU	4.1
1	B	-5	VAL	3.8
1	A	159	GLN	3.7
1	B	171	GLN	3.7
1	B	153	LEU	3.6
1	B	160	GLY	3.6
1	B	123	PRO	3.6
1	B	148	PHE	3.5
1	B	117	PHE	3.4
1	B	-6	LEU	3.3
1	A	59	GLU	3.3
1	A	157	VAL	3.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	120	ALA	3.1
1	A	-6	LEU	2.9
1	B	59	GLU	2.9
1	B	155	GLU	2.9
1	A	-5	VAL	2.9
1	B	10	ALA	2.7
1	B	247	SER	2.7
1	B	154	ARG	2.5
1	A	247	SER	2.4
1	A	161	HIS	2.4
1	B	-8	SER	2.4
1	A	11	ALA	2.3
1	A	121	GLY	2.3
1	A	154	ARG	2.3
1	B	129	ARG	2.2
1	B	11	ALA	2.0
1	A	148	PHE	2.0
1	A	160	GLY	2.0
1	B	263	GLY	2.0

## 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](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	IPA	B	501	4/4	0.67	0.24	40,44,44,46	0
2	IPA	A	501	4/4	0.87	0.14	23,35,37,39	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.