W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 14, 2017 – 08:51 pm GMT

PDB ID	:	3QFR
Title	:	Crystal Structure of Human NADPH-Cytochrome P450 Reductase (R457H Mutant)
Authors	:	Xia, C.; Marohnic, C.; Panda, S.P.; Masters, B.S.; Kim, J.-J.P.
Deposited on	:	2011-01-22
Resolution	:	2.40 Å(reported)

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

|		
MolProbity	:	4.02b-467
Mogul	:	1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
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)	:	recalc28949

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 2.40 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	Percentile Ranks	Value
Rfree		0.283
Clashscore		15
Ramachandran outliers		0.9%
Sidechain outliers		3.8%
RSRZ outliers		0.8%

Worse — *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution
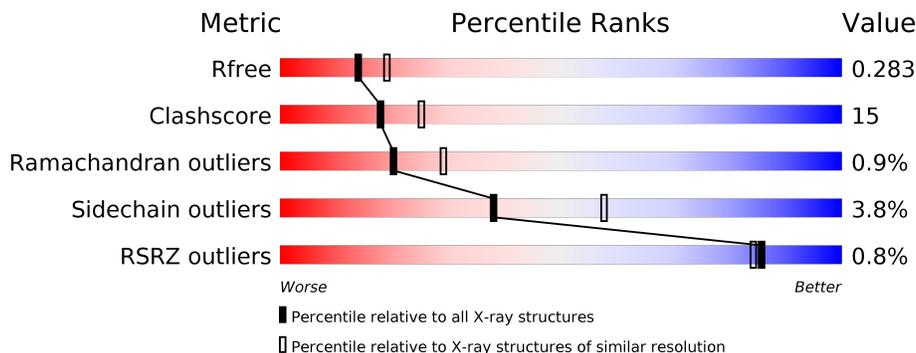
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	100719	3166 (2.40-2.40)
Clashscore	112137	3674 (2.40-2.40)
Ramachandran outliers	110173	3616 (2.40-2.40)
Sidechain outliers	110143	3617 (2.40-2.40)
RSRZ outliers	101464	3195 (2.40-2.40)

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 <=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	618	72% ... 24% ..
1	B	618	64% ... 31% ..

WORLDWIDE
wwPDB
PROTEIN DATA BANK

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 9964 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 NADPH–cytochrome P450 reductase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	604	Total 4813	C 3047	N 833	O 910	S 23	0	0	0
1	B	603	Total 4820	C 3049	N 835	O 913	S 23	0	0	0

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	63	GLY	-	EXPRESSION TAG	UNP P16435
A	64	SER	-	EXPRESSION TAG	UNP P16435
A	65	HIS	-	EXPRESSION TAG	UNP P16435
A	66	MET	-	EXPRESSION TAG	UNP P16435
A	457	HIS	ARG	ENGINEERED MUTATION	UNP P16435
B	63	GLY	-	EXPRESSION TAG	UNP P16435
B	64	SER	-	EXPRESSION TAG	UNP P16435
B	65	HIS	-	EXPRESSION TAG	UNP P16435
B	66	MET	-	EXPRESSION TAG	UNP P16435
B	457	HIS	ARG	ENGINEERED MUTATION	UNP P16435

- Molecule 2 is FLAVIN-ADENINE DINUCLEOTIDE (three-letter code: FAD) (formula: $C_{27}H_{33}N_9O_{15}P_2$).
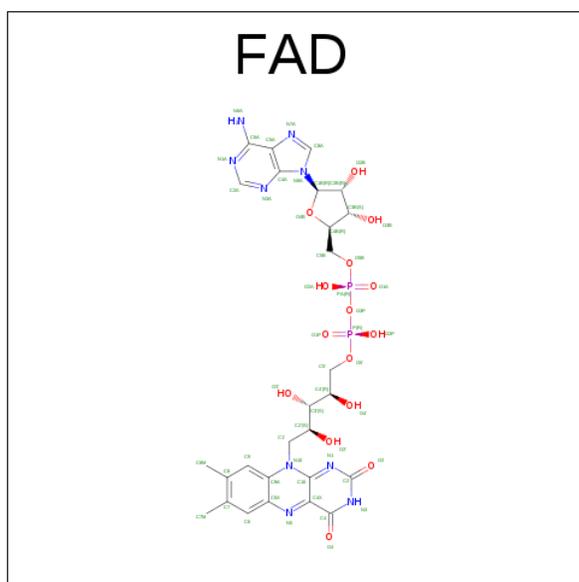
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			53	27	9	15	2		
2	B	1	Total	C	N	O	P	0	0
			53	27	9	15	2		

- Molecule 3 is FLAVIN MONONUCLEOTIDE (three-letter code: FMN) (formula: $C_{17}H_{21}N_4O_9P$).
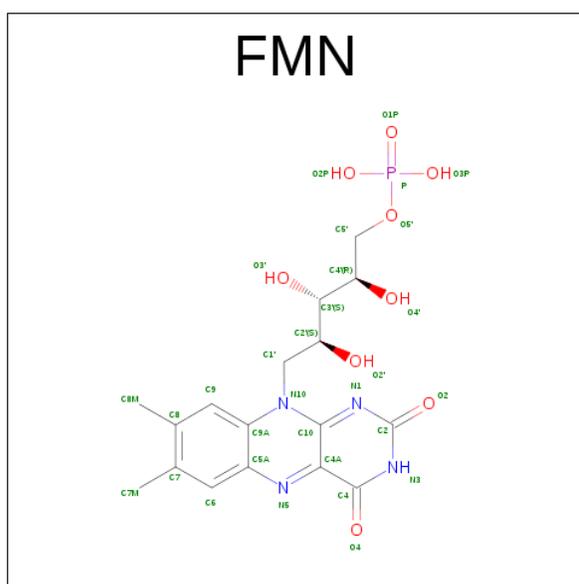
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			31	17	4	9	1		
3	B	1	Total	C	N	O	P	0	0
			31	17	4	9	1		

- Molecule 4 is NADP NICOTINAMIDE-ADENINE-DINUCLEOTIDE PHOSPHATE (three-letter code: NAP) (formula: $C_{21}H_{28}N_7O_{17}P_3$).
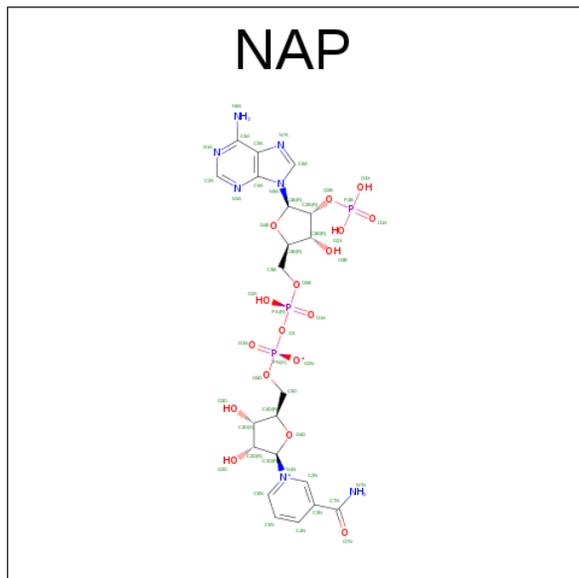
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
4	B	1	Total	C	N	O	P	0	0
			31	10	5	13	3		

- Molecule 5 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total	Ca	0	0
			1	1		

- Molecule 6 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	72	Total	O	0	0
			72	72		
6	B	28	Total	O	0	0
			28	28		

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 the various outlier classes 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: NADPH–cytochrome P450 reductase
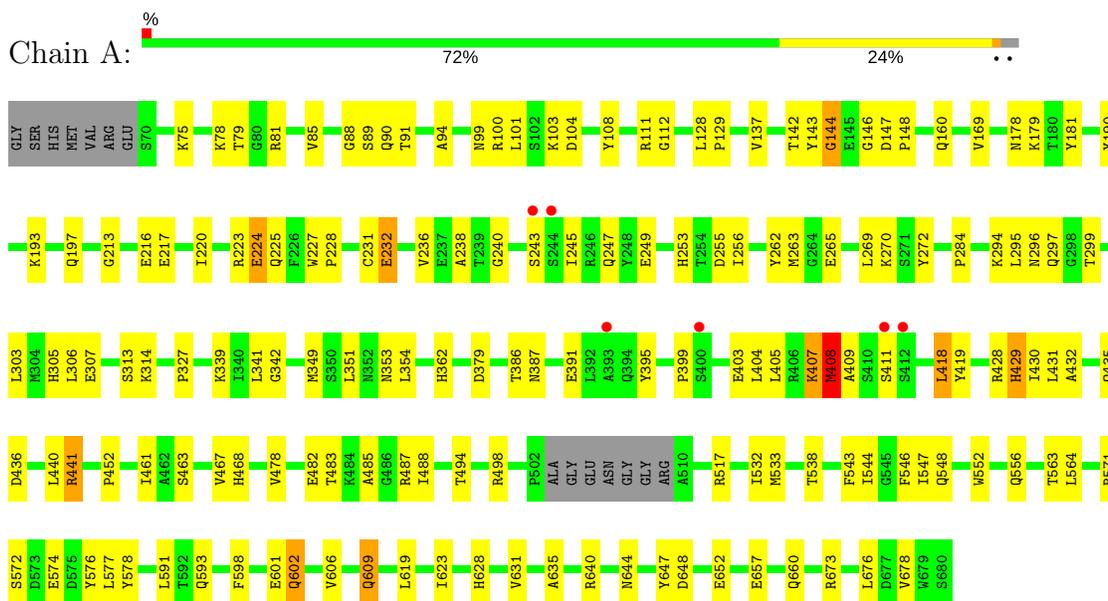
• Molecule 1: NADPH–cytochrome P450 reductase
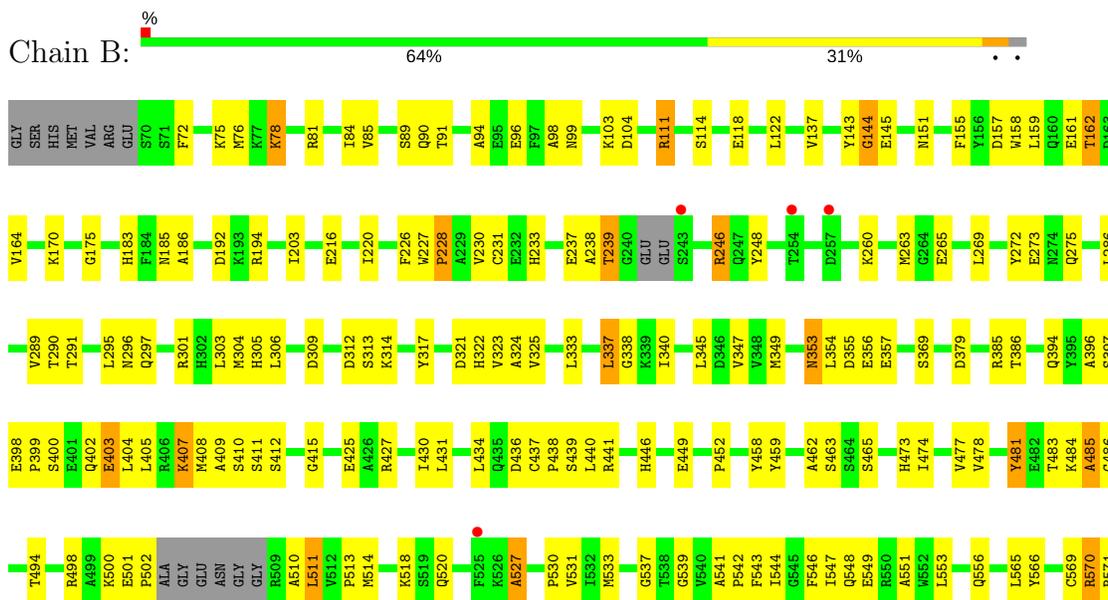
E574 D575 Y576 L577 Y578 E581 Q584 F585 H586 R587 A590 L591 T592 Q593 A597 F598 S599 R600 V606 Y607 V608 Q609 H610 L611 L612 K613 R616 E617 G626 A627 H628 I629 A635 R640 F646 L653 E657 Q660 R673 D677 V678 W679 S680

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	70.28Å 120.39Å 156.34Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	45.72 – 2.40 45.72 – 2.39	Depositor EDS
% Data completeness (in resolution range)	90.4 (45.72-2.40) 90.2 (45.72-2.39)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.07	Depositor
$< I/\sigma(I) >$ [1]	2.37 (at 2.39Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, R_{free}	0.227 , 0.285 0.227 , 0.283	Depositor DCC
R_{free} test set	2384 reflections (5.01%)	DCC
Wilson B-factor (Å^2)	49.5	Xtriage
Anisotropy	0.465	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 24.9	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	9964	wwPDB-VP
Average B, all atoms (Å^2)	59.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.11% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ 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, CA, NAP, FAD

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	0.41	0/4924	0.63	2/6667 (0.0%)
1	B	0.38	0/4930	0.62	0/6671
All	All	0.39	0/9854	0.62	2/13338 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	88	GLY	N-CA-C	-5.66	98.96	113.10
1	A	213	GLY	N-CA-C	-5.01	100.58	113.10

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	4813	0	4666	117	0
1	B	4820	0	4684	174	0
2	A	53	0	31	1	0
2	B	53	0	31	1	0
3	A	31	0	19	1	0
3	B	31	0	19	2	0
4	A	31	0	11	2	0

Continued on next page...

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	31	0	11	1	0
5	A	1	0	0	0	0
6	A	72	0	0	6	0
6	B	28	0	0	6	0
All	All	9964	0	9472	293	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 15.

The worst 5 of 293 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:441:ARG:HB3	1:A:441:ARG:HH11	1.20	1.05
1:B:275:GLN:HE22	1:B:286:LEU:H	1.12	0.97
1:A:657:GLU:H	1:A:660:GLN:HE21	1.11	0.93
1:B:379:ASP:HB3	1:B:452:PRO:HG2	1.52	0.89
1:A:99:ASN:HB3	6:A:21:HOH:O	1.74	0.88

There are no symmetry-related clashes.

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	600/618 (97%)	566 (94%)	32 (5%)	2 (0%)	44	60
1	B	597/618 (97%)	533 (89%)	55 (9%)	9 (2%)	12	16
All	All	1197/1236 (97%)	1099 (92%)	87 (7%)	11 (1%)	20	29

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	144	GLY
1	B	239	THR
1	B	356	GLU
1	A	408	MET
1	B	144	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	510/525 (97%)	492 (96%)	18 (4%)	41	61
1	B	514/525 (98%)	493 (96%)	21 (4%)	35	54
All	All	1024/1050 (98%)	985 (96%)	39 (4%)	38	58

5 of 39 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	609	GLN
1	B	203	ILE
1	B	591	LEU
1	B	78	LYS
1	B	111	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 48 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	660	GLN
1	B	185	ASN
1	B	586	HIS
1	B	90	GLN
1	B	153	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 ⓘ

Of 7 ligands modelled in this entry, 1 is monoatomic - leaving 6 for Mogul analysis.

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	$\#|Z| > 2$
3	FMN	A	751	-	31,33,33	2.58	12 (38%)	38,50,50	3.03	12 (31%)
2	FAD	A	752	-	51,58,58	2.84	22 (43%)	54,89,89	3.19	13 (24%)
4	NAP	A	753	-	28,33,52	2.53	8 (28%)	32,52,80	1.62	5 (15%)
3	FMN	B	751	-	31,33,33	2.49	11 (35%)	38,50,50	3.04	13 (34%)
2	FAD	B	752	-	51,58,58	2.86	23 (45%)	54,89,89	3.11	13 (24%)
4	NAP	B	753	-	28,33,52	2.34	6 (21%)	32,52,80	1.57	5 (15%)

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
3	FMN	A	751	-	-	0/16/18/18	0/3/3/3
2	FAD	A	752	-	-	0/28/50/50	0/6/6/6
4	NAP	A	753	-	-	0/17/37/67	0/3/3/5
3	FMN	B	751	-	-	0/16/18/18	0/3/3/3
2	FAD	B	752	-	-	0/28/50/50	0/6/6/6
4	NAP	B	753	-	-	0/17/37/67	0/3/3/5

The worst 5 of 82 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	752	FAD	C5'-C4'	-6.04	1.42	1.51
2	A	752	FAD	C5'-C4'	-5.40	1.43	1.51
3	A	751	FMN	C1'-N10	-3.42	1.44	1.48
3	B	751	FMN	C1'-N10	-2.94	1.45	1.48
4	B	753	NAP	PA-O5B	-2.67	1.47	1.59

The worst 5 of 61 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	752	FAD	N3A-C2A-N1A	-7.84	122.03	128.86
2	A	752	FAD	N3A-C2A-N1A	-7.80	122.06	128.86
2	A	752	FAD	C4X-C4-N3	-6.98	113.54	123.48
2	B	752	FAD	C4X-C4-N3	-6.84	113.74	123.48
3	A	751	FMN	C4-C4A-C10	-6.56	114.65	119.96

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	751	FMN	1	0
2	A	752	FAD	1	0
4	A	753	NAP	2	0
3	B	751	FMN	2	0
2	B	752	FAD	1	0
4	B	753	NAP	1	0

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, 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	604/618 (97%)	-0.32	6 (0%)	82	80	35, 52, 78, 92	0
1	B	603/618 (97%)	-0.14	4 (0%)	87	86	41, 65, 83, 90	0
All	All	1207/1236 (97%)	-0.23	10 (0%)	86	84	35, 59, 82, 92	0

The worst 5 of 10 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	244	SER	3.0
1	A	400	SER	2.6
1	A	411	SER	2.5
1	A	412	SER	2.5
1	B	243	SER	2.3

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, 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
2	FAD	B	752	53/53	0.93	0.19	0.62	58,73,82,83	0
3	FMN	B	751	31/31	0.95	0.17	0.57	61,67,73,74	0
4	NAP	A	753	31/48	0.92	0.15	0.27	81,85,107,111	0
3	FMN	A	751	31/31	0.98	0.16	0.25	45,50,57,59	0
2	FAD	A	752	53/53	0.97	0.14	-0.17	37,43,60,63	0
4	NAP	B	753	31/48	0.94	0.14	-0.37	69,73,93,97	0
5	CA	A	761	1/1	0.79	0.14	-	89,89,89,89	0

6.5 Other polymers ⓘ

There are no such residues in this entry.