



# Full wwPDB X-ray Structure Validation Report

Mar 31, 2014 – 07:56 PM EDT

PDB ID : 4L6H  
Title : Crystal structure of the Candida albicans Methionine Synthase in complex with Methotrexate and Homocysteine  
Authors : Ubhi, D.; Robertus, J.D.  
Deposited on : 2013-06-12  
Resolution : 1.75 Å(reported)

This is a full wwPDB 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 <http://wwpdb.org/ValidationPDFNotes.html>

---

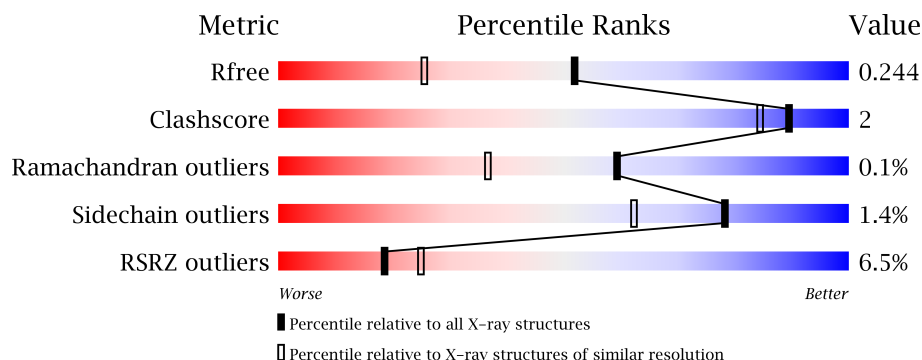
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1439  
EDS : stable22978  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22978

# 1 Overall quality at a glance

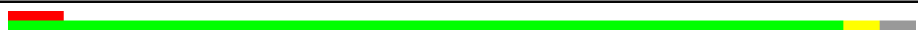
The reported resolution of this entry is 1.75 Å.

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))
$R_{free}$	66092	1134 (1.76-1.76)
Clashscore	79885	1304 (1.76-1.76)
Ramachandran outliers	78287	1288 (1.76-1.76)
Sidechain outliers	78261	1288 (1.76-1.76)
RSRZ outliers	66119	1135 (1.76-1.76)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	789	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
4	MTX	A	803[A]	-	X
4	MTX	A	803[B]	-	X

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 6486 atoms, of which 0 are hydrogen and 0 are deuterium.

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 5-methyltetrahydropteroyltriglutamate--homocysteinemethyltransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	761	5765	3696	958	1100	11	0	4	0

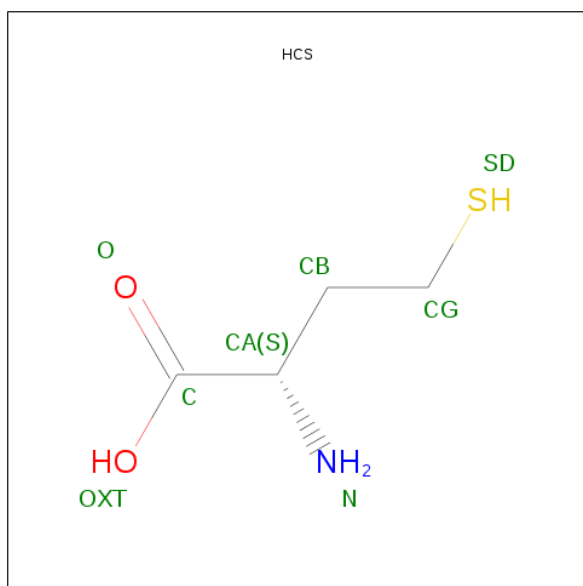
There are 25 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-21	MET	-	EXPRESSION TAG	UNP P82610
A	-20	HIS	-	EXPRESSION TAG	UNP P82610
A	-19	HIS	-	EXPRESSION TAG	UNP P82610
A	-18	HIS	-	EXPRESSION TAG	UNP P82610
A	-17	HIS	-	EXPRESSION TAG	UNP P82610
A	-16	HIS	-	EXPRESSION TAG	UNP P82610
A	-15	HIS	-	EXPRESSION TAG	UNP P82610
A	-14	SER	-	EXPRESSION TAG	UNP P82610
A	-13	SER	-	EXPRESSION TAG	UNP P82610
A	-12	GLY	-	EXPRESSION TAG	UNP P82610
A	-11	VAL	-	EXPRESSION TAG	UNP P82610
A	-10	ASP	-	EXPRESSION TAG	UNP P82610
A	-9	LEU	-	EXPRESSION TAG	UNP P82610
A	-8	GLY	-	EXPRESSION TAG	UNP P82610
A	-7	THR	-	EXPRESSION TAG	UNP P82610
A	-6	GLU	-	EXPRESSION TAG	UNP P82610
A	-5	ASN	-	EXPRESSION TAG	UNP P82610
A	-4	LEU	-	EXPRESSION TAG	UNP P82610
A	-3	TYR	-	EXPRESSION TAG	UNP P82610
A	-2	PHE	-	EXPRESSION TAG	UNP P82610
A	-1	GLN	-	EXPRESSION TAG	UNP P82610
A	0	SER	-	EXPRESSION TAG	UNP P82610
A	103	ALA	LYS	engineered mutation	UNP P82610
A	104	ALA	LYS	engineered mutation	UNP P82610
A	107	ALA	GLU	engineered mutation	UNP P82610

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

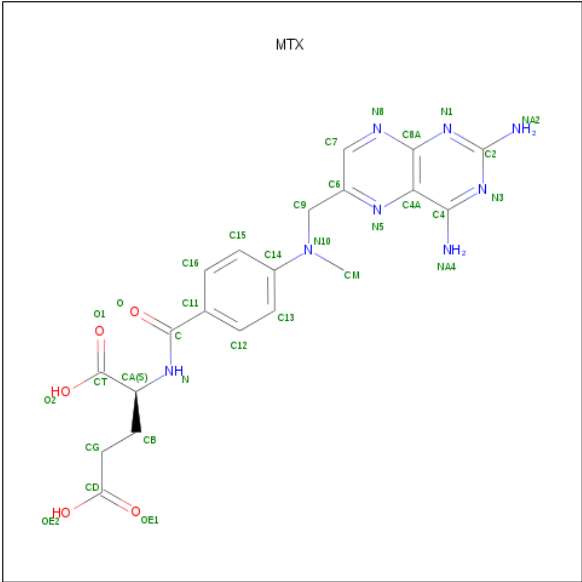
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Zn	0	0
			1	1		

- Molecule 3 is 2-AMINO-4-MERCAPTO-BUTYRICACID (three-letter code: HCS) (formula: C<sub>4</sub>H<sub>9</sub>NO<sub>2</sub>S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	S	0	0
			8	4	1	2	1		

- Molecule 4 is METHOTREXATE (three-letter code: MTX) (formula: C<sub>20</sub>H<sub>22</sub>N<sub>8</sub>O<sub>5</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	N	O	0	1
			66	40	16	10		

- Molecule 5 is water.

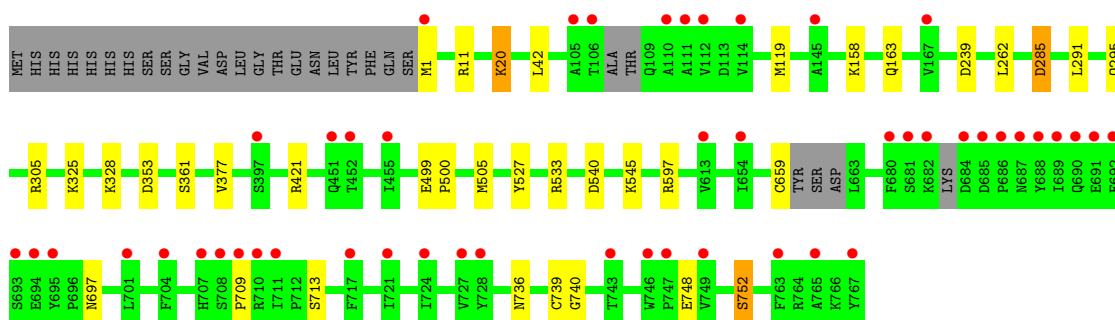
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	646	Total	O	0	0
			646	646		

### 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 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 ( $RSRZ > 2$ ). 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: 5-methyltetrahydropteroyltriglutamate--homocysteinemethyltransferase

Chain A: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	77.07Å 98.98Å 101.94Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	33.89 – 1.75 33.87 – 1.75	Depositor EDS
% Data completeness (in resolution range)	99.7 (33.89-1.75) 99.7 (33.87-1.75)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.86 (at 1.75Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, $R_{free}$	0.189 , 0.235 0.199 , 0.244	Depositor DCC
$R_{free}$ test set	3963 reflections (5.29%)	DCC
Wilson B-factor (Å <sup>2</sup> )	22.0	Xtriage
Anisotropy	0.179	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 35.9	EDS
Estimated twinning fraction	0.009 for -h,l,k	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	3 of 78865 reflections (0.004%)	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	6486	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 5.16% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: HCS, MTX, ZN

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	0.92	1/5885 (0.0%)	0.93	12/8018 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	361	SER	CA-CB	5.13	1.60	1.52

All (12) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	597	ARG	NE-CZ-NH2	7.96	124.28	120.30
1	A	305	ARG	NE-CZ-NH1	-6.61	116.99	120.30
1	A	285	ASP	CB-CG-OD1	6.58	124.22	118.30
1	A	533	ARG	NE-CZ-NH1	5.86	123.23	120.30
1	A	505	MET	CB-CG-SD	-5.78	95.07	112.40
1	A	505	MET	CG-SD-CE	-5.54	91.33	100.20
1	A	291	LEU	CA-CB-CG	5.44	127.81	115.30
1	A	540	ASP	CB-CG-OD1	5.37	123.13	118.30
1	A	239	ASP	CB-CG-OD1	5.31	123.08	118.30
1	A	421	ARG	NE-CZ-NH1	5.20	122.90	120.30
1	A	42	LEU	CB-CG-CD2	-5.09	102.34	111.00
1	A	285	ASP	CB-CG-OD2	-5.07	113.74	118.30

There are no chirality outliers.

There are no planarity outliers.



## 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	5765	0	0	13	0
2	A	1	0	0	0	0
3	A	8	0	7	0	0
4	A	66	0	40	0	0
5	A	646	0	0	4	0
All	All	6486	0	47	13	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 2.

All (13) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:1[B]:MET:CA	1:A:1[B]:MET:CE	2.65	0.73
1:A:748:GLU:O	1:A:752:SER:OG	2.24	0.55
1:A:119:MET:CB	1:A:527[A]:TYR:CZ	2.93	0.52
1:A:659:CYS:CB	5:A:1145:HOH:O	2.59	0.50
1:A:353:ASP:N	5:A:1213:HOH:O	2.47	0.47
1:A:1[B]:MET:CE	1:A:1[B]:MET:N	2.77	0.47
1:A:295:GLN:N	5:A:1185:HOH:O	2.46	0.47
1:A:739:CYS:SG	1:A:740:GLY:N	2.88	0.47
1:A:285:ASP:OD1	1:A:325:LYS:NZ	2.48	0.46
1:A:158:LYS:NZ	1:A:163:GLN:OE1	2.49	0.46
1:A:328:LYS:NZ	1:A:377:VAL:O	2.53	0.42
1:A:20[A]:LYS:CD	5:A:1328:HOH:O	2.68	0.41
1:A:499:GLU:N	1:A:500:PRO:CD	2.84	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone ⓘ

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	756/789 (96%)	739 (98%)	16 (2%)	1 (0%)	59	37

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	709	PRO

### 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	579/678 (85%)	570 (98%)	9 (2%)	75	55

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	11	ARG
1	A	20[A]	LYS
1	A	20[B]	LYS
1	A	262	LEU
1	A	545	LYS
1	A	697	ASN
1	A	713	SER
1	A	736	ASN
1	A	752	SER

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

### 5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 4 ligands modelled in this entry, 1 is monoatomic - leaving 3 for Mogul analysis.

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	$\# Z  > 2$
3	HCS	A	802	2	7,7,7	1.44	1 (14%)	8,8,8	1.16	0
4	MTX	A	803[A]	-	35,35,35	1.21	4 (11%)	49,49,49	1.86	15 (30%)
4	MTX	A	803[B]	-	35,35,35	1.28	4 (11%)	49,49,49	2.14	17 (34%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	HCS	A	802	2	-	0/7/7/7	0/0/0/0
4	MTX	A	803[A]	-	-	0/25/25/25	0/1/3/3
4	MTX	A	803[B]	-	-	0/25/25/25	0/1/3/3

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	803[B]	MTX	C4A-C8A	3.55	1.48	1.40
4	A	803[A]	MTX	OE1-CD	3.38	1.34	1.22
3	A	802	HCS	O-C	3.34	1.33	1.22
4	A	803[B]	MTX	OE1-CD	3.32	1.34	1.22

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	803[A]	MTX	C4A-C8A	3.31	1.47	1.40
4	A	803[B]	MTX	OE2-CD	-2.17	1.22	1.30
4	A	803[A]	MTX	OE2-CD	-2.11	1.22	1.30
4	A	803[B]	MTX	C6-N5	2.09	1.37	1.32
4	A	803[A]	MTX	C14-N10	2.02	1.43	1.38

All (32) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	803[B]	MTX	C2-N1-C8A	5.77	122.95	115.30
4	A	803[A]	MTX	C2-N1-C8A	5.33	122.36	115.30
4	A	803[B]	MTX	C6-C9-N10	-5.05	104.96	113.79
4	A	803[B]	MTX	N8-C8A-N1	4.89	122.78	116.19
4	A	803[A]	MTX	N8-C8A-N1	4.28	121.96	116.19
4	A	803[B]	MTX	C4A-C8A-N1	-3.79	116.40	122.26
4	A	803[B]	MTX	C8A-C4A-N5	-3.71	118.11	122.66
4	A	803[A]	MTX	C6-C9-N10	-3.67	107.38	113.79
4	A	803[A]	MTX	N1-C2-N3	-3.52	121.98	127.54
4	A	803[B]	MTX	C11-C-N	3.14	122.13	116.89
4	A	803[A]	MTX	C4-C4A-N5	3.07	123.47	120.31
4	A	803[A]	MTX	C11-C-N	3.04	121.95	116.89
4	A	803[B]	MTX	CM-N10-C9	3.02	122.17	114.12
4	A	803[B]	MTX	N1-C2-N3	-2.90	122.96	127.54
4	A	803[B]	MTX	CG-CB-CA	-2.71	107.68	112.99
4	A	803[B]	MTX	CA-N-C	-2.62	115.96	121.80
4	A	803[B]	MTX	C4-C4A-C8A	2.57	119.02	116.87
4	A	803[B]	MTX	C6-C7-N8	-2.51	120.52	123.16
4	A	803[B]	MTX	C4-C4A-N5	2.50	122.87	120.31
4	A	803[B]	MTX	CB-CA-CT	2.48	116.05	110.71
4	A	803[A]	MTX	O-C-C11	-2.46	116.70	121.01
4	A	803[B]	MTX	C7-C6-N5	2.44	122.44	120.75
4	A	803[A]	MTX	C7-N8-C8A	2.32	119.45	116.85
4	A	803[A]	MTX	C4A-C8A-N1	-2.24	118.81	122.26
4	A	803[B]	MTX	CT-CA-N	-2.20	104.87	110.53
4	A	803[A]	MTX	OE2-CD-OE1	-2.19	117.72	123.30
4	A	803[A]	MTX	OE2-CD-CG	2.16	121.86	114.22
4	A	803[A]	MTX	C9-N10-C14	2.12	123.77	119.36
4	A	803[B]	MTX	CM-N10-C14	-2.12	115.64	119.58
4	A	803[A]	MTX	C9-C6-C7	-2.08	117.77	121.49
4	A	803[A]	MTX	C8A-C4A-N5	-2.07	120.12	122.66
4	A	803[A]	MTX	C9-C6-N5	2.04	120.40	116.97

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 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 ⓘ

### 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	761/789 (96%)	0.23	49 (6%) 19 25	11, 24, 63, 118	0

All (49) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	689	ILE	6.4
1	A	707	HIS	5.9
1	A	709	PRO	5.8
1	A	111	ALA	4.9
1	A	692	PHE	4.9
1	A	767	TYR	4.9
1	A	727	VAL	4.9
1	A	708	SER	4.6
1	A	717	PHE	4.6
1	A	721	ILE	4.5
1	A	728	TYR	4.3
1	A	110	ALA	4.3
1	A	724	ILE	4.2
1	A	688	TYR	4.0
1	A	695	TYR	3.8
1	A	747	PRO	3.8
1	A	112	VAL	3.6
1	A	685	ASP	3.6
1	A	711	ILE	3.5
1	A	145	ALA	3.4
1	A	691	GLU	3.3
1	A	710	ARG	3.3
1	A	765	ALA	3.2
1	A	687	ASN	3.1
1	A	114	VAL	3.0
1	A	746	TRP	2.9
1	A	686	PRO	2.9

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	763	PHE	2.9
1	A	684	ASP	2.9
1	A	106	THR	2.8
1	A	167	VAL	2.8
1	A	693	SER	2.7
1	A	682	LYS	2.6
1	A	749	VAL	2.5
1	A	694	GLU	2.5
1	A	680	PHE	2.5
1	A	743	THR	2.5
1	A	690	GLN	2.5
1	A	105	ALA	2.4
1	A	451	GLN	2.3
1	A	613	VAL	2.3
1	A	681	SER	2.2
1	A	452	THR	2.2
1	A	455	ILE	2.1
1	A	654	ILE	2.1
1	A	1[A]	MET	2.1
1	A	397	SER	2.0
1	A	704	PHE	2.0
1	A	701	LEU	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 6.4 Ligands ⓘ

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	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
4	MTX	A	803[B]	33/33	0.26	8.15	23,34,40,44	33
4	MTX	A	803[A]	33/33	0.26	7.25	27,53,62,67	33
3	HCS	A	802	8/8	0.10	-0.31	24,29,33,34	0
2	ZN	A	801	1/1	0.06	-2.19	35,35,35,35	0

## 6.5 Other polymers

There are no such residues in this entry.