



# Full wwPDB X-ray Structure Validation Report

Feb 26, 2014 – 11:50 PM GMT

PDB ID : 3SFK  
Title : Crystal structure of Plasmodium falciparum dihydroorotate dehydrogenase bound with Inhibitor DSM267  
Authors : Deng, X.; Phillips, M.  
Deposited on : 2011-06-13  
Resolution : 2.95 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>

---

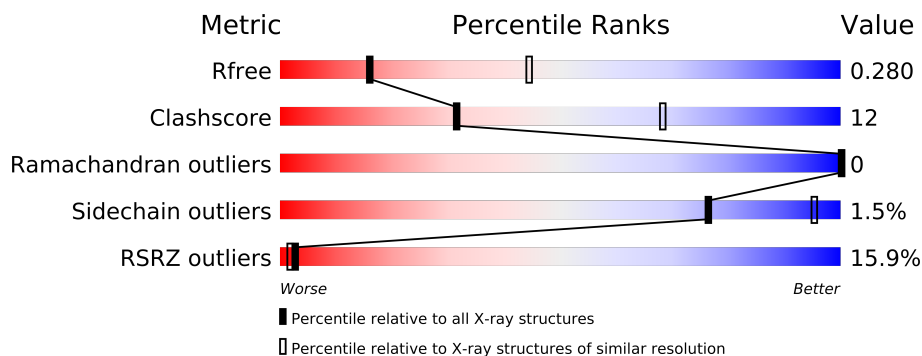
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
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) : stable22683

# 1 Overall quality at a glance


The reported resolution of this entry is 2.95 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	1587 (3.00-2.92)
Clashscore	79885	2029 (3.00-2.92)
Ramachandran outliers	78287	1955 (3.00-2.92)
Sidechain outliers	78261	1958 (3.00-2.92)
RSRZ outliers	66119	1588 (3.00-2.92)

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	400	

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 2993 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.

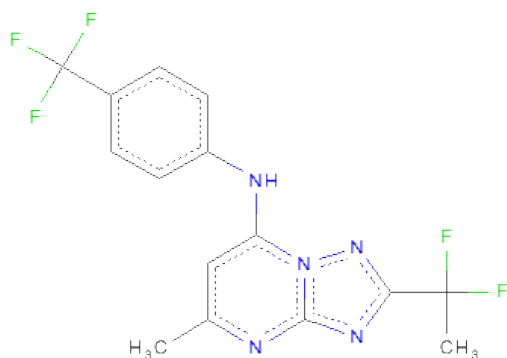
- Molecule 1 is a protein called Dihydroorotate dehydrogenase (quinone), mitochondrial.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	368	Total	C	N	O	S	0	0	0
			2921	1871	486	549	15			

There are 18 discrepancies between the modelled and reference sequences:

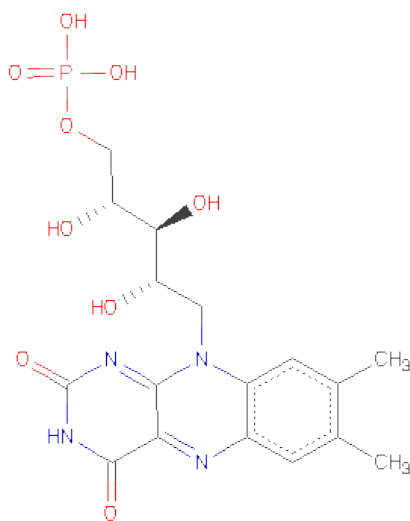
Chain	Residue	Modelled	Actual	Comment	Reference
A	140	MET	-	EXPRESSION TAG	UNP Q08210
A	141	GLY	-	EXPRESSION TAG	UNP Q08210
A	142	HIS	-	EXPRESSION TAG	UNP Q08210
A	143	HIS	-	EXPRESSION TAG	UNP Q08210
A	144	HIS	-	EXPRESSION TAG	UNP Q08210
A	145	HIS	-	EXPRESSION TAG	UNP Q08210
A	146	HIS	-	EXPRESSION TAG	UNP Q08210
A	147	HIS	-	EXPRESSION TAG	UNP Q08210
A	148	ALA	-	EXPRESSION TAG	UNP Q08210
A	149	GLU	-	EXPRESSION TAG	UNP Q08210
A	150	ASN	-	EXPRESSION TAG	UNP Q08210
A	151	LEU	-	EXPRESSION TAG	UNP Q08210
A	152	TYR	-	EXPRESSION TAG	UNP Q08210
A	153	PHE	-	EXPRESSION TAG	UNP Q08210
A	154	GLN	-	EXPRESSION TAG	UNP Q08210
A	155	GLY	-	EXPRESSION TAG	UNP Q08210
A	156	ALA	-	EXPRESSION TAG	UNP Q08210
A	157	PRO	-	EXPRESSION TAG	UNP Q08210

- Molecule 2 is 2-(1,1-DIFLUOROETHYL)-5-METHYL-N-[4-(TRIFLUOROMETHYL)PHENYL][1,2,4]TRIAZOLO[1,5-A]PYRIMIDIN-7-AMINE (three-letter code: D67) (formula: C<sub>15</sub>H<sub>12</sub>F<sub>5</sub>N<sub>5</sub>).



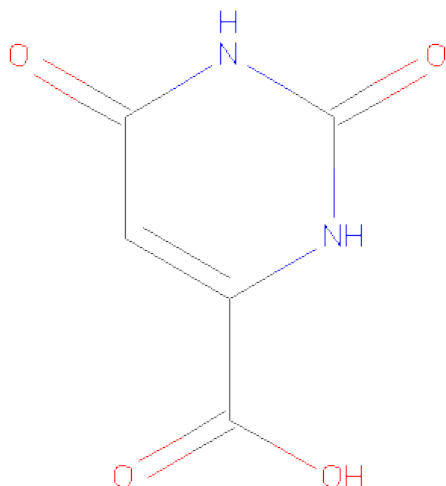
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	F	N		
2	A	1	25	15	5	5	0	0

- Molecule 3 is FLAVIN MONONUCLEOTIDE (three-letter code: FMN) (formula:  $C_{17}H_{21}N_4O_9P$ ).



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

- Molecule 4 is OROTIC ACID (three-letter code: ORO) (formula:  $C_5H_4N_2O_4$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	N	O	0	0
			11	5	2	4		

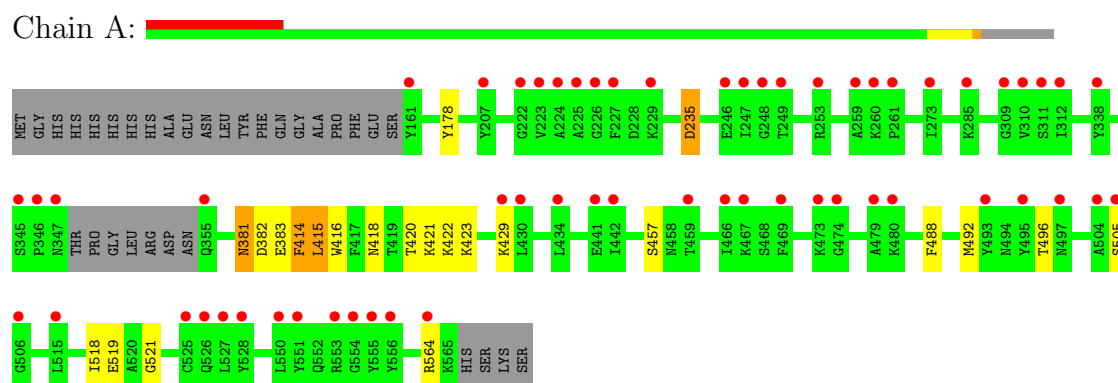
- Molecule 5 is water.

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

### 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: Dihydroorotate dehydrogenase (quinone), mitochondrial



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 64	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	86.59Å 86.59Å 138.17Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	50.00 – 2.95 39.24 – 2.89	Depositor EDS
% Data completeness (in resolution range)	99.7 (50.00-2.95) 99.7 (39.24-2.89)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.21 (at 2.90Å)	Xtriage
Refinement program	REFMAC 5.5.0072	Depositor
R, $R_{free}$	0.224 , 0.286 0.219 , 0.280	Depositor DCC
$R_{free}$ test set	646 reflections (5.17%)	DCC
Wilson B-factor (Å <sup>2</sup> )	98.4	Xtriage
Anisotropy	0.093	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 61.1	EDS
Estimated twinning fraction	0.109 for h,-h-k,-l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	0 of 13147 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	2993	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	114.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.28% 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: FMN, ORO, D67

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	0.40	0/2969	0.52	0/3995

There are no bond length outliers.

There are no bond angle outliers.

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	2921	0	0	35	1
2	A	25	0	12	0	0
3	A	31	0	19	1	0
4	A	11	0	3	0	0
5	A	5	0	0	0	0
All	All	2993	0	34	35	1

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 12.

All (35) close contacts within the same asymmetric unit are listed below.



Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:415:LEU:CD1	1:A:416:TRP:CD2	2.30	1.15
1:A:416:TRP:CD1	1:A:422:LYS:CA	2.30	1.13
1:A:415:LEU:CD1	1:A:416:TRP:CE3	2.35	1.09
1:A:415:LEU:CD1	1:A:416:TRP:CE2	2.37	1.07
1:A:416:TRP:CD1	1:A:421:LYS:C	2.33	1.02
1:A:415:LEU:CD1	1:A:416:TRP:CZ3	2.46	0.99
1:A:416:TRP:CD1	1:A:422:LYS:N	2.30	0.99
1:A:415:LEU:CD1	1:A:416:TRP:CZ2	2.53	0.91
1:A:415:LEU:CD1	1:A:416:TRP:CH2	2.54	0.91
1:A:414:PHE:N	1:A:414:PHE:CD2	2.42	0.85
1:A:416:TRP:CD1	1:A:421:LYS:O	2.30	0.84
1:A:416:TRP:NE1	1:A:422:LYS:C	2.30	0.84
1:A:416:TRP:CD1	1:A:422:LYS:C	2.52	0.83
1:A:416:TRP:CB	1:A:418:ASN:OD1	2.30	0.80
1:A:416:TRP:NE1	1:A:422:LYS:CA	2.46	0.76
1:A:416:TRP:NE1	1:A:423:LYS:N	2.37	0.73
1:A:416:TRP:C	1:A:418:ASN:OD1	2.30	0.70
1:A:518:ILE:O	1:A:521:GLY:N	2.30	0.64
1:A:416:TRP:CZ2	1:A:423:LYS:CA	2.81	0.63
1:A:418:ASN:N	1:A:418:ASN:OD1	2.30	0.62
1:A:457:SER:N	1:A:505:SER:O	2.33	0.62
1:A:429:LYS:NZ	3:A:1002:FMN:O2	2.33	0.61
1:A:420:THR:O	1:A:421:LYS:CB	2.53	0.57
1:A:416:TRP:CE2	1:A:423:LYS:N	2.73	0.56
1:A:519:GLU:O	1:A:564:ARG:N	2.39	0.55
1:A:382:ASP:O	1:A:415:LEU:N	2.39	0.55
1:A:235:ASP:OD2	1:A:235:ASP:N	2.42	0.53
1:A:416:TRP:CZ2	1:A:423:LYS:CB	2.92	0.51
1:A:381:ASN:N	1:A:381:ASN:ND2	2.61	0.48
1:A:416:TRP:O	1:A:418:ASN:OD1	2.31	0.47
1:A:416:TRP:CE2	1:A:423:LYS:CA	3.00	0.45
1:A:492:MET:O	1:A:496:THR:OG1	2.35	0.45
1:A:382:ASP:O	1:A:383:GLU:C	2.53	0.44
1:A:416:TRP:CE2	1:A:422:LYS:C	2.92	0.42
1:A:416:TRP:CA	1:A:418:ASN:OD1	2.68	0.41

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	Distance(Å)	Clash(Å)
1:A:178:TYR:O	1:A:381:ASN:ND2[6_554]	1.60	0.60

## 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	364/400 (91%)	349 (96%)	15 (4%)	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	326/353 (92%)	321 (98%)	5 (2%)	76 95

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	235	ASP
1	A	381	ASN
1	A	414	PHE
1	A	415	LEU
1	A	488	PHE

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 ⓘ

3 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	D67	A	1001	-	27,27,27	2.42	8 (29%)	42,42,42	1.78	5 (11%)
3	FMN	A	1002	-	33,33,33	0.93	2 (6%)	46,50,50	1.43	6 (13%)
4	ORO	A	1003	-	11,11,11	1.12	0	11,15,15	2.17	3 (27%)

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	D67	A	1001	-	-	0/14/16/16	0/1/3/3
3	FMN	A	1002	-	-	0/18/18/18	0/0/3/3
4	ORO	A	1003	-	-	0/4/4/4	0/1/1/1

All (10) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1001	D67	F4-C13	-7.23	1.31	1.37
2	A	1001	D67	F5-C13	-6.91	1.32	1.37
2	A	1001	D67	C11-N2	-4.49	1.33	1.38
2	A	1001	D67	C11-N5	3.02	1.35	1.32

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1001	D67	N3-N2	2.34	1.39	1.36
2	A	1001	D67	C2-N1	-2.33	1.35	1.40
3	A	1002	FMN	C1'-N10	2.25	1.50	1.48
2	A	1001	D67	C12-N3	2.16	1.35	1.33
2	A	1001	D67	C7-N1	-2.09	1.35	1.39
3	A	1002	FMN	C5A-N5	2.04	1.38	1.35

All (14) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1001	D67	F5-C13-F4	7.34	108.50	104.67
3	A	1002	FMN	C2-N1-C10	4.60	119.62	114.98
4	A	1003	ORO	C4-N3-C2	-3.93	120.73	125.73
2	A	1001	D67	N1-C7-N2	3.91	118.44	113.69
4	A	1003	ORO	C5-C4-N3	3.61	119.62	116.31
3	A	1002	FMN	C4A-N5-C5A	3.55	120.68	116.69
4	A	1003	ORO	C2-N1-C6	-3.38	120.69	123.46
2	A	1001	D67	C11-N2-N3	3.37	113.52	110.57
2	A	1001	D67	C12-N3-N2	-2.76	98.53	102.32
3	A	1002	FMN	C4A-C10-N1	-2.73	120.00	122.73
3	A	1002	FMN	C4-N3-C2	-2.71	119.83	125.39
3	A	1002	FMN	C5A-C9A-N10	2.69	119.46	116.80
3	A	1002	FMN	C1'-N10-C9A	2.36	121.17	118.87
2	A	1001	D67	C7-N2-C11	-2.00	120.83	122.82

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	368/400 (92%)	0.78	59 (16%) 3 2	74, 113, 154, 183	0

All (59) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	556	TYR	7.3
1	A	207	TYR	6.9
1	A	224	ALA	6.3
1	A	555	TYR	5.8
1	A	550	LEU	5.6
1	A	441	GLU	5.4
1	A	311	SER	5.2
1	A	225	ALA	5.2
1	A	469	PHE	4.9
1	A	222	GLY	4.5
1	A	551	TYR	4.4
1	A	505	SER	4.3
1	A	526	GLN	4.1
1	A	346	PRO	3.9
1	A	249	THR	3.9
1	A	347	ASN	3.8
1	A	246	GLU	3.8
1	A	506	GLY	3.8
1	A	355	GLN	3.8
1	A	223	VAL	3.7
1	A	259	ALA	3.7
1	A	466	ILE	3.7
1	A	261	PRO	3.7
1	A	515	LEU	3.6
1	A	248	GLY	3.4
1	A	525	CYS	3.3
1	A	495	TYR	3.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	442	ILE	3.2
1	A	480	LYS	3.1
1	A	345	SER	3.0
1	A	474	GLY	3.0
1	A	273	ILE	2.9
1	A	310	VAL	2.9
1	A	493	TYR	2.8
1	A	527	LEU	2.8
1	A	285	LYS	2.6
1	A	226	GLY	2.6
1	A	554	GLY	2.6
1	A	309	GLY	2.5
1	A	161	TYR	2.5
1	A	338	TYR	2.5
1	A	430	LEU	2.4
1	A	459	THR	2.4
1	A	467	LYS	2.4
1	A	253	ARG	2.3
1	A	564	ARG	2.3
1	A	247	ILE	2.3
1	A	312	ILE	2.2
1	A	479	ALA	2.2
1	A	504	ALA	2.2
1	A	434	LEU	2.1
1	A	528	TYR	2.1
1	A	553	ARG	2.1
1	A	229	LYS	2.1
1	A	473	LYS	2.1
1	A	260	LYS	2.1
1	A	227	PHE	2.0
1	A	429	LYS	2.0
1	A	497	ASN	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(Å <sup>2</sup> )	Q<0.9
3	FMN	A	1002	31/31	0.37	0.16	79,81,88,89	0
2	D67	A	1001	25/25	0.18	-0.50	90,94,97,97	0
4	ORO	A	1003	11/11	0.21	-0.67	95,96,97,98	0

## 6.5 Other polymers

There are no such residues in this entry.