



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 11, 2024 – 08:26 PM EST

PDB ID : 3DHG
Title : Crystal Structure of Toluene 4-Monooxygenase Hydroxylase
Authors : Bailey, L.J.; Mccoy, J.G.; Phillips Jr., G.N.; Fox, B.G.
Deposited on : 2008-06-17
Resolution : 1.85 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

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
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.36
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

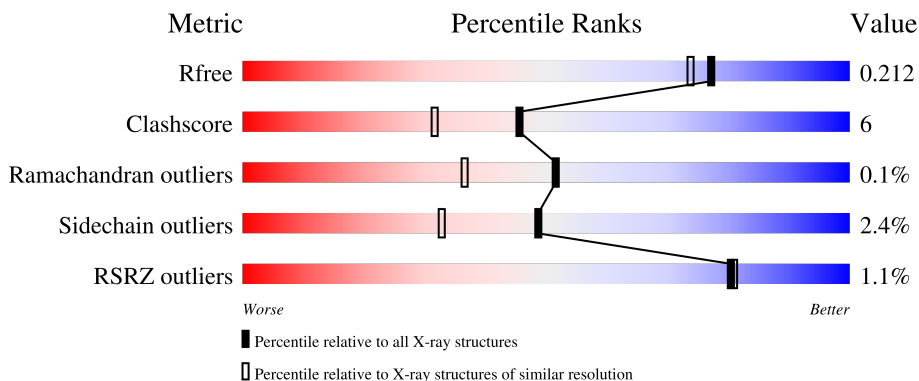
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.85 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	2469 (1.86-1.86)
Clashscore	141614	2625 (1.86-1.86)
Ramachandran outliers	138981	2592 (1.86-1.86)
Sidechain outliers	138945	2592 (1.86-1.86)
RSRZ outliers	127900	2436 (1.86-1.86)

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 ≥ 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	500	 90% 8% .
1	D	500	 86% 12% ..
2	B	327	 80% 12% . 7%
2	E	327	 82% 9% . 7%
3	C	84	 82% 13% ..

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	84	 <p>6% 76% 19% . .</p>

2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 16033 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 toluene 4-monooxygenase hydroxylase alpha subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	491	4050	2599	683	745	23	0	3	0
1	D	492	4101	2633	690	754	24	0	10	0

- Molecule 2 is a protein called toluene 4-monooxygenase hydroxylase beta subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	304	2547	1617	439	476	15	0	4	0
2	E	304	2551	1617	439	479	16	0	6	0

- Molecule 3 is a protein called toluene 4-monooxygenase hydroxylase gamma subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	82	664	420	117	123	4	0	2	0
3	F	81	645	407	116	118	4	0	0	0

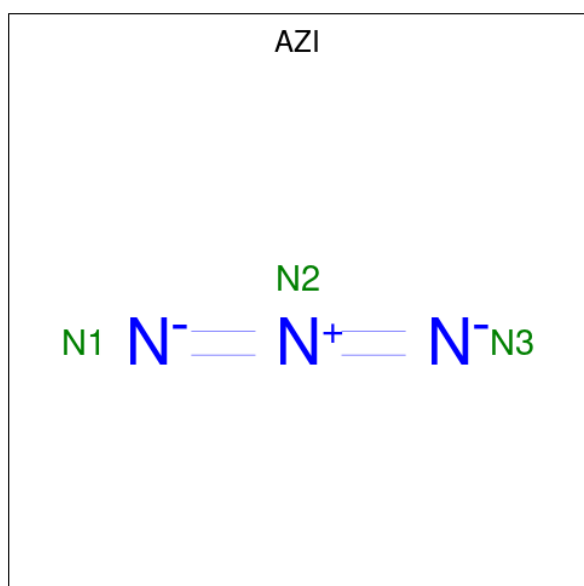
- Molecule 4 is FE (III) ION (three-letter code: FE) (formula: Fe).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	2	Total Fe 2 2	0	0
4	D	2	Total Fe 2 2	0	0

- Molecule 5 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	5	Total Ca 5 5	0	0
5	C	1	Total Ca 1 1	0	0
5	E	2	Total Ca 2 2	0	0
5	F	1	Total Ca 1 1	0	0

- Molecule 6 is AZIDE ION (three-letter code: AZI) (formula: N₃).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	D	1	Total N 3 3	0	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	352	Total O 352 352	0	0
7	B	302	Total O 302 302	0	0
7	C	42	Total O 42 42	0	0
7	D	408	Total O 408 408	0	0
7	E	306	Total O 306 306	0	0

Continued on next page...

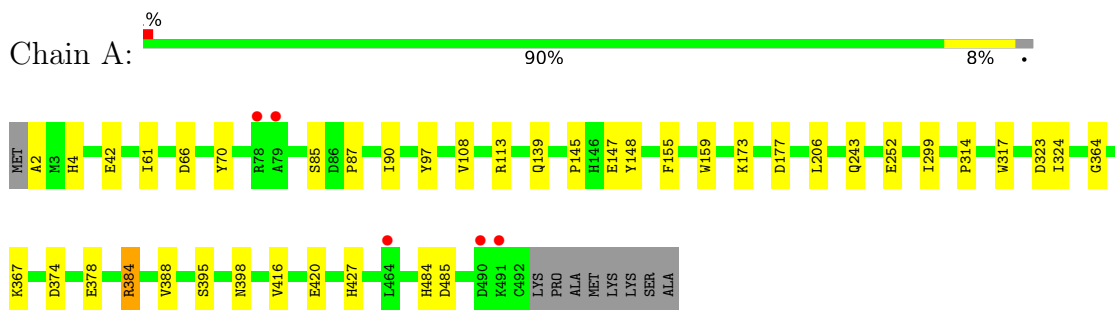
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	F	49	Total	O	0	0
			49	49		

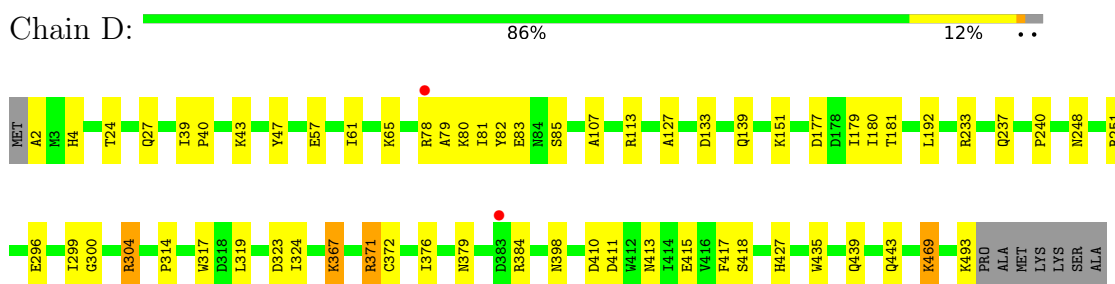
3 Residue-property plots [i](#)

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.

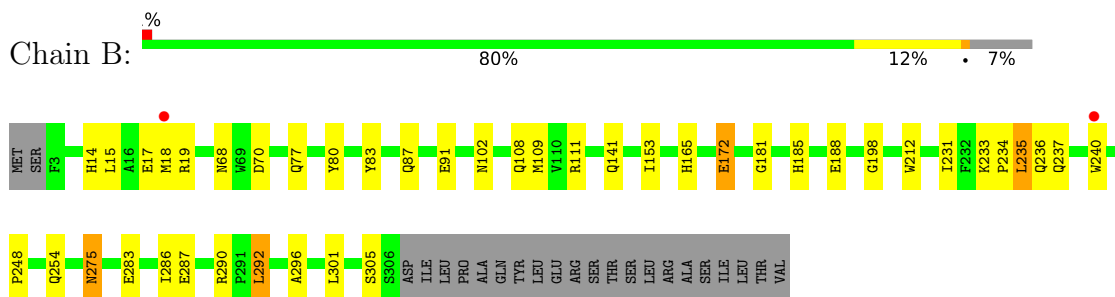
- Molecule 1: toluene 4-monooxygenase hydroxylase alpha subunit



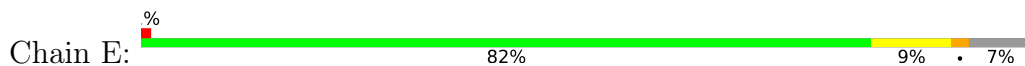
- Molecule 1: toluene 4-monooxygenase hydroxylase alpha subunit

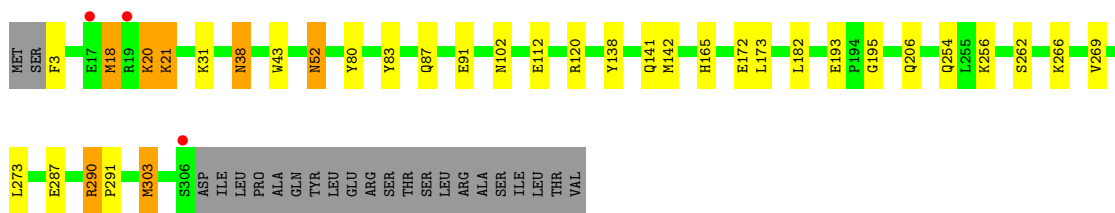


- Molecule 2: toluene 4-monooxygenase hydroxylase beta subunit

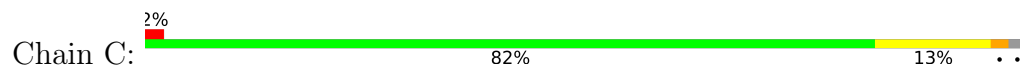


- Molecule 2: toluene 4-monooxygenase hydroxylase beta subunit

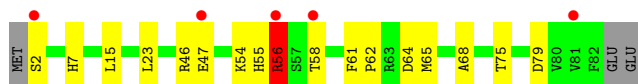
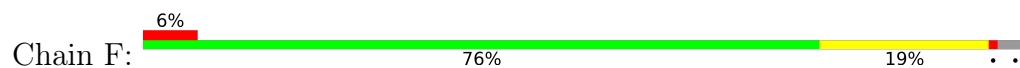




- Molecule 3: toluene 4-monooxygenase hydroxylase gamma subunit



- Molecule 3: toluene 4-monooxygenase hydroxylase gamma subunit



4 Data and refinement statistics i

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	56.82Å 181.48Å 89.92Å 90.00° 107.62° 90.00°	Depositor
Resolution (Å)	49.45 – 1.85 49.42 – 1.85	Depositor EDS
% Data completeness (in resolution range)	98.5 (49.45-1.85) 98.5 (49.42-1.85)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.02 (at 1.86Å)	Xtrriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.164 , 0.214 0.163 , 0.212	Depositor DCC
R_{free} test set	7265 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å ²)	22.1	Xtrriage
Anisotropy	0.475	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 46.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	0.036 for h,-k,-h-l	Xtrriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	16033	wwPDB-VP
Average B, all atoms (Å ²)	24.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.56% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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 [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: AZI, FE, CA

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	0.75	0/4188	0.70	0/5686
1	D	0.79	0/4254	0.74	3/5774 (0.1%)
2	B	0.79	1/2636 (0.0%)	0.76	1/3584 (0.0%)
2	E	0.81	2/2644 (0.1%)	0.74	1/3593 (0.0%)
3	C	0.71	0/682	0.78	1/924 (0.1%)
3	F	0.70	0/657	0.77	0/890
All	All	0.78	3/15061 (0.0%)	0.74	6/20451 (0.0%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	303[A]	MET	C-N	-5.47	1.21	1.34
2	E	303[B]	MET	C-N	-5.47	1.21	1.34
2	B	172	GLU	CD-OE2	-5.13	1.20	1.25

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	233	ARG	NE-CZ-NH1	-5.64	117.48	120.30
1	D	233	ARG	NE-CZ-NH2	5.54	123.07	120.30
2	B	235	LEU	CA-CB-CG	5.29	127.48	115.30
3	C	42	ARG	CG-CD-NE	-5.22	100.84	111.80
1	D	133	ASP	CB-CG-OD2	-5.11	113.70	118.30

There are no chirality outliers.

There are no planarity outliers.

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	4050	0	3820	36	0
1	D	4101	0	3882	58	0
2	B	2547	0	2428	45	0
2	E	2551	0	2439	39	0
3	C	664	0	664	13	0
3	F	645	0	643	18	0
4	A	2	0	0	0	0
4	D	2	0	0	0	0
5	B	5	0	0	0	0
5	C	1	0	0	0	0
5	E	2	0	0	0	0
5	F	1	0	0	0	0
6	D	3	0	0	0	0
7	A	352	0	0	5	0
7	B	302	0	0	14	0
7	C	42	0	0	1	0
7	D	408	0	0	7	0
7	E	306	0	0	6	0
7	F	49	0	0	2	0
All	All	16033	0	13876	182	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 6.

The worst 5 of 182 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)
2:E:193:GLU:CG	2:E:303[A]:MET:HE2	1.79	1.11
3:C:23:LEU:HA	3:C:67[A]:ILE:HD11	1.36	1.07
1:D:61[A]:ILE:HD11	7:D:992:HOH:O	1.59	1.02
2:B:18:MET:HG2	7:B:1429:HOH:O	1.61	1.00
2:E:193:GLU:HG3	2:E:303[A]:MET:HE2	1.42	1.00

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	492/500 (98%)	477 (97%)	15 (3%)	0	100	100
1	D	500/500 (100%)	486 (97%)	14 (3%)	0	100	100
2	B	306/327 (94%)	301 (98%)	5 (2%)	0	100	100
2	E	308/327 (94%)	302 (98%)	6 (2%)	0	100	100
3	C	82/84 (98%)	77 (94%)	5 (6%)	0	100	100
3	F	79/84 (94%)	75 (95%)	3 (4%)	1 (1%)	12	3
All	All	1767/1822 (97%)	1718 (97%)	48 (3%)	1 (0%)	51	36

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	F	56	ARG

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	419/423 (99%)	414 (99%)	5 (1%)	71	62
1	D	427/423 (101%)	416 (97%)	11 (3%)	46	30
2	B	279/296 (94%)	274 (98%)	5 (2%)	59	45
2	E	281/296 (95%)	274 (98%)	7 (2%)	47	31
3	C	75/75 (100%)	72 (96%)	3 (4%)	31	14
3	F	72/75 (96%)	66 (92%)	6 (8%)	11	2

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	1553/1588 (98%)	1516 (98%)	37 (2%)	49 33

5 of 37 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	E	52	ASN
3	F	58	THR
2	E	206	GLN
3	F	46	ARG
3	C	58	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 48 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	439	GLN
2	E	84	ASN
1	D	443	GLN
2	E	38	ASN
2	E	108	GLN

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](#)

Of 14 ligands modelled in this entry, 13 are monoatomic - leaving 1 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
6	AZI	D	510	4	0,2,2	-	-	0,1,1	-	-

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 [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, 95th 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 ²)	Q<0.9
1	A	491/500 (98%)	-0.33	5 (1%) 82 82	14, 23, 39, 52	0
1	D	492/500 (98%)	-0.41	2 (0%) 92 92	14, 21, 34, 50	0
2	B	304/327 (92%)	-0.39	2 (0%) 87 88	15, 21, 33, 46	0
2	E	304/327 (92%)	-0.45	3 (0%) 82 82	13, 20, 32, 43	0
3	C	82/84 (97%)	-0.15	2 (2%) 59 57	22, 31, 41, 48	0
3	F	81/84 (96%)	-0.10	5 (6%) 20 20	20, 30, 44, 49	0
All	All	1754/1822 (96%)	-0.36	19 (1%) 80 81	13, 22, 38, 52	0

The worst 5 of 19 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	18	MET	4.9
2	E	17	GLU	3.3
1	A	464	LEU	3.2
2	B	240[A]	TRP	3.0
3	F	56	ARG	2.9

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](#)

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, 95th 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(Å ²)	Q<0.9
6	AZI	D	510	3/3	0.91	0.23	33,33,41,42	0
5	CA	B	511	1/1	0.97	0.11	49,49,49,49	0
5	CA	B	504	1/1	0.97	0.12	39,39,39,39	0
4	FE	D	507	1/1	0.99	0.10	23,23,23,23	0
5	CA	B	505	1/1	0.99	0.04	28,28,28,28	0
5	CA	B	502	1/1	0.99	0.10	31,31,31,31	0
5	CA	C	506	1/1	0.99	0.05	38,38,38,38	0
5	CA	E	509	1/1	0.99	0.08	33,33,33,33	0
5	CA	E	512	1/1	0.99	0.06	29,29,29,29	0
5	CA	F	513	1/1	0.99	0.07	38,38,38,38	0
5	CA	B	503	1/1	0.99	0.03	24,24,24,24	0
4	FE	A	501	1/1	1.00	0.11	20,20,20,20	0
4	FE	D	508	1/1	1.00	0.10	18,18,18,18	0
4	FE	A	502	1/1	1.00	0.10	23,23,23,23	0

6.5 Other polymers [i](#)

There are no such residues in this entry.