



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 4, 2017 – 04:27 PM EDT

PDB ID : 2B67  
Title : Crystal structure of the Nitroreductase Family Protein from *Streptococcus pneumoniae* TIGR4  
Authors : Kim, Y.; Volkart, L.; Abdullah, J.; Collart, F.; Joachimiak, A.; Midwest Center for Structural Genomics (MCSG)  
Deposited on : unknown  
Resolution : 2.05 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
Mogul	:	1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20030345
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)	:	rb-20030345

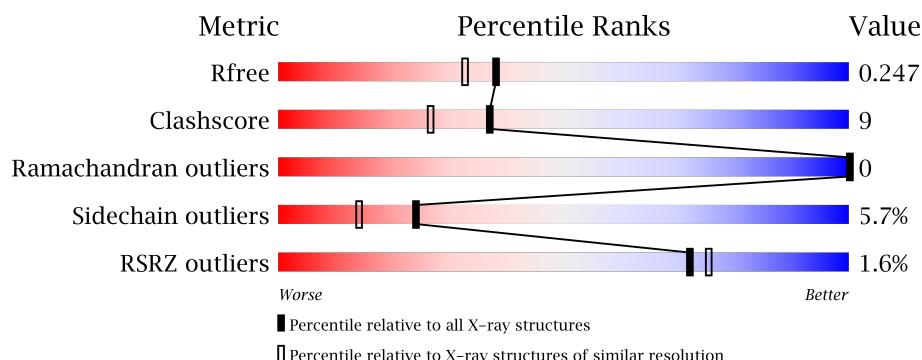
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.05 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	1316 (2.04-2.04)
Clashscore	112137	1394 (2.04-2.04)
Ramachandran outliers	110173	1383 (2.04-2.04)
Sidechain outliers	110143	1383 (2.04-2.04)
RSRZ outliers	101464	1319 (2.04-2.04)

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	204	
1	B	204	
1	C	204	
1	D	204	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	FMN	A	801	-	-	-	X
2	FMN	B	802	-	-	-	X
2	FMN	C	803	-	-	-	X
2	FMN	D	804	-	-	-	X

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 7396 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 COG0778: Nitroreductase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	201	Total	C	N	O	Se	0	4	0
			1638	1040	285	310	3			
1	B	200	Total	C	N	O	Se	0	1	0
			1602	1020	275	305	2			
1	C	200	Total	C	N	O	Se	0	5	0
			1636	1041	280	312	3			
1	D	201	Total	C	N	O	Se	0	4	0
			1639	1043	285	307	4			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	SER	-	CLONING ARTIFACT	UNP Q97S03
A	-1	ASN	-	CLONING ARTIFACT	UNP Q97S03
A	0	ALA	-	CLONING ARTIFACT	UNP Q97S03
A	1	MSE	MET	MODIFIED RESIDUE	UNP Q97S03
A	109	MSE	MET	MODIFIED RESIDUE	UNP Q97S03
A	137	MSE	MET	MODIFIED RESIDUE	UNP Q97S03
B	-2	SER	-	CLONING ARTIFACT	UNP Q97S03
B	-1	ASN	-	CLONING ARTIFACT	UNP Q97S03
B	0	ALA	-	CLONING ARTIFACT	UNP Q97S03
B	1	MSE	MET	MODIFIED RESIDUE	UNP Q97S03
B	109	MSE	MET	MODIFIED RESIDUE	UNP Q97S03
B	137	MSE	MET	MODIFIED RESIDUE	UNP Q97S03
C	-2	SER	-	CLONING ARTIFACT	UNP Q97S03
C	-1	ASN	-	CLONING ARTIFACT	UNP Q97S03
C	0	ALA	-	CLONING ARTIFACT	UNP Q97S03
C	1	MSE	MET	MODIFIED RESIDUE	UNP Q97S03
C	109	MSE	MET	MODIFIED RESIDUE	UNP Q97S03
C	137	MSE	MET	MODIFIED RESIDUE	UNP Q97S03
D	-2	SER	-	CLONING ARTIFACT	UNP Q97S03
D	-1	ASN	-	CLONING ARTIFACT	UNP Q97S03
D	0	ALA	-	CLONING ARTIFACT	UNP Q97S03

*Continued on next page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	1	MSE	MET	MODIFIED RESIDUE	UNP Q97S03
D	109	MSE	MET	MODIFIED RESIDUE	UNP Q97S03
D	137	MSE	MET	MODIFIED RESIDUE	UNP Q97S03

- 
- The image displays the chemical structure of Flavin Mononucleotide (FMN). It features an isoalloxazine ring system, which is a tricyclic aromatic heterocycle consisting of a benzene ring fused to two pyrimidine rings. The atoms in the ring are labeled: N1, N3, N5, N9, C2, C4, C4A, C5A, C6, C7, C8, C8A, C9A, and C10. The ribityl chain is attached to the N10 position of the isoalloxazine ring. The ribityl chain consists of three carbon atoms (C1', C2', C3') and a phosphate group. The phosphate group is shown as a phosphorus atom (P) bonded to four oxygen atoms: one double-bonded (O1P), and three single-bonded (O2P, O3P, O4P). The ribityl chain is shown with stereochemistry: the hydroxyl group on C3' is wedged, and the hydroxyl group on C2' is dashed. The phosphate group is attached to C1' via an oxygen atom (O5').

- Molecule 3 is ACETIC ACID (three-letter code: ACY) (formula:  $C_2H_4O_2$ ).



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

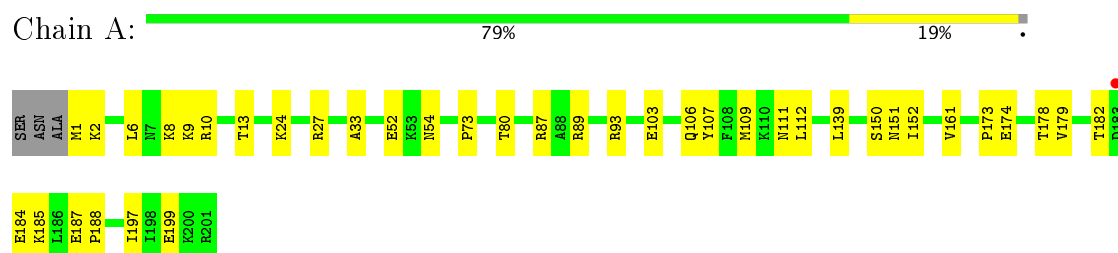
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	235	Total	O	0	0
			235	235		
4	B	170	Total	O	0	0
			170	170		
4	C	208	Total	O	0	0
			208	208		
4	D	140	Total	O	0	0
			140	140		

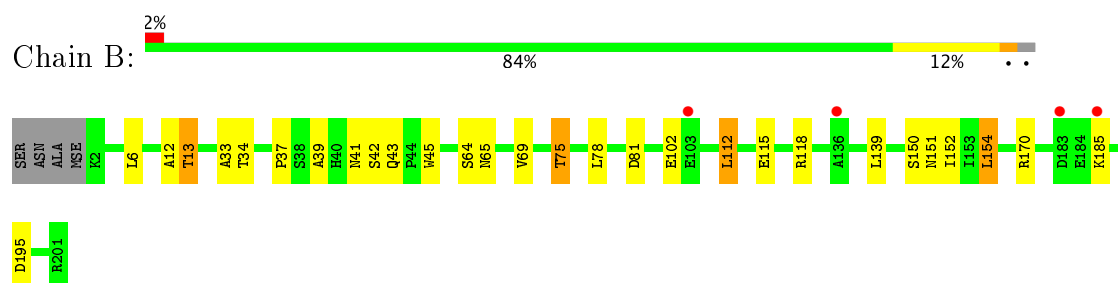
### 3 Residue-property plots [i](#)

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.

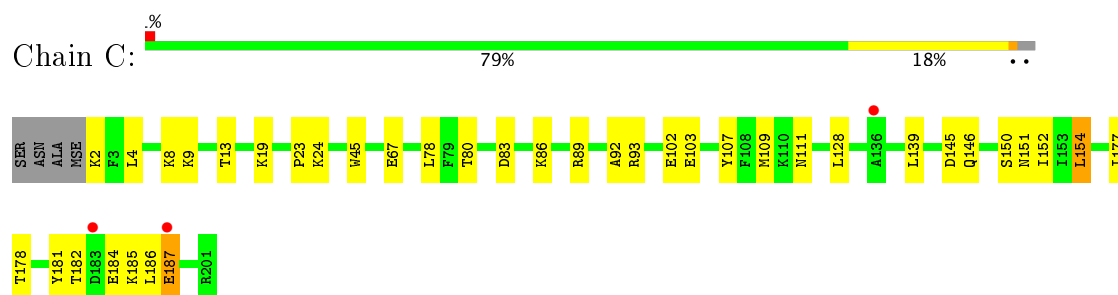
- Molecule 1: COG0778: Nitroreductase



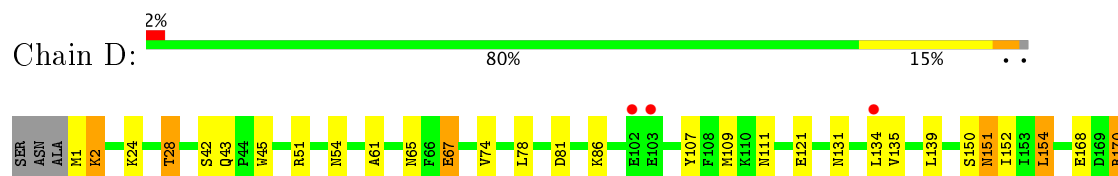
- Molecule 1: COG0778: Nitroreductase

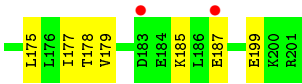


- Molecule 1: COG0778: Nitroreductase



- Molecule 1: COG0778: Nitroreductase







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	55.40 Å   80.28 Å   103.02 Å 90.00°   100.61°   90.00°	Depositor
Resolution (Å)	39.02 – 2.05 39.02 – 2.05	Depositor EDS
% Data completeness (in resolution range)	97.7 (39.02-2.05) 97.6 (39.02-2.05)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.21 (at 2.05 Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.177   ,   0.247 0.176   ,   0.247	Depositor DCC
$R_{free}$ test set	5451 reflections (11.15%)	DCC
Wilson B-factor (Å <sup>2</sup> )	22.2	Xtriage
Anisotropy	0.282	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 54.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	7396	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 65.13 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 7.4385e-06. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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 [i](#)

### 5.1 Standard geometry [i](#)

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

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	0.72	0/1664	0.72	2/2242 (0.1%)
1	B	0.69	0/1628	0.69	0/2196
1	C	0.70	0/1662	0.70	0/2242
1	D	0.62	0/1665	0.72	1/2243 (0.0%)
All	All	0.68	0/6619	0.71	3/8923 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	87	ARG	NE-CZ-NH1	5.59	123.10	120.30
1	A	87	ARG	NE-CZ-NH2	-5.33	117.63	120.30
1	D	170	ARG	NE-CZ-NH2	-5.26	117.67	120.30

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	1638	0	1659	27	0
1	B	1602	0	1619	30	0
1	C	1636	0	1650	37	0
1	D	1639	0	1668	33	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	31	0	19	7	0
2	B	31	0	19	3	0
2	C	31	0	19	5	0
2	D	31	0	19	6	0
3	A	4	0	3	0	0
4	A	235	0	0	10	0
4	B	170	0	0	5	0
4	C	208	0	0	6	0
4	D	140	0	0	5	0
All	All	7396	0	6675	119	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 9.

The worst 5 of 119 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:C:803:FMN:O2P	4:C:960:HOH:O	1.81	0.96
1:D:131:ASN:O	1:D:134[B]:LEU:HD22	1.65	0.94
1:B:154:LEU:HB2	2:B:802:FMN:H6	1.53	0.90
1:D:154:LEU:HB2	2:D:804:FMN:H6	1.55	0.88
1:D:74:VAL:HG13	1:D:179:VAL:HB	1.54	0.88

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	203/204 (100%)	199 (98%)	4 (2%)	0	100	100
1	B	199/204 (98%)	195 (98%)	4 (2%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	203/204 (100%)	197 (97%)	6 (3%)	0	100	100
1	D	203/204 (100%)	198 (98%)	5 (2%)	0	100	100
All	All	808/816 (99%)	789 (98%)	19 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	177/172 (103%)	170 (96%)	7 (4%)	36	27
1	B	173/172 (101%)	164 (95%)	9 (5%)	27	17
1	C	177/172 (103%)	164 (93%)	13 (7%)	16	8
1	D	177/172 (103%)	164 (93%)	13 (7%)	16	8
All	All	704/688 (102%)	662 (94%)	42 (6%)	24	13

5 of 42 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	45	TRP
1	C	154	LEU
1	D	168	GLU
1	C	78	LEU
1	C	102[B]	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	68	GLN
1	C	111	ASN
1	D	54	ASN
1	C	54	ASN
1	D	123	GLN

### 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 ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

5 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	FMN	A	801	-	31,33,33	1.49	5 (16%)	38,50,50	1.61	5 (13%)
3	ACY	A	805	-	1,3,3	1.44	0	0,3,3	0.00	-
2	FMN	B	802	-	31,33,33	1.35	5 (16%)	38,50,50	1.66	5 (13%)
2	FMN	C	803	-	31,33,33	1.43	5 (16%)	38,50,50	1.67	6 (15%)
2	FMN	D	804	-	31,33,33	1.48	5 (16%)	38,50,50	1.59	6 (15%)

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
2	FMN	A	801	-	-	0/16/18/18	0/3/3/3
3	ACY	A	805	-	-	0/0/0/0	0/0/0/0
2	FMN	B	802	-	-	0/16/18/18	0/3/3/3
2	FMN	C	803	-	-	0/16/18/18	0/3/3/3

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	FMN	D	804	-	-	0/16/18/18	0/3/3/3

The worst 5 of 20 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	802	FMN	C1'-N10	2.00	1.50	1.48
2	D	804	FMN	C1'-N10	2.01	1.50	1.48
2	B	802	FMN	C5A-N5	2.15	1.38	1.35
2	C	803	FMN	C5A-N5	2.32	1.38	1.35
2	A	801	FMN	C5A-N5	2.37	1.38	1.35

The worst 5 of 22 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	802	FMN	C4A-C4-N3	-3.63	118.31	123.48
2	C	803	FMN	C4A-C4-N3	-3.02	119.18	123.48
2	D	804	FMN	C4A-C4-N3	-2.69	119.65	123.48
2	A	801	FMN	C4A-C4-N3	-2.50	119.92	123.48
2	D	804	FMN	C1'-N10-C10	-2.02	116.43	118.50

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 21 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	801	FMN	7	0
2	B	802	FMN	3	0
2	C	803	FMN	5	0
2	D	804	FMN	6	0

## 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	198/204 (97%)	-0.25	1 (0%) 90 92	10, 18, 35, 46	0
1	B	198/204 (97%)	-0.15	4 (2%) 65 70	11, 24, 37, 42	0
1	C	198/204 (97%)	-0.28	3 (1%) 74 77	10, 20, 36, 46	0
1	D	198/204 (97%)	-0.08	5 (2%) 58 63	12, 28, 41, 47	0
All	All	792/816 (97%)	-0.19	13 (1%) 72 75	10, 22, 37, 47	0

The worst 5 of 13 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	134[A]	LEU	3.9
1	D	183	ASP	2.8
1	B	183	ASP	2.8
1	C	187[A]	GLU	2.4
1	D	103	GLU	2.3

### 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	FMN	A	801	31/31	0.87	0.37	6.70	20,25,27,27	31
2	FMN	B	802	31/31	0.83	0.33	6.24	14,21,23,23	31
2	FMN	D	804	31/31	0.88	0.31	4.65	13,22,23,25	31
2	FMN	C	803	31/31	0.86	0.33	3.19	14,23,28,28	31
3	ACY	A	805	4/4	0.99	0.09	-0.40	30,30,30,31	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.