

# Full wwPDB X-ray Structure Validation Report (i)

#### Dec 9, 2024 – 06:19 pm GMT

PDB ID	:	8RDG
Title	:	Crystal structure of UDP-galactose 4-epimerase from Pyrococcus horikoshii
		with bound NAD
Authors	:	Thunnissen, A.M.W.H.; Alvarez Quispe, C.; Desmet, T.
Deposited on		
Resolution	:	1.90  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 https://www.wwpdb.org/validation/2017/XrayValidationReportHelp with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at http://www.wwpdb.org/validation/2017/FAQs#types.

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

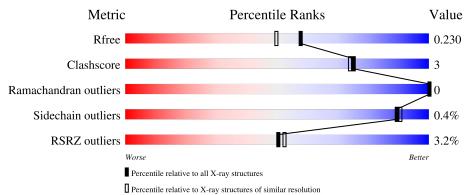
MolProbity	:	4.02b-467
5		1.8.4, CSD as541be (2020)
Xtriage (Phenix)		
EDS		3.0
buster-report	:	1.1.7 (2018)
		20231227.v01 (using entries in the PDB archive December 27th 2023)
		9.0.003 (Gargrove)
Density-Fitness		
Ideal geometry (proteins)		
Ideal geometry (DNA, RNA)		<u> </u>
Validation Pipeline (wwPDB-VP)		2.40

# 1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:  $X\text{-}RAY \, DIFFRACTION$ 

The reported resolution of this entry is 1.90 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	$\begin{array}{c} \textbf{Whole archive} \\ (\#\textbf{Entries}) \end{array}$	${f Similar\ resolution}\ (\#{ m Entries,\ resolution\ range}({ m \AA}))$		
$R_{free}$	164625	7293 (1.90-1.90)		
Clashscore	180529	8090 (1.90-1.90)		
Ramachandran outliers	177936	8022 (1.90-1.90)		
Sidechain outliers	177891	8022 (1.90-1.90)		
RSRZ outliers	164620	7292 (1.90-1.90)		

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 of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. 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	А	326	% <b>8</b> 9%	8%	•
1	В	326	5% 89%	8%	•



# 2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 5317 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.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	٨	315	Total	С	Ν	0	S	0	2	0
	1 A	515	2503	1595	429	471	8	0	2	0
1	D	315	Total	С	Ν	0	S	0	0	0
	D	515	2484	1585	424	467	8	0	0	0

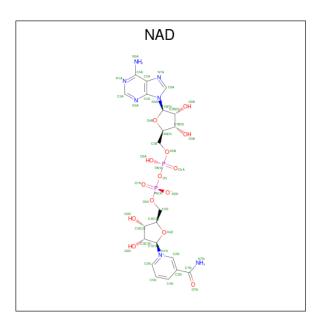
• Molecule 1 is a protein called SDR family NAD(P)-dependent oxidoreductase.

Chain	Residue	Modelled	Actual	Comment	Reference
А	319	LEU	-	expression tag	UNP A0A832T930
А	320	GLU	-	expression tag	UNP A0A832T930
А	321	HIS	-	expression tag	UNP A0A832T930
А	322	HIS	-	expression tag	UNP A0A832T930
А	323	HIS	-	expression tag	UNP A0A832T930
А	324	HIS	-	expression tag	UNP A0A832T930
А	325	HIS	-	expression tag	UNP A0A832T930
А	326	HIS	-	expression tag	UNP A0A832T930
В	319	LEU	-	expression tag	UNP A0A832T930
В	320	GLU	-	expression tag	UNP A0A832T930
В	321	HIS	-	expression tag	UNP A0A832T930
В	322	HIS	-	expression tag	UNP A0A832T930
В	323	HIS	-	expression tag	UNP A0A832T930
В	324	HIS	-	expression tag	UNP A0A832T930
В	325	HIS	-	expression tag	UNP A0A832T930
В	326	HIS	-	expression tag	UNP A0A832T930

There are 16 discrepancies between the modelled and reference sequences:

• Molecule 2 is NICOTINAMIDE-ADENINE-DINUCLEOTIDE (three-letter code: NAD) (formula: C<sub>21</sub>H<sub>27</sub>N<sub>7</sub>O<sub>14</sub>P<sub>2</sub>) (labeled as "Ligand of Interest" by depositor).





Mol	Chain	Residues	Atoms			ZeroOcc	AltConf		
0	٨	1	Total	С	Ν	Ο	Р	0	0
	Z A	1	44	21	7	14	2	0	0
0	В	1	Total	С	Ν	Ο	Р	0	0
	D	1	44	21	7	14	2	0	

• Molecule 3 is water.

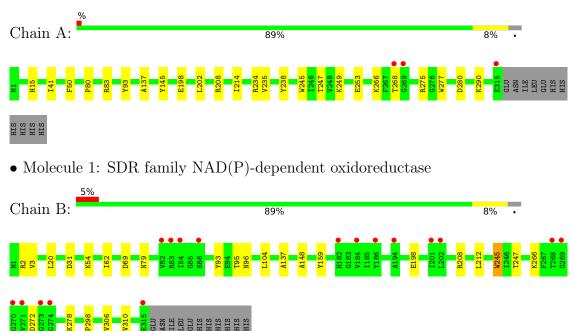
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	А	130	Total O 130 130	0	0
3	В	112	Total         O           112         112	0	0



# 3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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: SDR family NAD(P)-dependent oxidoreductase





# 4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants	61.38Å 78.09Å 128.28Å	Depositor
a, b, c, $\alpha$ , $\beta$ , $\gamma$	$90.00^{\circ}$ $90.00^{\circ}$ $90.00^{\circ}$	Depositor
Resolution (Å)	49.56 - 1.90	Depositor
Resolution (A)	49.56 - 1.90	EDS
% Data completeness	99.0 (49.56-1.90)	Depositor
(in resolution range)	99.6 (49.56 - 1.90)	EDS
R <sub>merge</sub>	0.21	Depositor
R <sub>sym</sub>	(Not available)	Depositor
$< I/\sigma(I) > 1$	$1.32 (at 1.90 \text{\AA})$	Xtriage
Refinement program	PHENIX (1.21rc1_4958: ???)	Depositor
D D.	0.193 , $0.230$	Depositor
$R, R_{free}$	0.193 , $0.230$	DCC
$R_{free}$ test set	2461 reflections $(4.98%)$	wwPDB-VP
Wilson B-factor $(Å^2)$	22.3	Xtriage
Anisotropy	0.256	Xtriage
Bulk solvent $k_{sol}(e/Å^3), B_{sol}(Å^2)$	0.36, 31.0	EDS
L-test for twinning <sup>2</sup>	$ \langle L  \rangle = 0.49, \langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	5317	wwPDB-VP
Average B, all atoms $(Å^2)$	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 6.73% of the height of the origin peak. No significant pseudotranslation is detected.

<sup>&</sup>lt;sup>2</sup>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.



<sup>&</sup>lt;sup>1</sup>Intensities estimated from amplitudes.

# 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: NAD

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		lengths	Bond angles		
Mol Chain		RMSZ	# Z  > 5	RMSZ	# Z  > 5	
1	А	0.38	0/2550	0.60	0/3443	
1	В	0.39	0/2531	0.59	0/3418	
All	All	0.39	0/5081	0.60	0/6861	

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 (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	А	2503	0	2509	14	0
1	В	2484	0	2494	17	0
2	А	44	0	26	0	0
2	В	44	0	26	0	0
3	А	130	0	0	1	0
3	В	112	0	0	0	0
All	All	5317	0	5055	29	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 3.

All (29) close contacts within the same asymmetric unit are listed below, sorted by their clash



magnitude.

Atom-1	Atom-2	Interatomic	Clash
Atom-1	Atom-2	distance (Å)	overlap (Å)
1:A:198:GLU:HB3	1:A:266:LYS:HE2	1.78	0.65
1:B:208:ARG:HD2	1:B:245:TRP:CE2	2.38	0.58
1:B:208:ARG:HG3	1:B:247:THR:HG22	1.87	0.55
1:A:208:ARG:HG2	1:A:247:THR:HG22	1.89	0.55
1:A:137:ALA:HB2	1:B:137:ALA:HB2	1.88	0.54
1:A:83:ARG:HG3	3:A:602:HOH:O	2.07	0.54
1:B:272:ASP:OD2	1:B:278:LYS:HD2	2.10	0.52
1:A:249:LYS:NZ	1:A:253:GLU:OE2	2.44	0.49
1:A:202:LEU:HD23	1:A:268:THR:HG21	1.94	0.48
1:A:83:ARG:HG2	1:A:277:TRP:CD2	2.49	0.47
1:B:96:ASN:HB2	1:B:148:ALA:HB1	1.97	0.46
1:A:238:TYR:CE2	1:A:290:LYS:HG2	2.51	0.46
1:B:198:GLU:HB3	1:B:266:LYS:HE2	1.97	0.45
1:B:62:ILE:HD12	1:B:62:ILE:HA	1.82	0.45
1:B:79:ASN:HB3	1:B:95:THR:OG1	2.16	0.45
1:B:3:VAL:HG21	1:B:20:LEU:HD13	1.99	0.45
1:B:31:ASP:O	1:B:54:LYS:HA	2.18	0.44
1:B:208:ARG:HG3	1:B:247:THR:CG2	2.48	0.44
1:B:212:LEU:HD11	1:B:298:PRO:HB3	2.00	0.44
1:A:234:ARG:HG3	1:A:235:VAL:HG23	2.00	0.44
1:B:208:ARG:HD2	1:B:245:TRP:CD2	2.54	0.43
1:A:275:ARG:HB3	1:A:280:ASP:HB3	2.00	0.43
1:A:80:PRO:HA	1:A:145:TYR:CE1	2.54	0.42
1:A:15:HIS:CD2	1:A:214:ILE:HG12	2.55	0.42
1:B:104:LEU:HB3	1:B:159:TYR:CD2	2.55	0.41
1:B:306:VAL:O	1:B:310:VAL:HG23	2.21	0.41
1:A:41:ILE:HD12	1:A:50:PHE:HZ	1.86	0.41
1:A:93:TYR:HH	1:B:93:TYR:HH	1.66	0.41
1:B:2:ARG:NE	1:B:69:ASP:O	2.49	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.



Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	А	315/326~(97%)	308~(98%)	7~(2%)	0	100	100
1	В	313/326~(96%)	306 (98%)	7 (2%)	0	100	100
All	All	628/652~(96%)	614 (98%)	14 (2%)	0	100	100

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

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	А	268/277~(97%)	267~(100%)	1 (0%)	89 90		
1	В	266/277~(96%)	265 (100%)	1 (0%)	89 90		
All	All	534/554~(96%)	532 (100%)	2~(0%)	89 90		

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	А	245	TRP
1	В	245	TRP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

#### 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 oligosaccharides 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	Trune	Chain	Dec	Link	Bo	ond leng	$\mathbf{ths}$	В	ond ang	les
Mol Type Ch	Chain	nain Res		Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z >2	
2	NAD	А	401	-	42,48,48	0.76	1 (2%)	50,73,73	0.98	3 (6%)
2	NAD	В	401	-	42,48,48	0.88	2 (4%)	50,73,73	0.89	1 (2%)

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	NAD	А	401	-	-	5/26/62/62	0/5/5/5
2	NAD	В	401	-	-	6/26/62/62	0/5/5/5

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Ζ	$\operatorname{Observed}(\operatorname{\AA})$	$\mathrm{Ideal}(\mathrm{\AA})$
2	В	401	NAD	C8A-N7A	-2.91	1.29	1.34
2	В	401	NAD	O4D-C1D	-2.16	1.38	1.41
2	А	401	NAD	C4A-N3A	-2.12	1.32	1.35

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
2	А	401	NAD	O4B-C1B-C2B	-2.60	103.13	106.93
2	А	401	NAD	O7N-C7N-C3N	-2.46	116.69	119.63
2	В	401	NAD	O2N-PN-O1N	2.28	123.49	112.24

Continued on next page...



Continued from previous page...

Mol	Chain	Res	Type	Atoms	$\mathbf{Z}$	$Observed(^{o})$	$Ideal(^{o})$
2	А	401	NAD	O2A-PA-O1A	2.02	122.25	112.24

There are no chirality outliers.

All (11) torsion outliers are listed below:

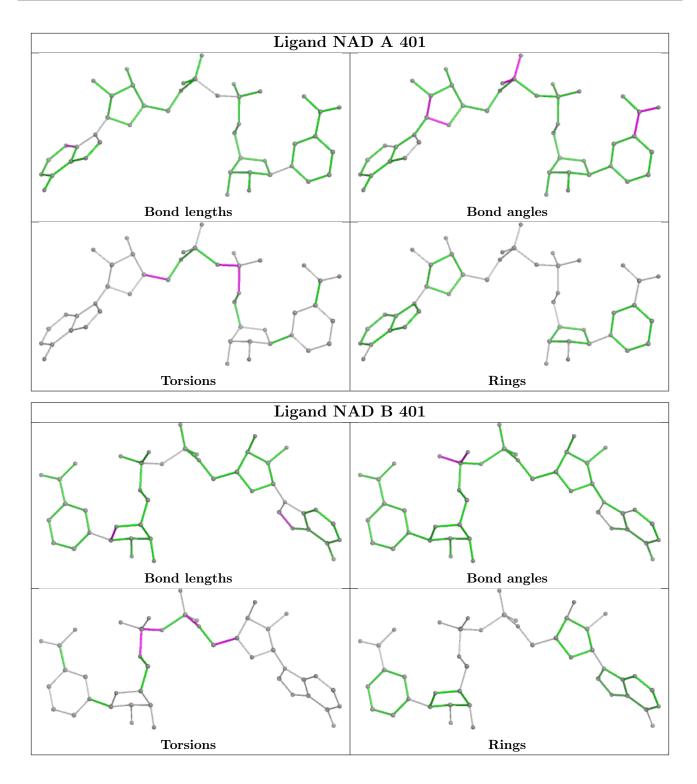
Mol	Chain	Res	Type	Atoms
2	А	401	NAD	C5D-O5D-PN-O1N
2	А	401	NAD	C5D-O5D-PN-O2N
2	В	401	NAD	C5B-O5B-PA-O1A
2	В	401	NAD	C5D-O5D-PN-O2N
2	В	401	NAD	C5D-O5D-PN-O3
2	В	401	NAD	C5D-O5D-PN-O1N
2	В	401	NAD	PA-O3-PN-O2N
2	А	401	NAD	C5D-O5D-PN-O3
2	А	401	NAD	O4B-C4B-C5B-O5B
2	В	401	NAD	O4B-C4B-C5B-O5B
2	А	401	NAD	PA-O3-PN-O2N

There are no ring outliers.

No monomer is involved in short contacts.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.





## 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,  $95^{th}$  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	$\langle RSRZ \rangle$	# RSRZ > 2	$OWAB(Å^2)$	Q<0.9
1	А	315/326~(96%)	-0.25	3 (0%) 79 81	10, 22, 41, 65	2 (0%)
1	В	315/326~(96%)	0.07	17 (5%) 32 34	12, 24, 51, 72	0
All	All	630/652~(96%)	-0.09	20 (3%) 50 52	10, 23, 47, 72	2 (0%)

All (20) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	А	268	THR	3.3
1	А	315	GLU	2.9
1	В	274	GLY	2.7
1	В	84	ILE	2.6
1	В	270	GLY	2.6
1	В	182	HIS	2.5
1	В	194	ALA	2.5
1	А	269	GLY	2.4
1	В	86	SER	2.4
1	В	184	VAL	2.3
1	В	202	LEU	2.3
1	В	315	GLU	2.3
1	В	271	VAL	2.3
1	В	268	THR	2.3
1	В	273	GLY	2.2
1	В	83	ARG	2.1
1	В	186	TYR	2.1
1	В	201	ILE	2.1
1	В	269	GLY	2.0
1	В	82	VAL	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 monosaccharides in this entry.

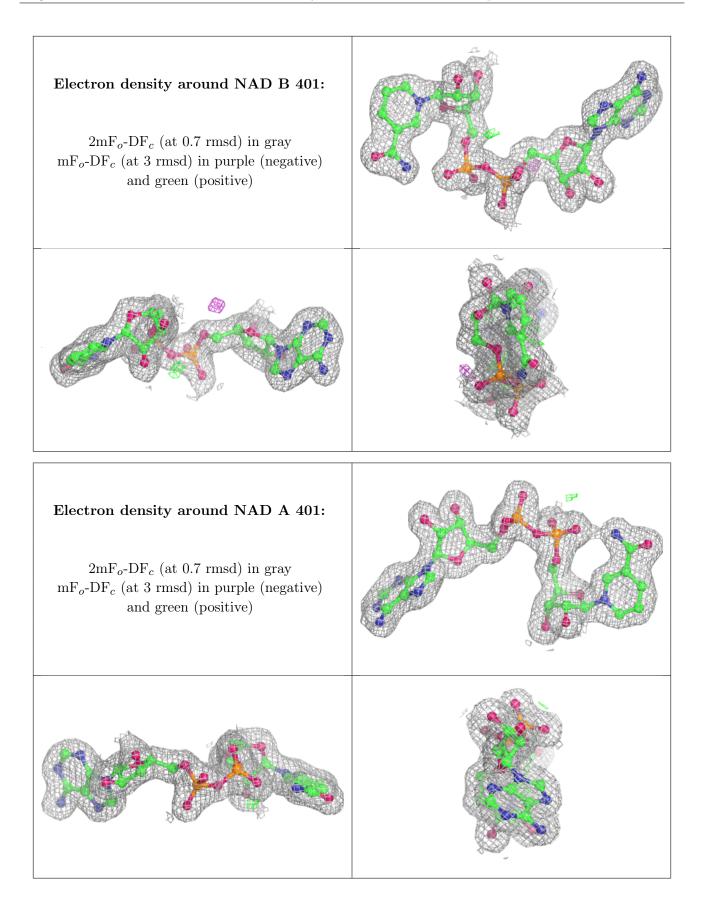
### 6.4 Ligands (i)

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. 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	$\mathbf{B} ext{-factors}(\mathbf{A}^2)$	Q<0.9
2	NAD	В	401	44/44	0.98	0.04	12,17,21,22	0
2	NAD	А	401	44/44	0.99	0.04	12,16,18,22	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.







## 6.5 Other polymers (i)

There are no such residues in this entry.

