



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 19, 2016 – 07:00 PM GMT

PDB ID : 4CW5  
Title : Crystal structure of the enoyl reductase domain of DfnA from *Bacillus amyloliquefaciens*  
Authors : Jakob, R.P.; Buhkari, H.S.T.; Maier, T.  
Deposited on : 2014-04-01  
Resolution : 2.30 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-20026982  
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-20026982

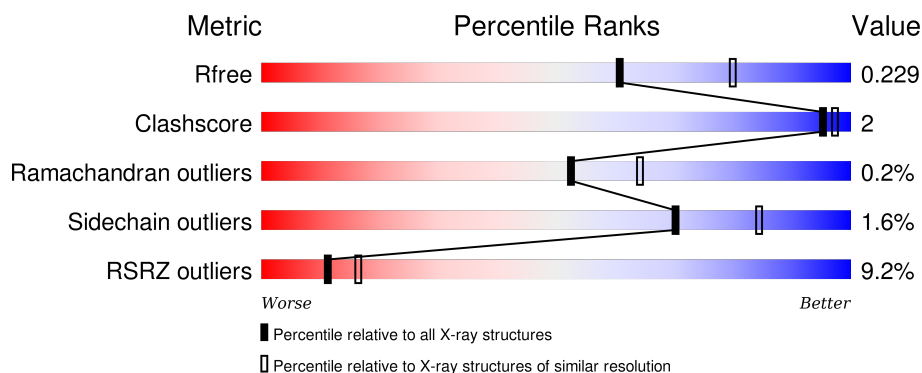
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.30 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	3852 (2.30-2.30)
Clashscore	102246	4452 (2.30-2.30)
Ramachandran outliers	100387	4410 (2.30-2.30)
Sidechain outliers	100360	4409 (2.30-2.30)
RSRZ outliers	91569	3857 (2.30-2.30)

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	454	<div> <div>4%</div> <div>90%</div> <div>6%</div> <div>.</div> </div>
1	B	454	<div> <div>13%</div> <div>93%</div> <div>.</div> <div>.</div> </div>

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 13855 atoms, of which 6893 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.

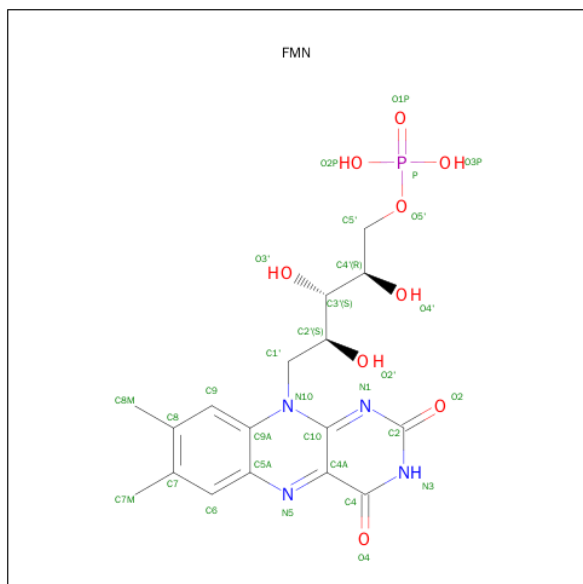
- Molecule 1 is a protein called DFNA.

Mol	Chain	Residues	Atoms								ZeroOcc	AltConf	Trace
1	A	437	Total	C	H	N	O	S	Se		0	1	0
			6830	2148	3430	598	631	3	20				
1	B	438	Total	C	H	N	O	S	Se		0	0	0
			6827	2148	3425	599	633	3	19				

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	299	SER	-	EXPRESSION TAG	UNP A7Z6E3
A	300	MSE	-	EXPRESSION TAG	UNP A7Z6E3
B	299	SER	-	EXPRESSION TAG	UNP A7Z6E3
B	300	MSE	-	EXPRESSION TAG	UNP A7Z6E3

- 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	H	N	O	P	0	0
			50	17	19	4	9	1		
2	B	1	Total	C	H	N	O	P	0	0
			50	17	19	4	9	1		

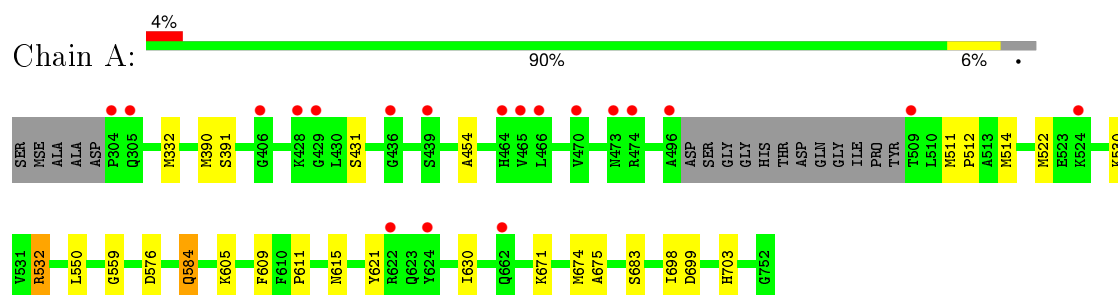
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	34	Total	O	0	0
			34	34		
3	B	64	Total	O	0	0
			64	64		

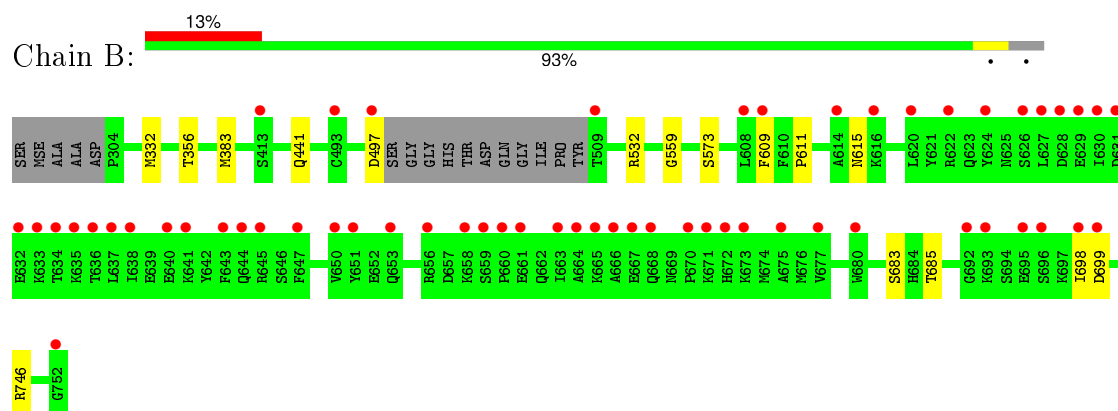
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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 ( $\text{RSRZ} > 2$ ). 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: DFNA



#### • Molecule 1: DFNA



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	80.85Å 94.00Å 144.53Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.00 – 2.30 47.00 – 2.25	Depositor EDS
% Data completeness (in resolution range)	100.0 (47.00-2.30) 99.9 (47.00-2.25)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.70 (at 2.24Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, $R_{free}$	0.204 , 0.231 0.203 , 0.229	Depositor DCC
$R_{free}$ test set	2505 reflections (5.05%)	DCC
Wilson B-factor (Å <sup>2</sup> )	46.7	Xtriage
Anisotropy	0.147	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 39.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 52945 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	13855	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	81.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.01% 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, 0.375 respectively for untwinned datasets, and 0.333, 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

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	0.24	0/3444	0.45	0/4597
1	B	0.23	0/3443	0.44	0/4598
All	All	0.24	0/6887	0.45	0/9195

There are no bond length outliers.

There are no bond angle outliers.

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	3400	3430	3430	16	0
1	B	3402	3425	3425	7	0
2	A	31	19	19	2	0
2	B	31	19	19	1	0
3	A	34	0	0	3	0
3	B	64	0	0	1	0
All	All	6962	6893	6893	22	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 2.

All (22) 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:A:550:LEU:O	1:B:746:ARG:NH1	2.20	0.75
1:B:685:THR:OG1	1:B:699:ASP:O	2.07	0.73
1:A:605:LYS:NZ	3:A:2024:HOH:O	2.16	0.65
1:A:454:ALA:CB	1:A:514:MSE:HE1	2.29	0.62
1:B:611:PRO:O	1:B:615:ASN:ND2	2.33	0.61
1:A:611:PRO:O	1:A:615:ASN:ND2	2.35	0.59
1:A:511:MSE:HB3	1:A:512:PRO:HD3	1.90	0.53
1:A:609:PHE:CD1	1:A:698:ILE:HG23	2.45	0.52
1:A:699:ASP:HB2	3:A:2027:HOH:O	2.08	0.52
1:A:576:ASP:OD1	3:A:2017:HOH:O	2.20	0.48
1:B:609:PHE:CD1	1:B:698:ILE:HG23	2.49	0.47
1:B:441:GLN:NE2	3:B:2039:HOH:O	2.36	0.46
1:A:584:GLN:HE21	1:A:584:GLN:H	1.62	0.46
1:A:332:MSE:HA	2:A:1753:FMN:N5	2.31	0.45
1:B:332:MSE:HA	2:B:1753:FMN:N5	2.32	0.44
1:A:703:HIS:HB2	2:A:1753:FMN:HM72	2.02	0.42
1:A:454:ALA:HB2	1:A:514:MSE:HE1	1.99	0.42
1:A:522:MSE:CE	1:A:530:LYS:HA	2.49	0.42
1:B:356:THR:CG2	1:B:383:MSE:HE3	2.50	0.41
1:A:621:TYR:CE1	1:A:675:ALA:HB2	2.56	0.41
1:A:630:ILE:HD11	1:A:674:MSE:HE3	2.03	0.40
1:A:532:ARG:H	1:A:532:ARG:HD2	1.87	0.40

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	434/454 (96%)	424 (98%)	9 (2%)	1 (0%)	52	64
1	B	434/454 (96%)	424 (98%)	9 (2%)	1 (0%)	52	64
All	All	868/908 (96%)	848 (98%)	18 (2%)	2 (0%)	52	64



All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	559	GLY
1	A	559	GLY

### 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	352/343 (103%)	345 (98%)	7 (2%)	63	79
1	B	352/343 (103%)	348 (99%)	4 (1%)	80	90
All	All	704/686 (103%)	693 (98%)	11 (2%)	70	84

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	390	MSE
1	A	391	SER
1	A	431	SER
1	A	532	ARG
1	A	584	GLN
1	A	671	LYS
1	A	683	SER
1	B	497	ASP
1	B	532	ARG
1	B	573	SER
1	B	683	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	582	ASN
1	A	584	GLN
1	A	713	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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

2 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	1753	-	31,33,33	1.09	2 (6%)	32,50,50	1.82	3 (9%)
2	FMN	B	1753	-	31,33,33	1.06	2 (6%)	32,50,50	1.79	3 (9%)

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	1753	-	-	0/18/18/18	0/3/3/3
2	FMN	B	1753	-	-	0/18/18/18	0/3/3/3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1753	FMN	C1'-N10	2.31	1.50	1.48

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1753	FMN	C1'-N10	2.46	1.51	1.48
2	B	1753	FMN	C4A-N5	2.92	1.37	1.33
2	A	1753	FMN	C4A-N5	3.19	1.38	1.33

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1753	FMN	C4A-N5-C5A	2.61	119.80	116.72
2	B	1753	FMN	C4A-N5-C5A	2.63	119.82	116.72
2	A	1753	FMN	C5A-C9A-N10	3.34	120.08	117.58
2	B	1753	FMN	C5A-C9A-N10	3.36	120.09	117.58
2	B	1753	FMN	C2-N1-C10	8.36	120.52	113.39
2	A	1753	FMN	C2-N1-C10	8.60	120.73	113.39

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1753	FMN	2	0
2	B	1753	FMN	1	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

### 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	418/454 (92%)	0.33	19 (4%) 37 46	35, 70, 120, 159	0
1	B	419/454 (92%)	0.69	58 (13%) 4 6	30, 56, 147, 175	0
All	All	837/908 (92%)	0.51	77 (9%) 11 16	30, 63, 138, 175	0

All (77) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	627	LEU	8.8
1	B	497	ASP	7.4
1	B	630	ILE	5.7
1	B	651	TYR	5.5
1	B	670	PRO	5.2
1	B	656	ARG	4.9
1	B	664	ALA	4.8
1	B	620	LEU	4.5
1	B	637	LEU	4.4
1	B	660	PRO	4.4
1	B	668	GLN	4.1
1	B	636	THR	4.1
1	B	695	GLU	4.0
1	B	628	ASP	4.0
1	B	643	PHE	3.9
1	B	616	LYS	3.9
1	B	509	THR	3.8
1	B	666	ALA	3.7
1	B	609	PHE	3.7
1	B	641	LYS	3.6
1	B	638	ILE	3.5
1	A	466	LEU	3.5
1	B	624	TYR	3.4
1	A	304	PRO	3.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	496	ALA	3.3
1	B	650	VAL	3.3
1	B	614	ALA	3.2
1	A	473	ASN	3.1
1	B	632	GLU	3.0
1	B	644	GLN	3.0
1	A	464	HIS	3.0
1	A	622	ARG	3.0
1	B	693	LYS	2.9
1	B	629	GLU	2.9
1	A	470	VAL	2.9
1	B	692	GLY	2.9
1	B	752	GLY	2.8
1	B	671	LYS	2.8
1	B	626	SER	2.8
1	A	406	GLY	2.7
1	B	663	ILE	2.7
1	B	665	LYS	2.7
1	B	698	ILE	2.7
1	B	659	SER	2.7
1	B	661	GLU	2.6
1	A	429	GLY	2.6
1	B	673	LYS	2.6
1	B	696	SER	2.6
1	B	608	LEU	2.6
1	B	647	PHE	2.5
1	B	640	GLU	2.5
1	B	658	LYS	2.5
1	B	675	ALA	2.5
1	B	680	TRP	2.5
1	A	436	GLY	2.5
1	A	474	ARG	2.4
1	B	633	LYS	2.4
1	B	667	GLU	2.3
1	A	509	THR	2.3
1	B	634	THR	2.3
1	B	493	CYS	2.2
1	A	305	GLN	2.2
1	A	465	VAL	2.2
1	B	635	LYS	2.2
1	A	524	LYS	2.1
1	B	645	ARG	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	631	ASP	2.1
1	B	413	SER	2.1
1	A	662	GLN	2.1
1	B	653	GLN	2.1
1	B	622	ARG	2.1
1	B	677	VAL	2.1
1	A	428	LYS	2.0
1	A	439	SER	2.0
1	B	672	HIS	2.0
1	B	699	ASP	2.0
1	A	624	TYR	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(Å <sup>2</sup> )	Q<0.9
2	FMN	B	1753	31/31	0.96	0.21	0.07	29,38,49,50	0
2	FMN	A	1753	31/31	0.96	0.14	-0.16	37,44,55,66	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.