



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 19, 2016 – 07:12 PM GMT

PDB ID : 4E2D  
Title : Structure of the old yellow enzyme from Trypanosoma cruzi  
Authors : Murakami, M.T.; Rodrigues, N.C.; Gava, L.M.; Canduri, F.; Oliva, G.; Barbosa, L.R.S.; Borgers, J.C.  
Deposited on : 2012-03-08  
Resolution : 2.00 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 : 1.7.1 (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

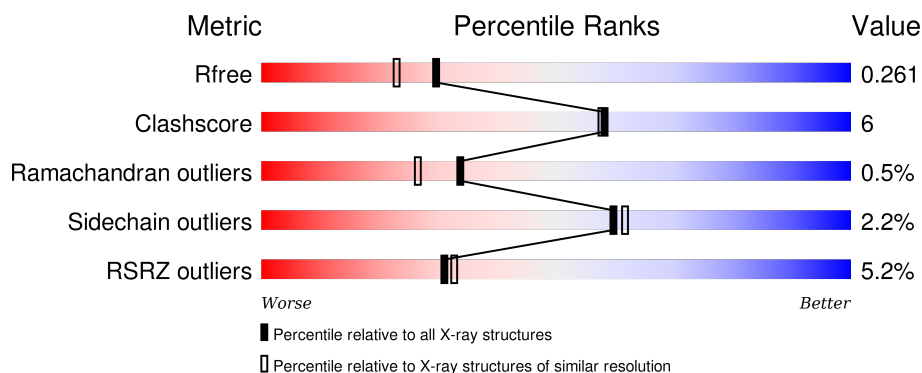
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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	6249 (2.00-2.00)
Clashscore	102246	7340 (2.00-2.00)
Ramachandran outliers	100387	7248 (2.00-2.00)
Sidechain outliers	100360	7247 (2.00-2.00)
RSRZ outliers	91569	6262 (2.00-2.00)

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	382	<div> <div>4%</div> <div>84%</div> <div>14%</div> <div>..</div> </div>
1	B	382	<div> <div>5%</div> <div>84%</div> <div>14%</div> <div>..</div> </div>
1	C	382	<div> <div>7%</div> <div>86%</div> <div>12%</div> <div>..</div> </div>
1	D	382	<div> <div>5%</div> <div>84%</div> <div>13%</div> <div>..</div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 12421 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.

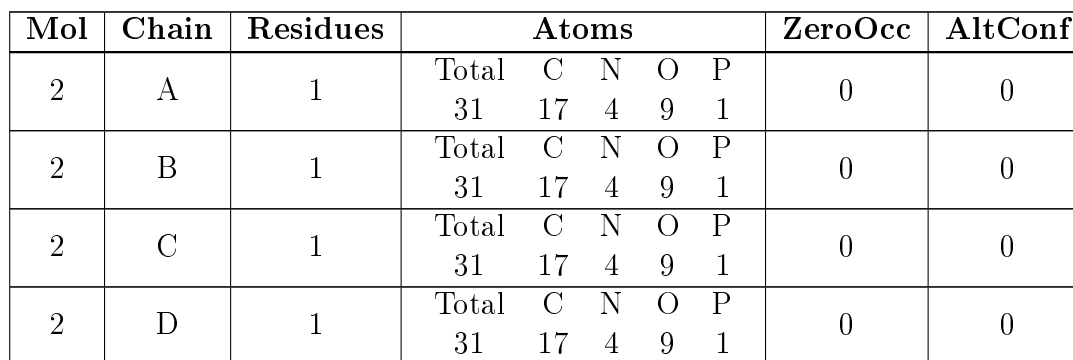
- Molecule 1 is a protein called Old Yellow Protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	377	Total	C	N	O	S	0	2	0
			2976	1886	523	558	9			
1	B	377	Total	C	N	O	S	3	2	0
			2971	1883	520	559	9			
1	C	377	Total	C	N	O	S	5	2	0
			2973	1885	520	558	10			
1	D	376	Total	C	N	O	S	0	0	0
			2952	1872	515	556	9			

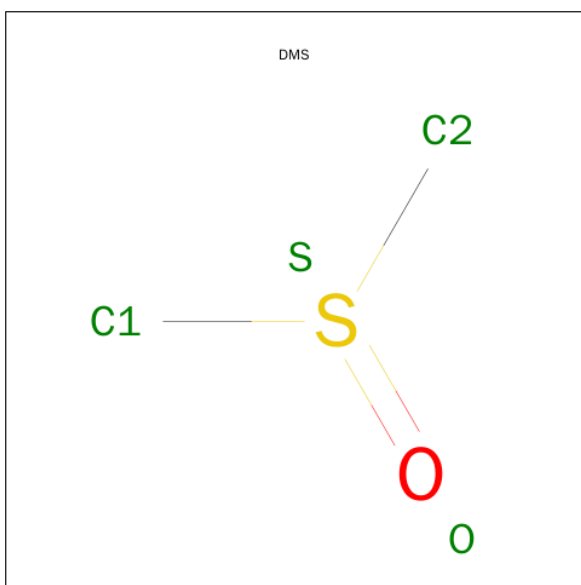
There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	GLY	-	EXPRESSION TAG	UNP Q2TJB8
A	-1	SER	-	EXPRESSION TAG	UNP Q2TJB8
A	0	HIS	-	EXPRESSION TAG	UNP Q2TJB8
B	-2	GLY	-	EXPRESSION TAG	UNP Q2TJB8
B	-1	SER	-	EXPRESSION TAG	UNP Q2TJB8
B	0	HIS	-	EXPRESSION TAG	UNP Q2TJB8
C	-2	GLY	-	EXPRESSION TAG	UNP Q2TJB8
C	-1	SER	-	EXPRESSION TAG	UNP Q2TJB8
C	0	HIS	-	EXPRESSION TAG	UNP Q2TJB8
D	-2	GLY	-	EXPRESSION TAG	UNP Q2TJB8
D	-1	SER	-	EXPRESSION TAG	UNP Q2TJB8
D	0	HIS	-	EXPRESSION TAG	UNP Q2TJB8

- Molecule 2 is FLAVIN MONONUCLEOTIDE (three-letter code: FMN) (formula: C<sub>17</sub>H<sub>21</sub>N<sub>4</sub>O<sub>9</sub>P).



- Molecule 3 is DIMETHYL SULFOXIDE (three-letter code: DMS) (formula:  $\text{C}_2\text{H}_6\text{OS}$ ).



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

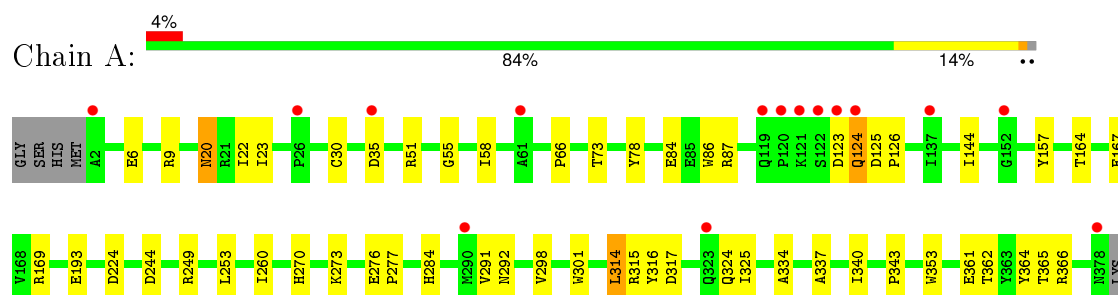
- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	109	Total O 109 109	0	0
4	B	91	Total O 91 91	0	0
4	C	107	Total O 107 107	0	0
4	D	102	Total O 102 102	0	0

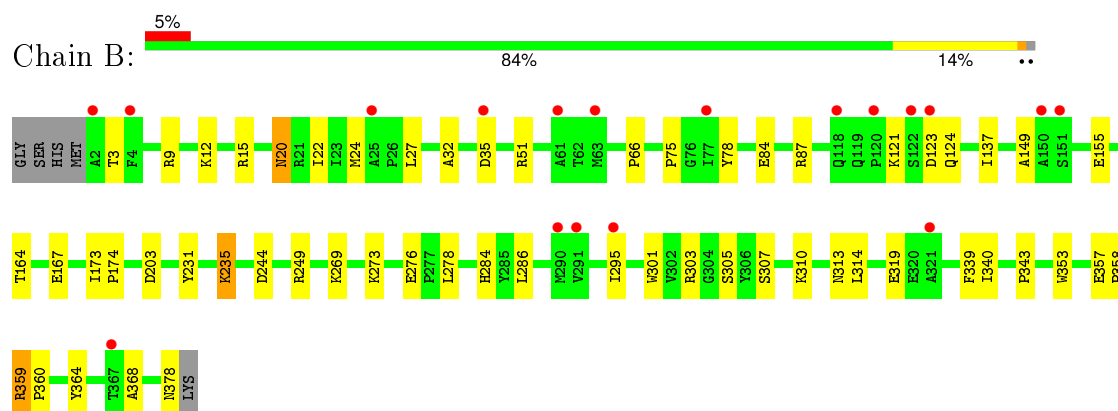
### 3 Residue-property plots [i](#)

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.

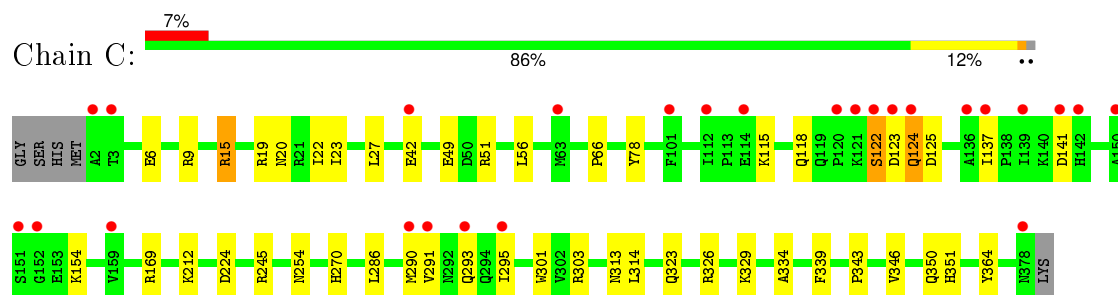
#### • Molecule 1: Old Yellow Protein



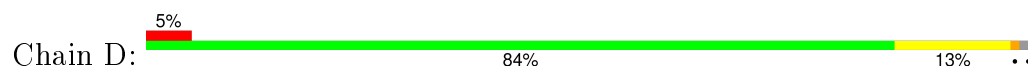
#### • Molecule 1: Old Yellow Protein

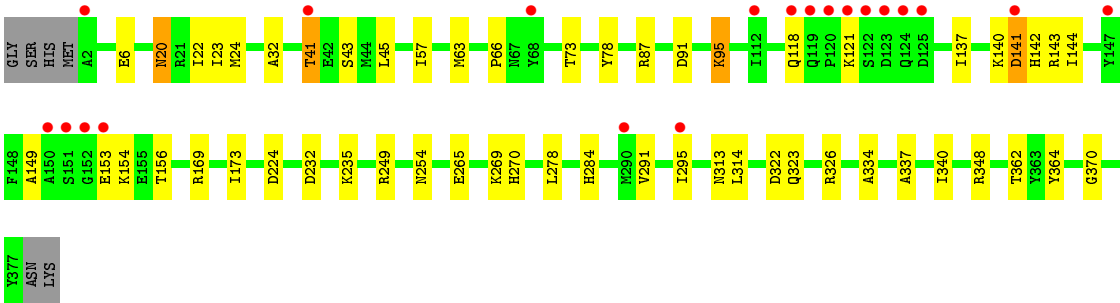


#### • Molecule 1: Old Yellow Protein



#### • Molecule 1: Old Yellow Protein





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	50.77Å 118.47Å 111.12Å 90.00° 90.09° 90.00°	Depositor
Resolution (Å)	40.00 – 2.00 36.43 – 2.00	Depositor EDS
% Data completeness (in resolution range)	90.1 (40.00-2.00) 91.3 (36.43-2.00)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.63 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, $R_{free}$	0.194 , 0.242 0.209 , 0.261	Depositor DCC
$R_{free}$ test set	4055 reflections (5.25%)	DCC
Wilson B-factor (Å <sup>2</sup> )	23.9	Xtriage
Anisotropy	0.036	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 41.9	EDS
Estimated twinning fraction	0.911 for H, K, L 0.089 for -H, -K, L 0.006 for h,-k,-l	Xtriage
Reported twinning fraction	0.911 for H, K, L 0.089 for -H, -K, L	Depositor
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.53$ , $\langle L^2 \rangle = 0.37$	Xtriage
Outliers	0 of 81443 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	12421	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	24.0	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 31.52 % 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 1.0920e-03. 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, 0.375 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, DMS

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	0.77	3/3049 (0.1%)	0.81	1/4135 (0.0%)
1	B	0.74	2/3044 (0.1%)	0.78	1/4130 (0.0%)
1	C	0.75	1/3046 (0.0%)	0.82	5/4132 (0.1%)
1	D	0.72	0/3019	0.79	2/4096 (0.0%)
All	All	0.75	6/12158 (0.0%)	0.80	9/16493 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	B	0	2

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	301	TRP	CD2-CE2	5.66	1.48	1.41
1	A	301	TRP	CD2-CE2	5.48	1.48	1.41
1	C	301	TRP	CD2-CE2	5.42	1.47	1.41
1	B	353	TRP	CD2-CE2	5.14	1.47	1.41
1	A	86	TRP	CD2-CE2	5.07	1.47	1.41

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	15	ARG	NE-CZ-NH2	-8.25	116.18	120.30
1	D	232	ASP	CB-CG-OD1	6.35	124.02	118.30
1	C	245	ARG	NE-CZ-NH1	-6.22	117.19	120.30
1	B	203	ASP	CB-CG-OD1	5.94	123.65	118.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	A	317	ASP	CB-CG-OD1	5.62	123.35	118.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	123	ASP	Peptide
1	B	124	GLN	Peptide

## 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	2976	0	2950	32	1
1	B	2971	0	2942	39	0
1	C	2973	0	2946	34	1
1	D	2952	0	2918	37	0
2	A	31	0	19	1	0
2	B	31	0	19	4	0
2	C	31	0	19	3	0
2	D	31	0	19	2	0
3	A	4	0	6	0	0
3	B	4	0	6	1	0
3	C	4	0	6	0	0
3	D	4	0	6	0	0
4	A	109	0	0	6	0
4	B	91	0	0	8	0
4	C	107	0	0	3	0
4	D	102	0	0	5	0
All	All	12421	0	11856	140	1

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 140 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:D:295:ILE:HB	4:D:576:HOH:O	1.38	1.23
1:B:295:ILE:HB	4:B:555:HOH:O	1.37	1.19
1:B:358:PRO:O	1:B:360:PRO:HD3	1.46	1.16
1:C:122:SER:HB2	1:C:123:ASP:HA	1.30	1.09
1:D:235:LYS:HE3	4:D:582:HOH:O	1.56	1.05

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	Interatomic distance (Å)	Clash overlap (Å)
1:A:325:ILE:O	1:C:15:ARG:NH2[2_546]	2.14	0.06

## 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	377/382 (99%)	362 (96%)	13 (3%)	2 (0%)	34	26
1	B	377/382 (99%)	365 (97%)	10 (3%)	2 (0%)	34	26
1	C	377/382 (99%)	360 (96%)	15 (4%)	2 (0%)	34	26
1	D	374/382 (98%)	356 (95%)	16 (4%)	2 (0%)	34	26
All	All	1505/1528 (98%)	1443 (96%)	54 (4%)	8 (0%)	34	26

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	314	LEU
1	B	314	LEU
1	C	122	SER
1	C	314	LEU
1	D	314	LEU

### 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	313/315 (99%)	309 (99%)	4 (1%)	76	79
1	B	313/315 (99%)	305 (97%)	8 (3%)	54	54
1	C	313/315 (99%)	305 (97%)	8 (3%)	54	54
1	D	310/315 (98%)	301 (97%)	9 (3%)	50	49
All	All	1249/1260 (99%)	1220 (98%)	29 (2%)	60	60

5 of 29 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	124	GLN
1	C	290[A]	MET
1	D	143	ARG
1	C	141	ASP
1	C	290[B]	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 28 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	20	ASN
1	C	198	ASN
1	D	270	HIS
1	C	118	GLN
1	C	142	HIS

### 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](#)

8 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	401	-	32,33,33	1.34	5 (15%)	34,50,50	2.29	9 (26%)
3	DMS	A	402	-	3,3,3	0.57	0	3,3,3	0.65	0
2	FMN	B	401	-	32,33,33	1.69	7 (21%)	34,50,50	2.00	9 (26%)
3	DMS	B	402	-	3,3,3	0.53	0	3,3,3	0.90	0
2	FMN	C	401	-	32,33,33	1.66	7 (21%)	34,50,50	2.72	13 (38%)
3	DMS	C	402	-	3,3,3	0.47	0	3,3,3	0.69	0
2	FMN	D	401	-	32,33,33	1.70	5 (15%)	34,50,50	2.44	10 (29%)
3	DMS	D	402	-	3,3,3	0.46	0	3,3,3	0.51	0

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	401	-	-	0/18/18/18	0/3/3/3
3	DMS	A	402	-	-	0/0/0/0	0/0/0/0
2	FMN	B	401	-	-	0/18/18/18	0/3/3/3
3	DMS	B	402	-	-	0/0/0/0	0/0/0/0
2	FMN	C	401	-	-	0/18/18/18	0/3/3/3
3	DMS	C	402	-	-	0/0/0/0	0/0/0/0
2	FMN	D	401	-	-	0/18/18/18	0/3/3/3
3	DMS	D	402	-	-	0/0/0/0	0/0/0/0

The worst 5 of 24 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	401	FMN	O5'-C5'	-2.03	1.36	1.44
2	B	401	FMN	C4A-N5	2.28	1.36	1.33
2	C	401	FMN	C9A-N10	2.37	1.42	1.38
2	A	401	FMN	C9A-C5A	2.39	1.47	1.42
2	A	401	FMN	C4A-C10	2.53	1.45	1.40

The worst 5 of 41 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	401	FMN	C4-C4A-C10	-6.46	115.81	119.94
2	D	401	FMN	C4-C4A-C10	-5.27	116.57	119.94
2	D	401	FMN	C4A-C10-N10	-4.73	117.09	120.52
2	C	401	FMN	O5'-P-O1P	-4.60	95.51	107.08
2	A	401	FMN	C4-C4A-C10	-4.37	117.14	119.94

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	401	FMN	1	0
2	B	401	FMN	4	0
3	B	402	DMS	1	0
2	C	401	FMN	3	0
2	D	401	FMN	2	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	377/382 (98%)	0.31	15 (3%) 42 44	11, 20, 36, 74	28 (7%)
1	B	377/382 (98%)	0.49	18 (4%) 34 36	12, 23, 41, 66	29 (7%)
1	C	377/382 (98%)	0.58	26 (6%) 20 21	15, 24, 43, 67	34 (9%)
1	D	376/382 (98%)	0.50	20 (5%) 30 32	13, 23, 43, 77	30 (7%)
All	All	1507/1528 (98%)	0.47	79 (5%) 31 33	11, 22, 42, 77	121 (8%)

The worst 5 of 79 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	2	ALA	11.2
1	D	68	TYR	7.3
1	D	121	LYS	6.2
1	C	152	GLY	6.0
1	A	2	ALA	5.9

### 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 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	401	31/31	0.96	0.14	-0.05	12,14,17,18	0
2	FMN	B	401	31/31	0.94	0.15	-0.09	17,20,22,23	0
2	FMN	D	401	31/31	0.94	0.13	-0.31	15,18,19,21	0
2	FMN	C	401	31/31	0.95	0.14	-0.47	15,19,21,23	0
3	DMS	B	402	4/4	0.95	0.16	-	36,38,40,41	0
3	DMS	C	402	4/4	0.90	0.19	-	39,43,45,50	0
3	DMS	A	402	4/4	0.95	0.15	-	34,36,41,47	0
3	DMS	D	402	4/4	0.96	0.18	-	44,47,48,48	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.