Full wwPDB X-ray Structure Validation Report (i)

Oct 3, 2016 – 01:05 PM EDT

PDB ID	:	5JCI
Title	:	Structure and catalytic mechanism of monodehydroascorbate reductase, MDHAR, from Oryza sativa L. japonica
Authors	:	Park, A.K.; Kim, H.W.
Deposited on	:	2016-04-15
Resolution	:	1.70 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*
A user guide is available at
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

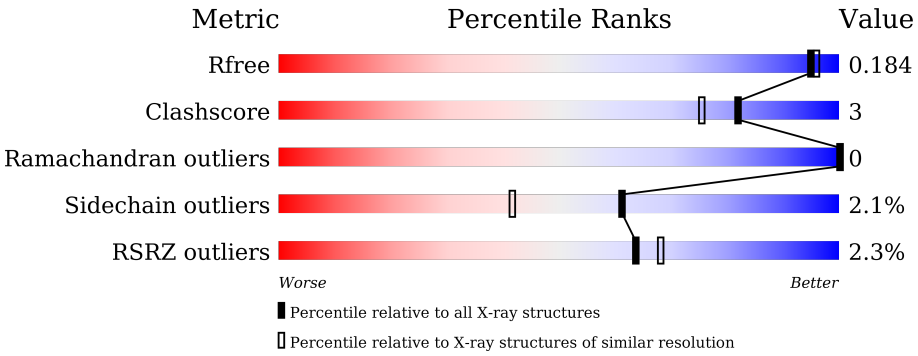
MolProbity	:	4.02b-467
Mogul	:	1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20027939
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20027939

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.70 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	3190 (1.70-1.70)
Clashscore	102246	3585 (1.70-1.70)
Ramachandran outliers	100387	3527 (1.70-1.70)
Sidechain outliers	100360	3527 (1.70-1.70)
RSRZ outliers	91569	3200 (1.70-1.70)

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	453	2% 80% 14% • 5%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 3812 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.

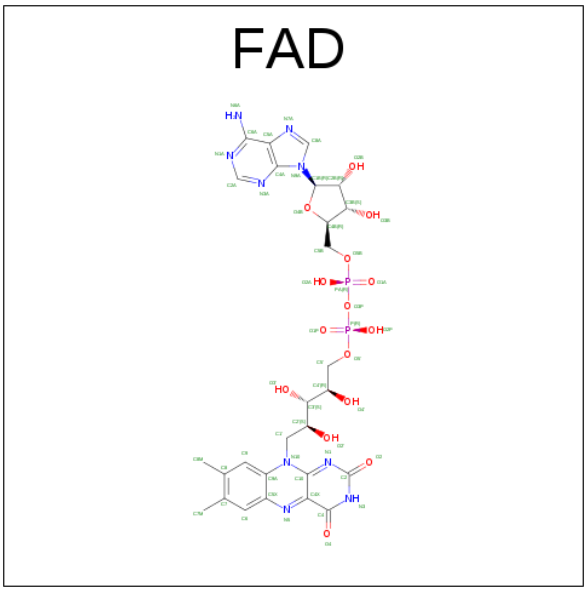
- Molecule 1 is a protein called Os09g0567300 protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	432	Total	C	N	O	S	0	0	0
			3278	2112	535	625	6			

There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-17	HIS	-	expression tag	UNP Q652L6
A	-16	HIS	-	expression tag	UNP Q652L6
A	-15	HIS	-	expression tag	UNP Q652L6
A	-14	HIS	-	expression tag	UNP Q652L6
A	-13	HIS	-	expression tag	UNP Q652L6
A	-12	HIS	-	expression tag	UNP Q652L6
A	-11	ALA	-	expression tag	UNP Q652L6
A	-10	SER	-	expression tag	UNP Q652L6
A	-9	GLU	-	expression tag	UNP Q652L6
A	-8	ASN	-	expression tag	UNP Q652L6
A	-7	LEU	-	expression tag	UNP Q652L6
A	-6	TYR	-	expression tag	UNP Q652L6
A	-5	PHE	-	expression tag	UNP Q652L6
A	-4	GLN	-	expression tag	UNP Q652L6
A	-3	GLY	-	expression tag	UNP Q652L6
A	-2	ALA	-	expression tag	UNP Q652L6
A	-1	MET	-	expression tag	UNP Q652L6
A	0	VAL	-	expression tag	UNP Q652L6
A	1	MET	-	expression tag	UNP Q652L6
A	2	ALA	-	expression tag	UNP Q652L6
A	3	SER	-	expression tag	UNP Q652L6

- 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		

- Molecule 3 is water.

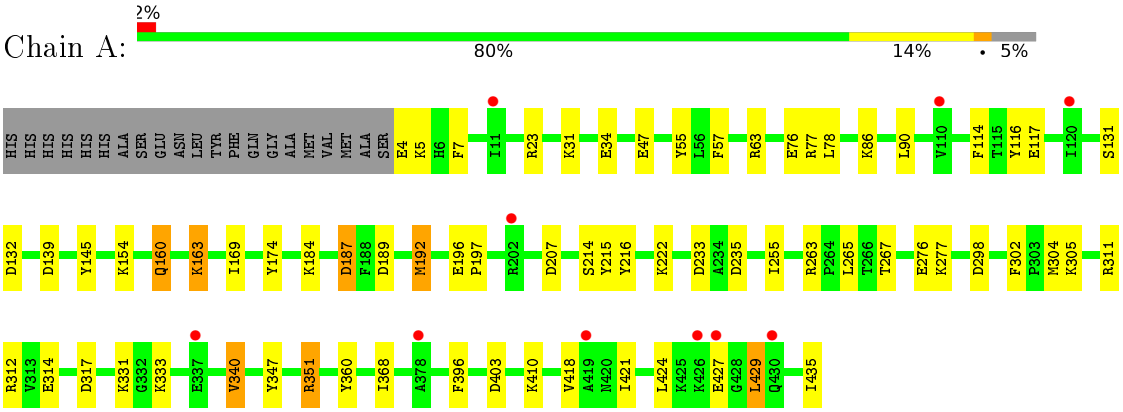
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	481	Total	O	0	0
			481	481		

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 (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: Os09g0567300 protein

Chain A:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α, β, γ	81.27Å 81.27Å 120.67Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 1.70 24.22 – 1.70	Depositor EDS
% Data completeness (in resolution range)	99.0 (50.00-1.70) 99.3 (24.22-1.70)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	5.96 (at 1.70Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.143 , 0.180 0.148 , 0.184	Depositor DCC
R_{free} test set	2139 reflections (5.01%)	DCC
Wilson B-factor (Å^2)	18.5	Xtriage
Anisotropy	0.022	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.41 , 50.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	3812	wwPDB-VP
Average B, all atoms (Å^2)	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.56% 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: 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	1.51	23/3348 (0.7%)	1.45	42/4531 (0.9%)

All (23) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	139	ASP	CB-CG	10.27	1.73	1.51
1	A	347	TYR	CE1-CZ	-7.98	1.28	1.38
1	A	174	TYR	CE1-CZ	7.92	1.48	1.38
1	A	117	GLU	CD-OE2	7.38	1.33	1.25
1	A	314	GLU	CD-OE2	7.11	1.33	1.25
1	A	76	GLU	CD-OE2	7.07	1.33	1.25
1	A	139	ASP	CG-OD2	6.97	1.41	1.25
1	A	139	ASP	CG-OD1	6.27	1.39	1.25
1	A	4	GLU	CD-OE2	6.21	1.32	1.25
1	A	4	GLU	CG-CD	6.11	1.61	1.51
1	A	47	GLU	CD-OE1	6.10	1.32	1.25
1	A	63	ARG	CZ-NH1	-6.06	1.25	1.33
1	A	216	TYR	CG-CD2	-5.66	1.31	1.39
1	A	214	SER	CB-OG	-5.43	1.35	1.42
1	A	55	TYR	CG-CD2	-5.32	1.32	1.39
1	A	184	LYS	CE-NZ	5.29	1.62	1.49
1	A	76	GLU	CD-OE1	5.27	1.31	1.25
1	A	160	GLN	CG-CD	5.24	1.63	1.51
1	A	174	TYR	CG-CD2	5.24	1.46	1.39
1	A	131	SER	N-CA	5.16	1.56	1.46
1	A	77	ARG	CZ-NH1	-5.13	1.26	1.33
1	A	427	GLU	CG-CD	5.08	1.59	1.51
1	A	403	ASP	CB-CG	5.03	1.62	1.51

All (42) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	351	ARG	NE-CZ-NH1	-14.70	112.95	120.30
1	A	312	ARG	NE-CZ-NH1	10.60	125.60	120.30
1	A	139	ASP	CB-CG-OD1	9.94	127.25	118.30
1	A	312	ARG	NE-CZ-NH2	-8.59	116.00	120.30
1	A	298	ASP	CB-CG-OD2	-8.55	110.61	118.30
1	A	340	VAL	CG1-CB-CG2	8.38	124.31	110.90
1	A	298	ASP	CB-CG-OD1	8.21	125.69	118.30
1	A	403	ASP	CB-CG-OD2	-7.87	111.22	118.30
1	A	187	ASP	CB-CG-OD1	-7.67	111.39	118.30
1	A	31	LYS	CD-CE-NZ	7.61	129.19	111.70
1	A	360	TYR	CD1-CE1-CZ	-7.52	113.03	119.80
1	A	263	ARG	NE-CZ-NH2	-6.87	116.87	120.30
1	A	117	GLU	OE1-CD-OE2	6.68	131.32	123.30
1	A	216	TYR	CD1-CE1-CZ	-6.63	113.83	119.80
1	A	63	ARG	NE-CZ-NH1	6.53	123.57	120.30
1	A	351	ARG	NH1-CZ-NH2	6.52	126.57	119.40
1	A	77	ARG	NE-CZ-NH1	6.25	123.42	120.30
1	A	23	ARG	NE-CZ-NH2	6.19	123.39	120.30
1	A	57	PHE	CB-CG-CD2	-6.19	116.47	120.80
1	A	235	ASP	CB-CG-OD2	-6.19	112.73	118.30
1	A	78	LEU	CB-CG-CD2	6.14	121.44	111.00
1	A	317	ASP	CB-CG-OD2	6.14	123.83	118.30
1	A	34	GLU	OE1-CD-OE2	-5.86	116.27	123.30
1	A	215	TYR	CB-CG-CD2	-5.82	117.51	121.00
1	A	132	ASP	CB-CG-OD2	5.80	123.52	118.30
1	A	192	MET	CG-SD-CE	-5.79	90.94	100.20
1	A	90	LEU	CA-CB-CG	5.74	128.51	115.30
1	A	340	VAL	CA-CB-CG1	5.62	119.33	110.90
1	A	55	TYR	CB-CG-CD2	5.59	124.36	121.00
1	A	145	TYR	CB-CG-CD2	-5.38	117.77	121.00
1	A	347	TYR	CD1-CE1-CZ	5.35	124.61	119.80
1	A	263	ARG	NE-CZ-NH1	5.34	122.97	120.30
1	A	311	ARG	NE-CZ-NH2	-5.33	117.64	120.30
1	A	429	LEU	CB-CG-CD1	5.30	120.01	111.00
1	A	302	PHE	CB-CG-CD2	5.21	124.45	120.80
1	A	233	ASP	CB-CG-OD2	-5.20	113.62	118.30
1	A	304	MET	CG-SD-CE	-5.19	91.90	100.20
1	A	207	ASP	CB-CG-OD2	-5.17	113.65	118.30
1	A	360	TYR	CB-CG-CD1	-5.14	117.92	121.00
1	A	189	ASP	CB-CG-OD1	5.13	122.91	118.30
1	A	396	PHE	CB-CG-CD2	-5.12	117.22	120.80
1	A	184	LYS	CD-CE-NZ	-5.09	99.99	111.70

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	3278	0	3280	21	0
2	A	53	0	31	0	0
3	A	481	0	0	7	3
All	All	3812	0	3311	21	3

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 3.

All (21) 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:351:ARG:NH1	3:A:601:HOH:O	2.07	0.86
1:A:368:ILE:HD11	1:A:424:LEU:HD23	1.70	0.74
1:A:410:LYS:HE3	1:A:435:ILE:O	1.97	0.64
1:A:163:LYS:HB2	3:A:926:HOH:O	2.03	0.58
1:A:160:GLN:HG2	3:A:668:HOH:O	2.08	0.52
1:A:7:PHE:O	1:A:116:TYR:HA	2.11	0.51
1:A:368:ILE:HD11	1:A:424:LEU:CD2	2.38	0.50
1:A:154:LYS:NZ	1:A:154:LYS:HB2	2.27	0.50
1:A:410:LYS:HE3	1:A:435:ILE:C	2.34	0.47
1:A:276:GLU:HG2	1:A:277:LYS:HG3	1.96	0.47
1:A:331:LYS:NZ	3:A:614:HOH:O	2.48	0.46
1:A:418:VAL:HG11	1:A:421:ILE:HD13	1.98	0.45
1:A:187:ASP:OD2	3:A:602:HOH:O	2.21	0.43
1:A:196:GLU:HB3	1:A:197:PRO:HD2	2.00	0.43
1:A:305:LYS:NZ	3:A:615:HOH:O	2.50	0.43
1:A:192:MET:HB3	1:A:192:MET:HE3	1.76	0.43
1:A:255:ILE:HD13	1:A:255:ILE:HG21	1.76	0.42
1:A:410:LYS:HG2	1:A:435:ILE:HG22	2.01	0.41
1:A:5:LYS:O	1:A:114:PHE:HA	2.21	0.41
1:A:333:LYS:NZ	3:A:608:HOH:O	2.40	0.41
1:A:265:LEU:HA	1:A:265:LEU:HD23	1.85	0.40

All (3) 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)
3:A:647:HOH:O	3:A:1015:HOH:O[7_555]	2.11	0.09
3:A:800:HOH:O	3:A:1040:HOH:O[4_454]	2.11	0.09
3:A:603:HOH:O	3:A:929:HOH:O[7_455]	2.12	0.08

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	430/453 (95%)	415 (96%)	15 (4%)	0	100 100

There are no Ramachandran outliers to report.

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	338/355 (95%)	331 (98%)	7 (2%)	61 42

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	86	LYS
1	A	163	LYS
1	A	169	ILE
1	A	222	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	267	THR
1	A	340	VAL
1	A	429	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	141	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)

1 ligand is 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	FAD	A	500	-	52,58,58	1.68	11 (21%)	52,89,89	2.47	14 (26%)

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	FAD	A	500	-	-	0/30/50/50	0/6/6/6

All (11) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	500	FAD	C4A-N3A	-4.47	1.29	1.35
2	A	500	FAD	C2B-C1B	-4.02	1.47	1.53
2	A	500	FAD	C10-N10	-3.55	1.35	1.39
2	A	500	FAD	C2'-C3'	-2.77	1.48	1.53
2	A	500	FAD	C4X-N5	2.07	1.36	1.33
2	A	500	FAD	O3B-C3B	2.20	1.48	1.43
2	A	500	FAD	C8A-N7A	2.54	1.39	1.34
2	A	500	FAD	C9A-C5X	2.63	1.48	1.42
2	A	500	FAD	C1'-N10	2.84	1.51	1.48
2	A	500	FAD	C4X-C10	3.26	1.46	1.40
2	A	500	FAD	C2A-N3A	3.77	1.38	1.32

All (14) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	500	FAD	N3A-C2A-N1A	-9.52	121.39	128.87
2	A	500	FAD	C5X-C9A-N10	-5.46	113.48	117.58
2	A	500	FAD	C4X-C4-N3	-3.98	118.32	123.52
2	A	500	FAD	N3-C2-N1	-3.30	122.13	127.69
2	A	500	FAD	C4-C4X-C10	-2.37	118.42	119.94
2	A	500	FAD	C7M-C7-C8	-2.24	115.91	120.73
2	A	500	FAD	O3B-C3B-C4B	-2.13	104.66	111.01
2	A	500	FAD	O5B-C5B-C4B	-2.04	101.72	109.09
2	A	500	FAD	O4'-C4'-C5'	-2.04	105.65	110.09
2	A	500	FAD	C2A-N1A-C6A	2.55	123.33	118.77
2	A	500	FAD	C4X-N5-C5X	3.61	120.98	116.72
2	A	500	FAD	C4-C4X-N5	3.82	123.35	118.70
2	A	500	FAD	N6A-C6A-N1A	3.87	125.01	118.52
2	A	500	FAD	C4-N3-C2	7.63	121.52	115.16

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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^{th} 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	432/453 (95%)	-0.22	10 (2%) 64 68	12, 19, 37, 64	0

All (10) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	419	ALA	5.3
1	A	427	GLU	3.4
1	A	426	LYS	3.2
1	A	430	GLN	2.7
1	A	11	ILE	2.6
1	A	202	ARG	2.4
1	A	378	ALA	2.4
1	A	337	GLU	2.4
1	A	120	ILE	2.2
1	A	110	VAL	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^{th} 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	A	500	53/53	0.99	0.05	-1.18	11,13,16,18	0

6.5 Other polymers (i)

There are no such residues in this entry.