



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 01:37 am GMT

PDB ID : 3UVA  
Title : Crystal structure of L-rhamnose isomerase mutant W38F from *Bacillus halodurans* in complex with Mn  
Authors : Doan, T.T.N.; Prabhu, P.; Jeya, M.; Kim, J.K.; Kang, L.W.; Lee, J.K.  
Deposited on : 2011-11-29  
Resolution : 2.69 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  
Xtriage (Phenix) : 1.9-1692  
EDS : trunk28620  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
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) : recalc28949

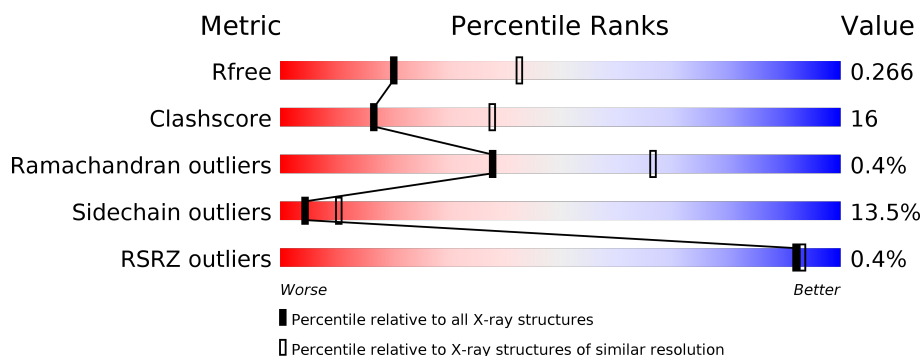
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.69 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))
$R_{free}$	100719	2259 (2.70-2.70)
Clashscore	112137	2590 (2.70-2.70)
Ramachandran outliers	110173	2550 (2.70-2.70)
Sidechain outliers	110143	2550 (2.70-2.70)
RSRZ outliers	101464	2275 (2.70-2.70)

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	424	 67% 22% 5% 5%
1	B	424	 58% 30% 7% 5%
1	C	424	 67% 23% 6% 5%
1	D	424	 61% 29% 5% 5%

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 13649 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 L-Rhamnose isomerase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	402	Total	C	N	O	S	0	0	0
			3289	2108	565	607	9			
1	B	404	Total	C	N	O	S	0	0	0
			3301	2115	567	609	10			
1	C	405	Total	C	N	O	S	0	0	0
			3307	2118	568	611	10			
1	D	403	Total	C	N	O	S	0	0	0
			3291	2110	564	608	9			

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-5	HIS	-	EXPRESSION TAG	UNP Q9KCL9
A	-4	HIS	-	EXPRESSION TAG	UNP Q9KCL9
A	-3	HIS	-	EXPRESSION TAG	UNP Q9KCL9
A	-2	HIS	-	EXPRESSION TAG	UNP Q9KCL9
A	-1	HIS	-	EXPRESSION TAG	UNP Q9KCL9
A	0	HIS	-	EXPRESSION TAG	UNP Q9KCL9
A	38	PHE	TRP	ENGINEERED MUTATION	UNP Q9KCL9
B	-5	HIS	-	EXPRESSION TAG	UNP Q9KCL9
B	-4	HIS	-	EXPRESSION TAG	UNP Q9KCL9
B	-3	HIS	-	EXPRESSION TAG	UNP Q9KCL9
B	-2	HIS	-	EXPRESSION TAG	UNP Q9KCL9
B	-1	HIS	-	EXPRESSION TAG	UNP Q9KCL9
B	0	HIS	-	EXPRESSION TAG	UNP Q9KCL9
B	38	PHE	TRP	ENGINEERED MUTATION	UNP Q9KCL9
C	-5	HIS	-	EXPRESSION TAG	UNP Q9KCL9
C	-4	HIS	-	EXPRESSION TAG	UNP Q9KCL9
C	-3	HIS	-	EXPRESSION TAG	UNP Q9KCL9
C	-2	HIS	-	EXPRESSION TAG	UNP Q9KCL9
C	-1	HIS	-	EXPRESSION TAG	UNP Q9KCL9
C	0	HIS	-	EXPRESSION TAG	UNP Q9KCL9
C	38	PHE	TRP	ENGINEERED MUTATION	UNP Q9KCL9

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	-5	HIS	-	EXPRESSION TAG	UNP Q9KCL9
D	-4	HIS	-	EXPRESSION TAG	UNP Q9KCL9
D	-3	HIS	-	EXPRESSION TAG	UNP Q9KCL9
D	-2	HIS	-	EXPRESSION TAG	UNP Q9KCL9
D	-1	HIS	-	EXPRESSION TAG	UNP Q9KCL9
D	0	HIS	-	EXPRESSION TAG	UNP Q9KCL9
D	38	PHE	TRP	ENGINEERED MUTATION	UNP Q9KCL9

- Molecule 2 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	2	Total Mn 2 2	0	0
2	A	2	Total Mn 2 2	0	0
2	D	2	Total Mn 2 2	0	0
2	C	2	Total Mn 2 2	0	0

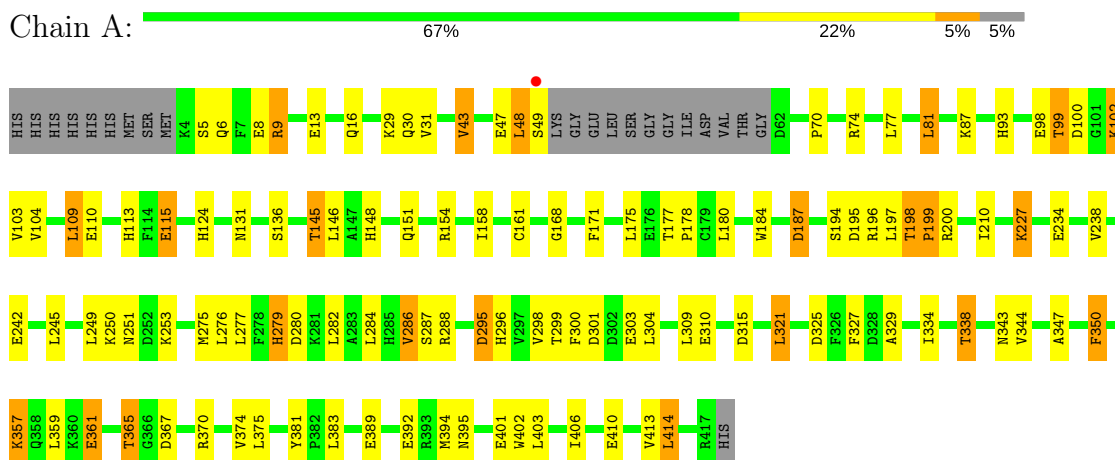
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	133	Total O 133 133	0	0
3	B	103	Total O 103 103	0	0
3	C	131	Total O 131 131	0	0
3	D	86	Total O 86 86	0	0

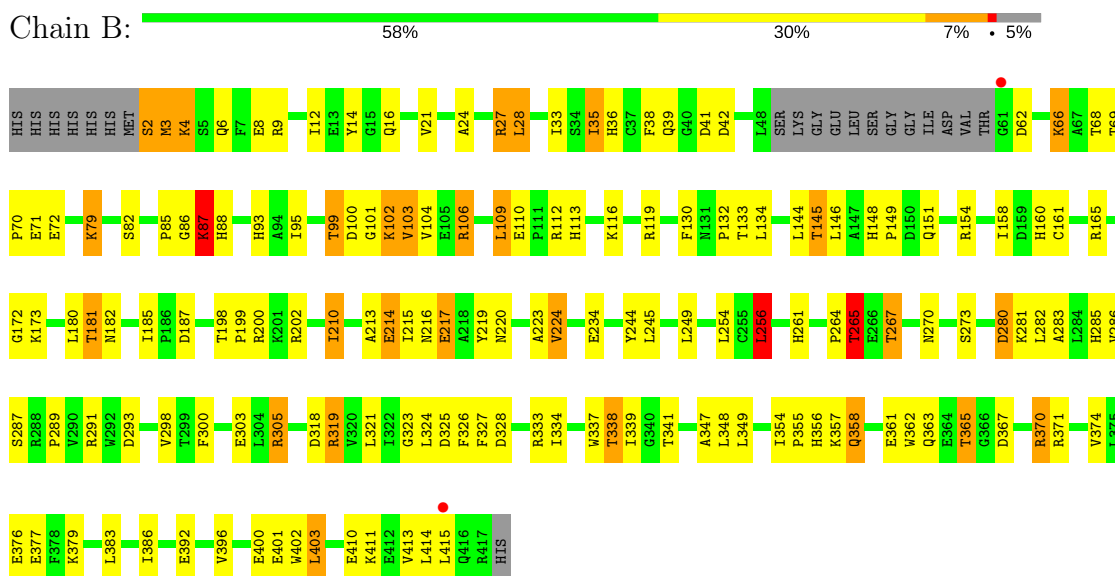
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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: L-Rhamnose isomerase



#### • Molecule 1: L-Rhamnose isomerase



#### • Molecule 1: L-Rhamnose isomerase





E410	D293	S194	K79	HIS
K411	S294		A80	HIS
E412	D295	R200	L81	HIS
V413	H296	K201	S82	HIS
L414			L83	HIS
L415	F300	E205	H88	HIS
Q416	L309	S206	R89	MET
ARG		L207		SER
HIS	R313		I95	KET
		T210		K4
	R319		K102	S5
	V320	E217		R9
	L321		L109	A10
		N220	E110	K11
	L324	A223	P111	I12
	D325	V224	R112	F13
	F326	E225	H113	Y14
	F327			
	D328	T231	K116	V21
			W17	
	I334			L25
	A335	S240	P132	F26
			T133	R27
	T338	V244		L28
	I339	L245	S136	K29
		S246	H137	Q30
	V344		E138	V31
		K250	K139	P32
	A347			I33
		L256	G143	S34
	I354	L257	L144	I35
	P355	D258	T145	H36
	H356		L146	C37
		H261	A147	F38
	K360	L262	H148	Q39
		H263	P149	
	Q363		D150	V43
	E364	E266	Q151	
				S49
	R371	K271	R154	K50
		L272	Q155	GLY
	V374	S273	F156	GLU
		A274	W157	LEU
	I386	N275	I158	SER
		L276	D159	GLY
	E392	L277	H160	ILE
		F278		ASP
	P397	H279	Y170	VAL
	I398	D280		THR
	K399	K281	K173	
	A400	L282		G61
	E401	A283	C179	D62
	W402	L284	L180	Y63
	L403	H285	T181	P64
	K404	V286	N182	
	E405	S287		L73
		R288	D187	R74
	I408		G188	M75
	Y409	Y292	Y198	

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	83.37Å 164.78Å 92.82Å 90.00° 117.01° 90.00°	Depositor
Resolution (Å)	45.87 – 2.69 45.87 – 2.69	Depositor EDS
% Data completeness (in resolution range)	88.0 (45.87-2.69) 88.0 (45.87-2.69)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.88 (at 2.69Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, $R_{free}$	0.185 , 0.266 0.189 , 0.266	Depositor DCC
$R_{free}$ test set	2778 reflections (5.39%)	DCC
Wilson B-factor (Å <sup>2</sup> )	36.6	Xtriage
Anisotropy	0.072	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 32.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.021 for h,-k,-h-l	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	13649	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	32.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.34% 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: MN

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	0.82	0/3369	0.85	5/4560 (0.1%)
1	B	0.76	1/3381 (0.0%)	0.84	3/4575 (0.1%)
1	C	0.80	5/3387 (0.1%)	0.84	0/4583
1	D	0.82	0/3371	0.82	1/4562 (0.0%)
All	All	0.80	6/13508 (0.0%)	0.84	9/18280 (0.0%)

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

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	266	GLU	CD-OE2	-5.56	1.19	1.25
1	C	266	GLU	CD-OE1	-5.55	1.19	1.25
1	C	262	TYR	CE1-CZ	-5.44	1.31	1.38
1	C	205	GLU	CB-CG	5.31	1.62	1.52
1	C	392	GLU	CG-CD	5.22	1.59	1.51

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	87	LYS	N-CA-C	-5.69	95.65	111.00
1	D	415	LEU	CA-CB-CG	5.63	128.24	115.30
1	B	415	LEU	CA-CB-CG	5.34	127.59	115.30
1	A	301	ASP	CB-CG-OD1	5.33	123.10	118.30

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	367	ASP	CB-CG-OD1	5.26	123.04	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	47	GLU	Peptide

## 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	3289	0	3245	92	1
1	B	3301	0	3257	153	0
1	C	3307	0	3262	81	0
1	D	3291	0	3248	105	1
2	A	2	0	0	0	0
2	B	2	0	0	0	0
2	C	2	0	0	0	0
2	D	2	0	0	0	0
3	A	133	0	0	5	0
3	B	103	0	0	4	0
3	C	131	0	0	6	0
3	D	86	0	0	4	0
All	All	13649	0	13012	415	1

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 415 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:102:LYS:HA	1:B:103:VAL:CG2	1.14	1.53
1:B:102:LYS:CA	1:B:103:VAL:HG23	1.49	1.40
1:B:102:LYS:CA	1:B:103:VAL:CG2	2.06	1.32

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:27:ARG:O	1:D:30:GLN:HG2	1.30	1.27
1:A:48:LEU:HD13	1:A:48:LEU:C	1.54	1.22

All (1) 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)
1:A:381:TYR:OH	1:D:400:GLU:OE2[1_655]	2.15	0.05

## 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	398/424 (94%)	365 (92%)	32 (8%)	1 (0%)	44 73
1	B	400/424 (94%)	371 (93%)	27 (7%)	2 (0%)	32 60
1	C	401/424 (95%)	374 (93%)	26 (6%)	1 (0%)	51 79
1	D	399/424 (94%)	372 (93%)	24 (6%)	3 (1%)	22 49
All	All	1598/1696 (94%)	1482 (93%)	109 (7%)	7 (0%)	38 66

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	265	THR
1	D	136	SER
1	C	259	THR
1	D	32	PRO
1	D	293	ASP

### 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	352/370 (95%)	311 (88%)	41 (12%)	6	15
1	B	353/370 (95%)	297 (84%)	56 (16%)	3	7
1	C	354/370 (96%)	308 (87%)	46 (13%)	5	11
1	D	352/370 (95%)	304 (86%)	48 (14%)	4	10
All	All	1411/1480 (95%)	1220 (86%)	191 (14%)	4	10

5 of 191 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	338	THR
1	C	81	LEU
1	D	284	LEU
1	B	357	LYS
1	B	411	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 39 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	131	ASN
1	C	241	HIS
1	D	356	HIS
1	C	137	HIS
1	C	160	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

There are no carbohydrates in this entry.

## 5.6 Ligand geometry

Of 8 ligands modelled in this entry, 8 are monoatomic - leaving 0 for Mogul analysis.

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

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 [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	402/424 (94%)	-0.59	1 (0%) 94 96	17, 29, 48, 68	0
1	B	404/424 (95%)	-0.53	2 (0%) 90 92	19, 32, 53, 77	0
1	C	405/424 (95%)	-0.58	3 (0%) 87 88	18, 31, 46, 84	0
1	D	403/424 (95%)	-0.51	1 (0%) 94 96	19, 31, 51, 72	0
All	All	1614/1696 (95%)	-0.55	7 (0%) 92 93	17, 31, 50, 84	0

The worst 5 of 7 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	3	MET	5.0
1	C	2	SER	4.3
1	B	415	LEU	2.5
1	A	49	SER	2.3
1	C	49	SER	2.2

### 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	MN	A	420	1/1	0.94	0.20	2.00	81,81,81,81	0
2	MN	D	420	1/1	0.93	0.17	1.50	75,75,75,75	0
2	MN	B	420	1/1	0.93	0.14	0.14	74,74,74,74	0
2	MN	C	420	1/1	0.97	0.13	-0.47	67,67,67,67	0
2	MN	D	419	1/1	0.98	0.06	-3.12	71,71,71,71	0
2	MN	A	419	1/1	0.96	0.07	-3.28	69,69,69,69	0
2	MN	C	419	1/1	0.95	0.13	-	81,81,81,81	0
2	MN	B	419	1/1	0.94	0.10	-	82,82,82,82	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.