



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 10, 2017 – 06:50 PM EST

PDB ID : 5GSN
Title : Tmm in complex with methimazole
Authors : Zhang, Y.Z.; Li, C.Y.
Deposited on : 2016-08-16
Resolution : 2.20 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

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-20028442
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-20028442

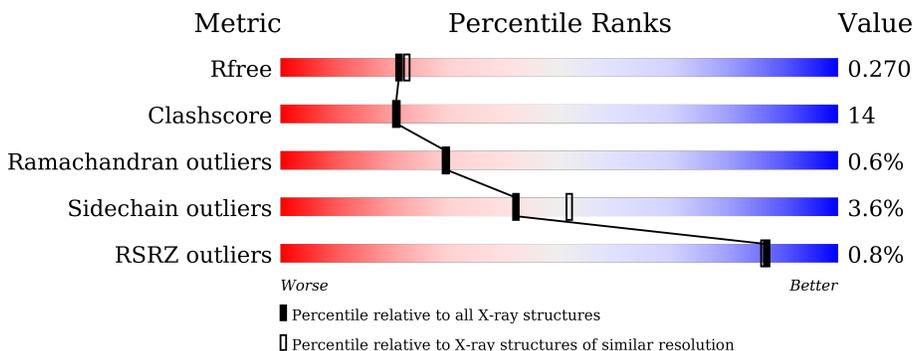
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.20 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	3774 (2.20-2.20)
Clashscore	102246	4477 (2.20-2.20)
Ramachandran outliers	100387	4404 (2.20-2.20)
Sidechain outliers	100360	4405 (2.20-2.20)
RSRZ outliers	91569	3781 (2.20-2.20)

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 ≥ 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	453	
1	B	453	
1	C	453	
1	D	453	

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	NAP	A	501	-	-	-	X
4	MMZ	A	503	-	-	-	X
4	MMZ	C	503	-	-	X	-

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 15548 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 Flavin-containing monooxygenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	445	Total 3582	C 2287	N 607	O 671	S 17	0	1	0
1	B	445	Total 3607	C 2303	N 613	O 674	S 17	0	4	0
1	C	445	Total 3573	C 2282	N 605	O 669	S 17	0	0	0
1	D	445	Total 3582	C 2287	N 606	O 672	S 17	0	1	0

There are 36 discrepancies between the modelled and reference sequences:

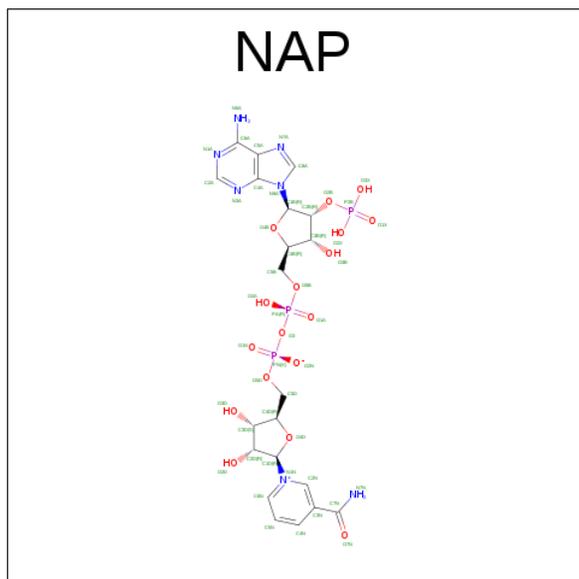
Chain	Residue	Modelled	Actual	Comment	Reference
A	153	ALA	GLU	engineered mutation	UNP A3SLM3
A	154	ALA	ASP	engineered mutation	UNP A3SLM3
A	207	SER	TYR	engineered mutation	UNP A3SLM3
A	448	HIS	-	expression tag	UNP A3SLM3
A	449	HIS	-	expression tag	UNP A3SLM3
A	450	HIS	-	expression tag	UNP A3SLM3
A	451	HIS	-	expression tag	UNP A3SLM3
A	452	HIS	-	expression tag	UNP A3SLM3
A	453	HIS	-	expression tag	UNP A3SLM3
B	153	ALA	GLU	engineered mutation	UNP A3SLM3
B	154	ALA	ASP	engineered mutation	UNP A3SLM3
B	207	SER	TYR	engineered mutation	UNP A3SLM3
B	448	HIS	-	expression tag	UNP A3SLM3
B	449	HIS	-	expression tag	UNP A3SLM3
B	450	HIS	-	expression tag	UNP A3SLM3
B	451	HIS	-	expression tag	UNP A3SLM3
B	452	HIS	-	expression tag	UNP A3SLM3
B	453	HIS	-	expression tag	UNP A3SLM3
C	153	ALA	GLU	engineered mutation	UNP A3SLM3
C	154	ALA	ASP	engineered mutation	UNP A3SLM3
C	207	SER	TYR	engineered mutation	UNP A3SLM3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	448	HIS	-	expression tag	UNP A3SLM3
C	449	HIS	-	expression tag	UNP A3SLM3
C	450	HIS	-	expression tag	UNP A3SLM3
C	451	HIS	-	expression tag	UNP A3SLM3
C	452	HIS	-	expression tag	UNP A3SLM3
C	453	HIS	-	expression tag	UNP A3SLM3
D	153	ALA	GLU	engineered mutation	UNP A3SLM3
D	154	ALA	ASP	engineered mutation	UNP A3SLM3
D	207	SER	TYR	engineered mutation	UNP A3SLM3
D	448	HIS	-	expression tag	UNP A3SLM3
D	449	HIS	-	expression tag	UNP A3SLM3
D	450	HIS	-	expression tag	UNP A3SLM3
D	451	HIS	-	expression tag	UNP A3SLM3
D	452	HIS	-	expression tag	UNP A3SLM3
D	453	HIS	-	expression tag	UNP A3SLM3

- Molecule 2 is NADP NICOTINAMIDE-ADENINE-DINUCLEOTIDE PHOSPHATE (three-letter code: NAP) (formula: C₂₁H₂₈N₇O₁₇P₃).



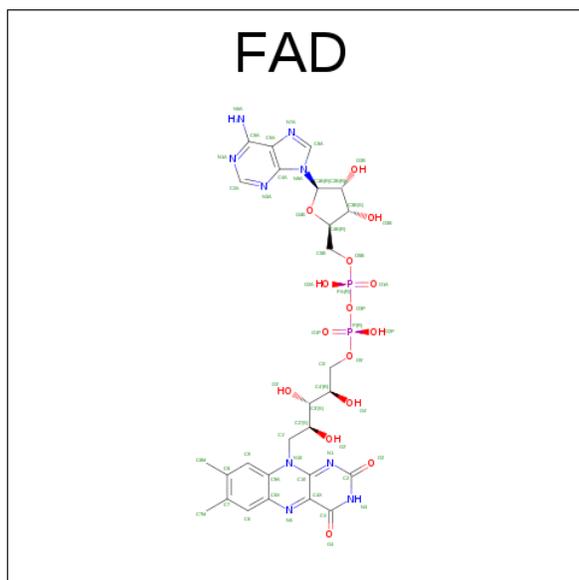
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
			Total	C	N	O			P
2	A	1	Total	C	N	O	P	0	0
			48	21	7	17	3		
2	B	1	Total	C	N	O	P	0	0
			48	21	7	17	3		
2	C	1	Total	C	N	O	P	0	0
			48	21	7	17	3		

Continued on next page...

Continued from previous page...

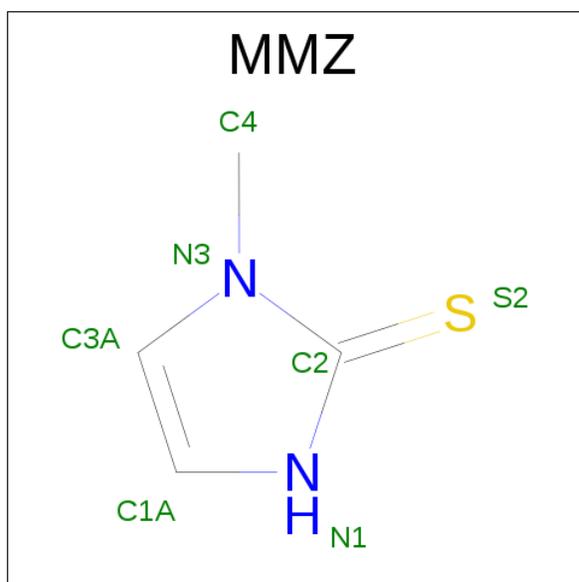
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
			Total	C	N	O			P
2	D	1	48	21	7	17	3	0	0

- Molecule 3 is FLAVIN-ADENINE DINUCLEOTIDE (three-letter code: FAD) (formula: $C_{27}H_{33}N_9O_{15}P_2$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
			Total	C	N	O			P
3	A	1	53	27	9	15	2	0	0
3	B	1	53	27	9	15	2	0	0
3	C	1	53	27	9	15	2	0	0
3	D	1	53	27	9	15	2	0	0

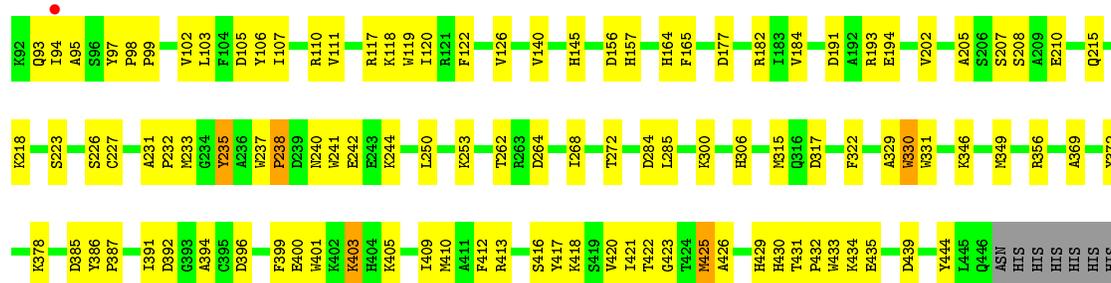
- Molecule 4 is 1-METHYL-1,3-DIHYDRO-2H-IMIDAZOLE-2-THIONE (three-letter code: MMZ) (formula: $C_4H_6N_2S$).



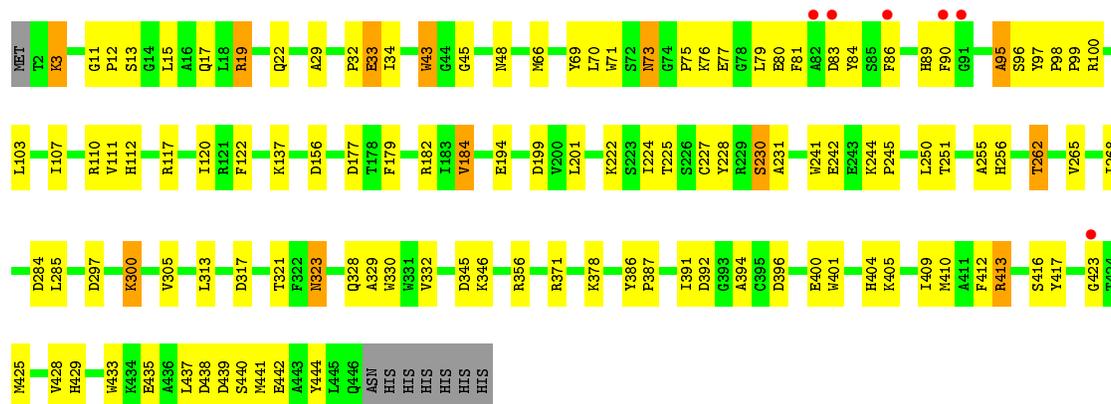
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	N	S	0	0
			7	4	2	1		
4	C	1	Total	C	N	S	0	0
			7	4	2	1		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	186	Total	O	0	0
			186	186		
5	B	202	Total	O	0	0
			202	202		
5	C	206	Total	O	0	0
			206	206		
5	D	192	Total	O	0	0
			192	192		



• Molecule 1: Flavin-containing monooxygenase



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	60.85Å 207.58Å 72.52Å 90.00° 90.31° 90.00°	Depositor
Resolution (Å)	46.74 - 2.20 46.74 - 2.20	Depositor EDS
% Data completeness (in resolution range)	89.1 (46.74-2.20) 91.4 (46.74-2.20)	Depositor EDS
R_{merge}	0.20	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.73 (at 2.20Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: 1.6.4_486)	Depositor
R, R_{free}	0.203 , 0.274 0.199 , 0.270	Depositor DCC
R_{free} test set	4239 reflections (5.05%)	DCC
Wilson B-factor (Å ²)	28.9	Xtrriage
Anisotropy	0.540	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.28 , 14.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.338 for h,-k,-l	Xtrriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	15548	wwPDB-VP
Average B, all atoms (Å ²)	33.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 33.25 % 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 8.2477e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹ Intensities estimated from amplitudes.

² 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: NAP, FAD, MMZ

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	0.42	0/3691	0.55	0/5010
1	B	0.44	0/3716	0.56	0/5044
1	C	0.44	1/3682 (0.0%)	0.57	1/4998 (0.0%)
1	D	0.42	0/3691	0.55	0/5010
All	All	0.43	1/14780 (0.0%)	0.56	1/20062 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	232	PRO	N-CD	5.04	1.54	1.47

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	231	ALA	C-N-CD	5.74	140.46	128.40

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	3582	0	3363	71	0
1	B	3607	0	3391	79	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	3573	0	3356	130	0
1	D	3582	0	3361	115	0
2	A	48	0	24	10	0
2	B	48	0	25	6	0
2	C	48	0	24	12	0
2	D	48	0	24	4	0
3	A	53	0	31	5	0
3	B	53	0	31	2	0
3	C	53	0	31	8	0
3	D	53	0	31	4	0
4	A	7	0	6	3	0
4	C	7	0	6	11	0
5	A	186	0	0	5	1
5	B	202	0	0	13	1
5	C	206	0	0	11	2
5	D	192	0	0	4	0
All	All	15548	0	13704	412	2

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 14.

The worst 5 of 412 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:323:ASN:H	1:D:323:ASN:HD22	1.01	1.01
1:A:87:GLU:OE2	1:A:93:GLN:NE2	1.98	0.97
1:D:323:ASN:HD22	1:D:323:ASN:N	1.64	0.95
1:C:421:ILE:HG13	1:C:422:THR:HG23	1.52	0.92
1:C:45:GLY:HA2	3:C:502:FAD:O3B	1.70	0.91

All (2) 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)
5:A:772:HOH:O	5:C:779:HOH:O[1_556]	1.94	0.26
5:B:794:HOH:O	5:C:803:HOH:O[1_455]	2.19	0.01

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	444/453 (98%)	417 (94%)	25 (6%)	2 (0%)	34	35
1	B	447/453 (99%)	423 (95%)	24 (5%)	0	100	100
1	C	443/453 (98%)	411 (93%)	27 (6%)	5 (1%)	17	14
1	D	444/453 (98%)	405 (91%)	35 (8%)	4 (1%)	21	19
All	All	1778/1812 (98%)	1656 (93%)	111 (6%)	11 (1%)	30	29

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	300	LYS
1	C	27	GLN
1	D	95	ALA
1	A	84	TYR
1	C	235	TYR

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	366/373 (98%)	352 (96%)	14 (4%)	40	49
1	B	369/373 (99%)	359 (97%)	10 (3%)	52	64
1	C	365/373 (98%)	353 (97%)	12 (3%)	45	56
1	D	366/373 (98%)	348 (95%)	18 (5%)	31	36
All	All	1466/1492 (98%)	1412 (96%)	54 (4%)	42	50

5 of 54 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	4	ARG
1	C	84	TYR
1	D	262	THR
1	C	19	ARG
1	C	50	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	323	ASN
1	C	415	ASN
1	D	256	HIS
1	C	256	HIS
1	D	323	ASN

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](#)

10 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	NAP	A	501	-	45,52,52	2.97	16 (35%)	55,80,80	2.06	8 (14%)
3	FAD	A	502	-	52,58,58	1.32	6 (11%)	52,89,89	2.04	9 (17%)
4	MMZ	A	503	-	5,7,7	1.68	1 (20%)	6,9,9	5.52	4 (66%)
2	NAP	B	501	-	45,52,52	2.88	17 (37%)	55,80,80	2.14	12 (21%)
3	FAD	B	502	-	52,58,58	1.31	5 (9%)	52,89,89	2.07	11 (21%)
2	NAP	C	501	-	45,52,52	2.90	16 (35%)	55,80,80	2.14	13 (23%)
3	FAD	C	502	-	52,58,58	1.31	5 (9%)	52,89,89	2.08	10 (19%)
4	MMZ	C	503	-	5,7,7	1.79	1 (20%)	6,9,9	5.02	5 (83%)
2	NAP	D	501	-	45,52,52	2.90	17 (37%)	55,80,80	1.93	10 (18%)
3	FAD	D	502	-	52,58,58	1.30	5 (9%)	52,89,89	2.09	9 (17%)

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	NAP	A	501	-	-	0/27/67/67	0/5/5/5
3	FAD	A	502	-	-	0/30/50/50	0/6/6/6
4	MMZ	A	503	-	-	0/0/0/0	0/1/1/1
2	NAP	B	501	-	-	0/27/67/67	0/5/5/5
3	FAD	B	502	-	-	0/30/50/50	0/6/6/6
2	NAP	C	501	-	-	0/27/67/67	0/5/5/5
3	FAD	C	502	-	-	0/30/50/50	0/6/6/6
4	MMZ	C	503	-	-	0/0/0/0	0/1/1/1
2	NAP	D	501	-	-	0/27/67/67	0/5/5/5
3	FAD	D	502	-	-	0/30/50/50	0/6/6/6

The worst 5 of 89 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	501	NAP	C3B-C2B	-6.94	1.37	1.53
2	B	501	NAP	C3B-C2B	-6.87	1.37	1.53
2	D	501	NAP	C3B-C2B	-6.64	1.38	1.53
2	C	501	NAP	C3B-C2B	-6.47	1.38	1.53
2	C	501	NAP	C2B-C1B	-3.07	1.44	1.53

The worst 5 of 91 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	501	NAP	N3A-C2A-N1A	-10.38	120.72	128.87
4	A	503	MMZ	C3A-N3-C2	-9.18	105.63	111.23
2	D	501	NAP	N3A-C2A-N1A	-8.95	121.84	128.87
2	C	501	NAP	N3A-C2A-N1A	-8.54	122.16	128.87
2	A	501	NAP	N3A-C2A-N1A	-8.30	122.35	128.87

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

10 monomers are involved in 51 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	501	NAP	10	0
3	A	502	FAD	5	0
4	A	503	MMZ	3	0
2	B	501	NAP	6	0
3	B	502	FAD	2	0
2	C	501	NAP	12	0
3	C	502	FAD	8	0
4	C	503	MMZ	11	0
2	D	501	NAP	4	0
3	D	502	FAD	4	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, 95th 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 ²)	Q < 0.9
1	A	445/453 (98%)	-0.51	2 (0%) 93 93	17, 30, 49, 66	0
1	B	445/453 (98%)	-0.51	3 (0%) 89 88	14, 30, 50, 65	0
1	C	445/453 (98%)	-0.36	4 (0%) 85 85	16, 32, 63, 86	0
1	D	445/453 (98%)	-0.34	6 (1%) 79 78	17, 31, 65, 86	0
All	All	1780/1812 (98%)	-0.43	15 (0%) 87 87	14, 30, 57, 86	0

The worst 5 of 15 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	82	ALA	6.1
1	C	91	GLY	5.4
1	D	86	PHE	3.4
1	D	423	GLY	3.4
1	C	82	ALA	3.1

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, 95th 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 ²)	Q<0.9
4	MMZ	A	503	7/7	0.84	0.17	4.43	35,41,46,59	0
2	NAP	A	501	48/48	0.93	0.14	2.04	25,34,53,61	0
2	NAP	C	501	48/48	0.94	0.13	1.36	25,35,53,56	0
3	FAD	A	502	53/53	0.95	0.12	0.49	13,22,36,37	0
2	NAP	B	501	48/48	0.95	0.11	0.21	21,32,44,49	0
3	FAD	B	502	53/53	0.96	0.11	0.16	14,22,30,37	0
2	NAP	D	501	48/48	0.95	0.12	0.13	23,33,46,52	0
3	FAD	C	502	53/53	0.96	0.11	0.10	14,24,40,41	0
3	FAD	D	502	53/53	0.95	0.11	-0.02	15,26,33,35	0
4	MMZ	C	503	7/7	0.84	0.19	-	37,41,48,64	0

6.5 Other polymers [i](#)

There are no such residues in this entry.