



# Full wwPDB X-ray Structure Validation Report ⓘ

Sep 26, 2016 – 07:30 PM EDT

PDB ID : 5J62  
Title : FMN-dependent Nitroreductase (CDR20291\_0684) from *Clostridium difficile* R20291  
Authors : Wang, B.; Powell, S.M.; Hessami, N.; Najjar, F.Z.; Thomas, L.M.; West, A.H.; Karr, E.A.; Richter-Addo, G.B.  
Deposited on : 2016-04-04  
Resolution : 2.15 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

<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 : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20027939  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
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-20027939



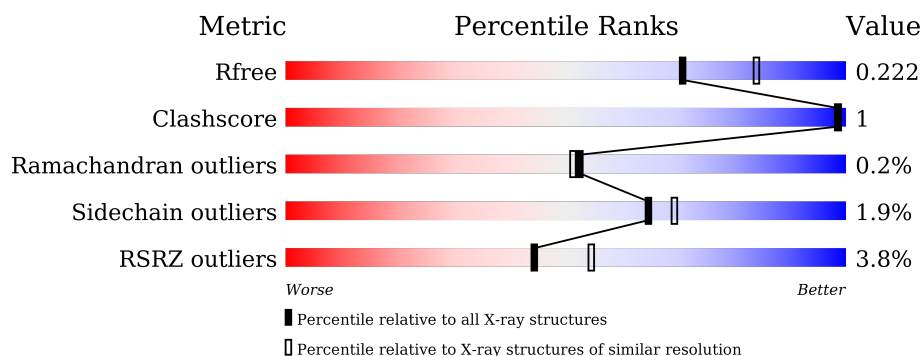
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.15 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))
$R_{free}$	91344	1045 (2.16-2.16)
Clashscore	102246	1152 (2.16-2.16)
Ramachandran outliers	100387	1131 (2.16-2.16)
Sidechain outliers	100360	1131 (2.16-2.16)
RSRZ outliers	91569	1050 (2.16-2.16)

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	231	<div> <div>3%</div> <div>86%</div> <div>10%</div> </div>
1	B	231	<div> <div>4%</div> <div>84%</div> <div>5%</div> <div>10%</div> </div>

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 criteria:



Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	GOL	A	302	-	-	-	X
4	PO4	A	303	-	-	-	X



## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 3478 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 Putative reductase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	208	Total	C	N	O	S	0	1	0
			1635	1034	275	321	5			
1	B	209	Total	C	N	O	S	0	0	0
			1632	1033	274	320	5			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	214	ALA	-	expression tag	UNP C9YJD4
A	215	GLU	-	expression tag	UNP C9YJD4
A	216	ASN	-	expression tag	UNP C9YJD4
A	217	LEU	-	expression tag	UNP C9YJD4
A	218	TYR	-	expression tag	UNP C9YJD4
A	219	PHE	-	expression tag	UNP C9YJD4
A	220	GLN	-	expression tag	UNP C9YJD4
A	221	GLY	-	expression tag	UNP C9YJD4
A	222	HIS	-	expression tag	UNP C9YJD4
A	223	HIS	-	expression tag	UNP C9YJD4
A	224	HIS	-	expression tag	UNP C9YJD4
A	225	HIS	-	expression tag	UNP C9YJD4
A	226	HIS	-	expression tag	UNP C9YJD4
A	227	HIS	-	expression tag	UNP C9YJD4
A	228	HIS	-	expression tag	UNP C9YJD4
A	229	HIS	-	expression tag	UNP C9YJD4
A	230	HIS	-	expression tag	UNP C9YJD4
A	231	HIS	-	expression tag	UNP C9YJD4
B	214	ALA	-	expression tag	UNP C9YJD4
B	215	GLU	-	expression tag	UNP C9YJD4
B	216	ASN	-	expression tag	UNP C9YJD4
B	217	LEU	-	expression tag	UNP C9YJD4
B	218	TYR	-	expression tag	UNP C9YJD4
B	219	PHE	-	expression tag	UNP C9YJD4
B	220	GLN	-	expression tag	UNP C9YJD4

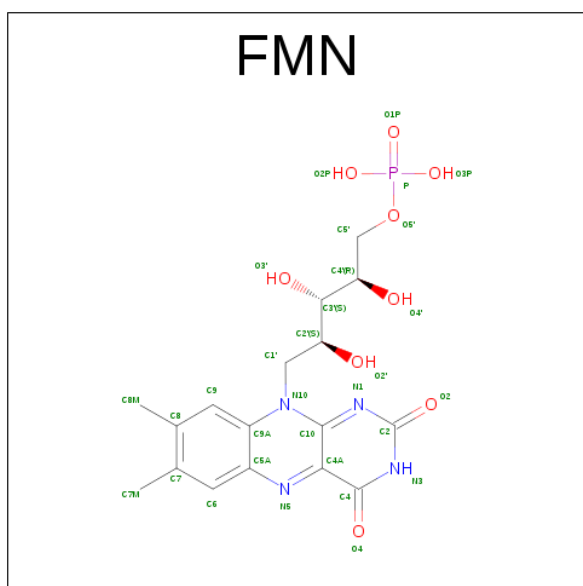
*Continued on next page...*



Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	221	GLY	-	expression tag	UNP C9YJD4
B	222	HIS	-	expression tag	UNP C9YJD4
B	223	HIS	-	expression tag	UNP C9YJD4
B	224	HIS	-	expression tag	UNP C9YJD4
B	225	HIS	-	expression tag	UNP C9YJD4
B	226	HIS	-	expression tag	UNP C9YJD4
B	227	HIS	-	expression tag	UNP C9YJD4
B	228	HIS	-	expression tag	UNP C9YJD4
B	229	HIS	-	expression tag	UNP C9YJD4
B	230	HIS	-	expression tag	UNP C9YJD4
B	231	HIS	-	expression tag	UNP C9YJD4

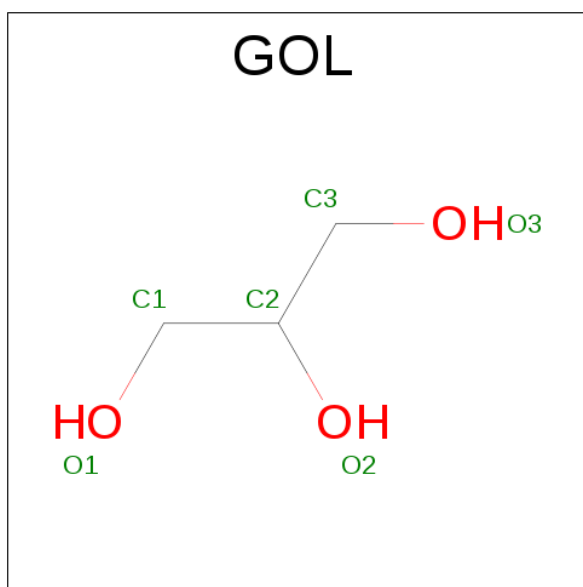
- Molecule 2 is FLAVIN MONONUCLEOTIDE (three-letter code: FMN) (formula:  $C_{17}H_{21}N_4O_9P$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			31	17	4	9	1		
2	B	1	Total	C	N	O	P	0	0
			31	17	4	9	1		

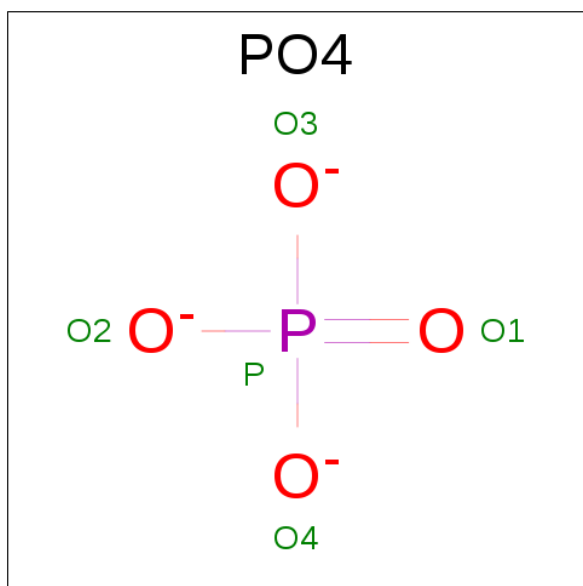
- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).





Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			6	3	3		

- Molecule 4 is PHOSPHATE ION (three-letter code: PO4) (formula:  $O_4P$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	O	P	0	0
			5	4	1		

- Molecule 5 is water.



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	74	Total 74	O 74	0	0
5	B	64	Total 64	O 64	0	0





- Molecule 1: Putative reductase



HIS	MET
HIS	ASU
HIS	ASN
	PHE
	GLN
	D7
	R17
	P44
	W52
	R73
	N79
	L96
	F97
	I98
	S104
	I111
	G112
	I113
	L114
	V159
	K160
	E161
	S162
	E163
	K166
	L179
	P188
	V207
	R212
	V213
	A214
	E215
	ASN
	LEU
	TYA
	PHE
	GLN
	GLY
	HIS
	HIS
	HIS
	HIS
	HIS



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	100.32Å 100.32Å 99.92Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	35.47 – 2.15 35.47 – 2.15	Depositor EDS
% Data completeness (in resolution range)	99.8 (35.47-2.15) 99.9 (35.47-2.15)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.84 (at 2.16Å)	Xtriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.167 , 0.217 0.175 , 0.222	Depositor DCC
$R_{free}$ test set	1415 reflections (5.26%)	DCC
Wilson B-factor (Å <sup>2</sup> )	39.7	Xtriage
Anisotropy	0.087	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 32.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.016 for -h,-l,-k 0.009 for l,-k,h	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	3478	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.10% 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: FMN, GOL, PO4

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	1/1664 (0.1%)	0.90	3/2256 (0.1%)
1	B	0.86	1/1661 (0.1%)	0.91	5/2252 (0.2%)
All	All	0.89	2/3325 (0.1%)	0.91	8/4508 (0.2%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	104	SER	CB-OG	5.48	1.49	1.42
1	A	105	SER	CB-OG	-5.47	1.35	1.42

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	17	ARG	NE-CZ-NH2	-6.89	116.85	120.30
1	A	105	SER	N-CA-CB	-6.08	101.38	110.50
1	B	212	ARG	NE-CZ-NH2	-6.07	117.26	120.30
1	B	73	ARG	NE-CZ-NH1	5.99	123.29	120.30
1	A	158	ASP	N-CA-C	5.88	126.87	111.00
1	B	73	ARG	CG-CD-NE	-5.33	100.60	111.80
1	B	73	ARG	NE-CZ-NH2	-5.21	117.69	120.30
1	A	110	ASP	CB-CG-OD1	5.11	122.89	118.30

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	1635	0	1630	3	0
1	B	1632	0	1630	4	0
2	A	31	0	19	0	0
2	B	31	0	19	0	0
3	A	6	0	8	0	0
4	A	5	0	0	0	0
5	A	74	0	0	0	0
5	B	64	0	0	0	0
All	All	3478	0	3306	6	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 (6) 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:52:TRP:HB3	1:B:111:ILE:HD11	1.86	0.57
1:B:44:PRO:HD2	1:B:113:ILE:HG21	1.91	0.52
1:A:50:GLN:HE22	1:B:207:VAL:HG22	1.83	0.44
1:B:98:ILE:O	1:B:188:PRO:HD2	2.20	0.41
1:A:50:GLN:N	1:A:51:PRO:HD3	2.36	0.41
1:A:8:ASN:CG	1:A:11:ILE:HD12	2.41	0.41

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	207/231 (90%)	205 (99%)	1 (0%)	1 (0%)	34 26

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	207/231 (90%)	206 (100%)	1 (0%)	0	100	100
All	All	414/462 (90%)	411 (99%)	2 (0%)	1 (0%)	52	51

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	158	ASP

### 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	185/205 (90%)	183 (99%)	2 (1%)	80	85
1	B	184/205 (90%)	179 (97%)	5 (3%)	52	53
All	All	369/410 (90%)	362 (98%)	7 (2%)	65	69

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	52	TRP
1	A	102	GLU
1	B	52	TRP
1	B	96	LEU
1	B	111	ILE
1	B	166	LYS
1	B	179	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	34	ASN
1	A	50	GLN
1	B	38	HIS

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	B	50	GLN
1	B	58	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 ⓘ

4 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	301	-	32,33,33	1.34	3 (9%)	34,50,50	3.01	6 (17%)
3	GOL	A	302	-	5,5,5	0.92	0	5,5,5	0.80	0
4	PO4	A	303	-	4,4,4	0.86	0	6,6,6	0.24	0
2	FMN	B	301	-	32,33,33	1.38	4 (12%)	34,50,50	3.22	7 (20%)

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	301	-	-	0/18/18/18	0/3/3/3
3	GOL	A	302	-	-	0/4/4/4	0/0/0/0
4	PO4	A	303	-	-	0/0/0/0	0/0/0/0
2	FMN	B	301	-	-	0/18/18/18	0/3/3/3

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	301	FMN	C10-N10	-3.98	1.34	1.39
2	B	301	FMN	C10-N1	-2.79	1.30	1.35
2	B	301	FMN	C10-N10	-2.23	1.36	1.39
2	A	301	FMN	P-O1P	3.37	1.61	1.50
2	A	301	FMN	C4-C4A	3.64	1.48	1.41
2	B	301	FMN	P-O1P	4.16	1.63	1.50
2	B	301	FMN	C4-C4A	4.40	1.50	1.41

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	301	FMN	C4A-C4-N3	-9.44	111.18	123.52
2	A	301	FMN	C4A-C4-N3	-8.38	112.57	123.52
2	A	301	FMN	C4A-C10-N10	-4.26	117.42	120.52
2	B	301	FMN	C4A-C10-N10	-4.01	117.61	120.52
2	B	301	FMN	N3-C2-N1	-4.00	120.95	127.69
2	A	301	FMN	N3-C2-N1	-3.98	120.99	127.69
2	A	301	FMN	C4-C4A-C10	-2.10	118.59	119.94
2	B	301	FMN	C4-C4A-C10	-2.05	118.63	119.94
2	B	301	FMN	C5A-C9A-N10	2.01	119.08	117.58
2	B	301	FMN	C1'-N10-C9A	2.45	121.67	118.83
2	A	301	FMN	O2P-P-O5'	2.83	114.97	106.72
2	A	301	FMN	C4-N3-C2	13.19	126.16	115.16
2	B	301	FMN	C4-N3-C2	14.38	127.16	115.16

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	208/231 (90%)	-0.20	6 (2%) 55 65	26, 38, 76, 103	0
1	B	209/231 (90%)	-0.04	10 (4%) 34 45	27, 43, 79, 100	0
All	All	417/462 (90%)	-0.12	16 (3%) 44 54	26, 40, 79, 103	0

All (16) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	7	ASP	4.7
1	B	163	GLU	4.1
1	A	213	VAL	3.9
1	A	160	LYS	3.5
1	B	214	ALA	3.4
1	B	166	LYS	3.0
1	B	161	GLU	2.7
1	B	114	LEU	2.6
1	A	81	GLU	2.6
1	B	79	ASN	2.6
1	A	6	GLN	2.5
1	A	7[A]	ASP	2.4
1	B	159	VAL	2.2
1	B	113	ILE	2.1
1	A	79	ASN	2.1
1	B	215	GLU	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. 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
4	PO4	A	303	5/5	0.89	0.23	5.71	50,54,57,59	5
3	GOL	A	302	6/6	0.90	0.21	5.11	38,57,60,61	0
2	FMN	A	301	31/31	0.97	0.11	-0.53	27,29,31,32	0
2	FMN	B	301	31/31	0.96	0.09	-0.74	32,36,40,41	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.