



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 21, 2020 – 07:25 PM EST

PDB ID : 6T2B  
Title : Glycoside hydrolase family 109 from Akkermansia muciniphila in complex with GalNAc and NAD<sup>+</sup>.  
Authors : Chaberski, E.K.; Fredslund, F.; Teze, D.; Shuoker, B.; Kunstmann, S.; Karlsson, E.N.; Hachem, M.A.; Welner, D.H.  
Deposited on : 2019-10-08  
Resolution : 2.13 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](mailto: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 ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.0 (224370), CSD as540be (2019)  
Xtriage (Phenix) : 1.13  
EDS : 2.8  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.8

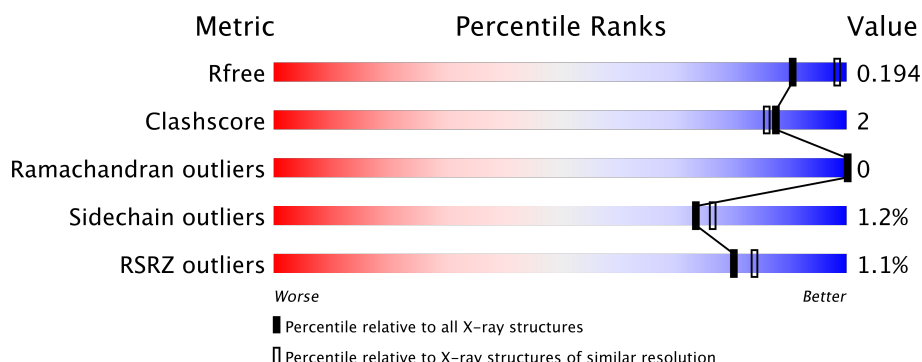
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.13 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))
$R_{free}$	111664	2128 (2.16-2.12)
Clashscore	122126	2253 (2.16-2.12)
Ramachandran outliers	120053	2223 (2.16-2.12)
Sidechain outliers	120020	2222 (2.16-2.12)
RSRZ outliers	108989	2086 (2.16-2.12)

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 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  $\leq 5\%$ . 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	456	<div> <div>3%</div> <div> <div></div> <div>89%</div> <div>7%</div> <div></div> </div> <div></div> </div>
1	B	456	<div> <div>%</div> <div> <div></div> <div>90%</div> <div>6%</div> <div></div> </div> <div></div> </div>
1	C	456	<div> <div></div> <div> <div></div> <div>90%</div> <div>5%</div> <div></div> </div> <div></div> </div>
1	D	456	<div> <div>%</div> <div> <div></div> <div>89%</div> <div>7%</div> <div></div> </div> <div></div> </div>

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 15130 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 Glycosyl hydrolase family 109 protein 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	438	Total	C	N	O	S	0	3	0
			3413	2149	612	622	30			
1	B	439	Total	C	N	O	S	0	2	0
			3413	2148	613	623	29			
1	C	436	Total	C	N	O	S	0	2	0
			3394	2137	610	618	29			
1	D	437	Total	C	N	O	S	0	2	0
			3402	2141	612	620	29			

There are 36 discrepancies between the modelled and reference sequences:

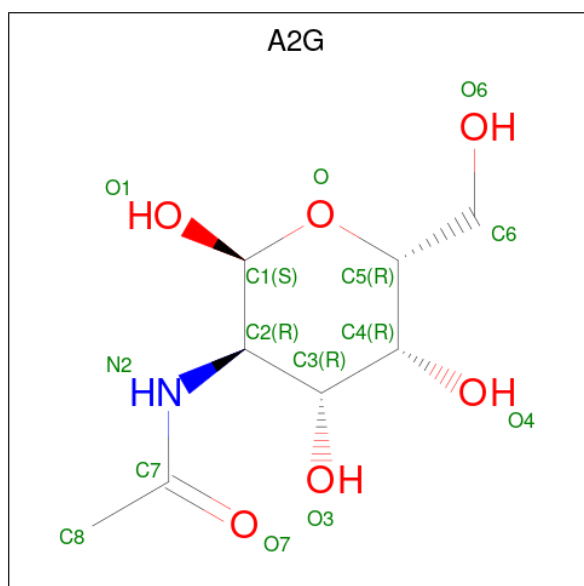
Chain	Residue	Modelled	Actual	Comment	Reference
A	26	MET	-	initiating methionine	UNP A0A2N8I799
A	474	LEU	-	expression tag	UNP A0A2N8I799
A	475	GLU	-	expression tag	UNP A0A2N8I799
A	476	HIS	-	expression tag	UNP A0A2N8I799
A	477	HIS	-	expression tag	UNP A0A2N8I799
A	478	HIS	-	expression tag	UNP A0A2N8I799
A	479	HIS	-	expression tag	UNP A0A2N8I799
A	480	HIS	-	expression tag	UNP A0A2N8I799
A	481	HIS	-	expression tag	UNP A0A2N8I799
B	26	MET	-	initiating methionine	UNP A0A2N8I799
B	474	LEU	-	expression tag	UNP A0A2N8I799
B	475	GLU	-	expression tag	UNP A0A2N8I799
B	476	HIS	-	expression tag	UNP A0A2N8I799
B	477	HIS	-	expression tag	UNP A0A2N8I799
B	478	HIS	-	expression tag	UNP A0A2N8I799
B	479	HIS	-	expression tag	UNP A0A2N8I799
B	480	HIS	-	expression tag	UNP A0A2N8I799
B	481	HIS	-	expression tag	UNP A0A2N8I799
C	26	MET	-	initiating methionine	UNP A0A2N8I799
C	474	LEU	-	expression tag	UNP A0A2N8I799
C	475	GLU	-	expression tag	UNP A0A2N8I799

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	476	HIS	-	expression tag	UNP A0A2N8I799
C	477	HIS	-	expression tag	UNP A0A2N8I799
C	478	HIS	-	expression tag	UNP A0A2N8I799
C	479	HIS	-	expression tag	UNP A0A2N8I799
C	480	HIS	-	expression tag	UNP A0A2N8I799
C	481	HIS	-	expression tag	UNP A0A2N8I799
D	26	MET	-	initiating methionine	UNP A0A2N8I799
D	474	LEU	-	expression tag	UNP A0A2N8I799
D	475	GLU	-	expression tag	UNP A0A2N8I799
D	476	HIS	-	expression tag	UNP A0A2N8I799
D	477	HIS	-	expression tag	UNP A0A2N8I799
D	478	HIS	-	expression tag	UNP A0A2N8I799
D	479	HIS	-	expression tag	UNP A0A2N8I799
D	480	HIS	-	expression tag	UNP A0A2N8I799
D	481	HIS	-	expression tag	UNP A0A2N8I799

- Molecule 2 is N-ACETYL-2-DEOXY-2-AMINO-GALACTOSE (three-letter code: A2G) (formula:  $C_8H_{15}NO_6$ ).



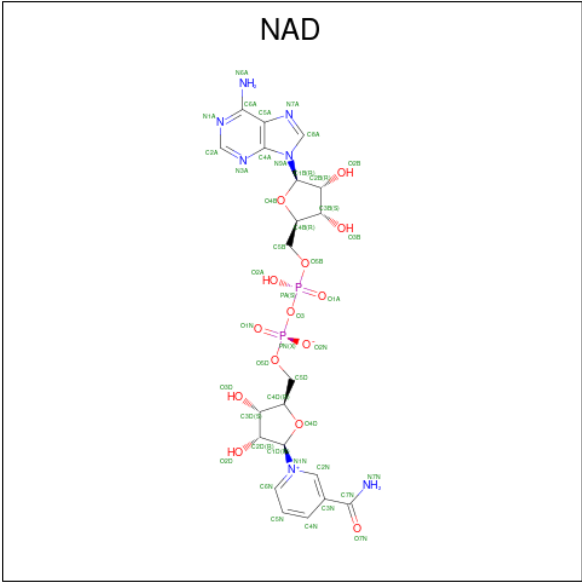
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			15	8	1	6		
2	B	1	Total	C	N	O	0	0
			15	8	1	6		
2	C	1	Total	C	N	O	0	0
			15	8	1	6		

Continued on next page...

Continued from previous page...

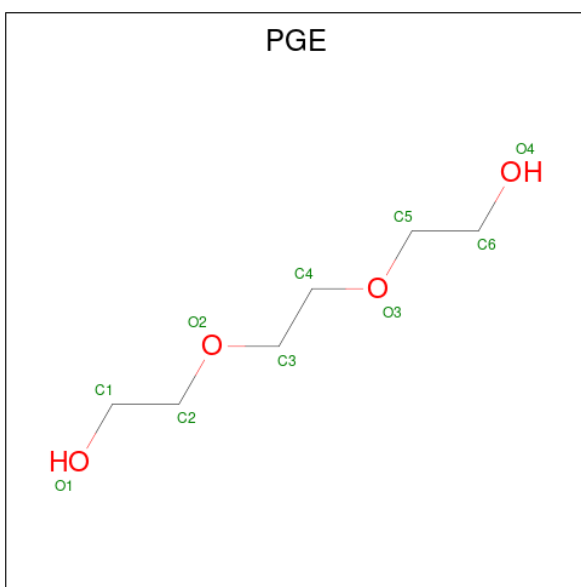
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	D	1	Total	C	N	O	0	0
			15	8	1	6		

- Molecule 3 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>).



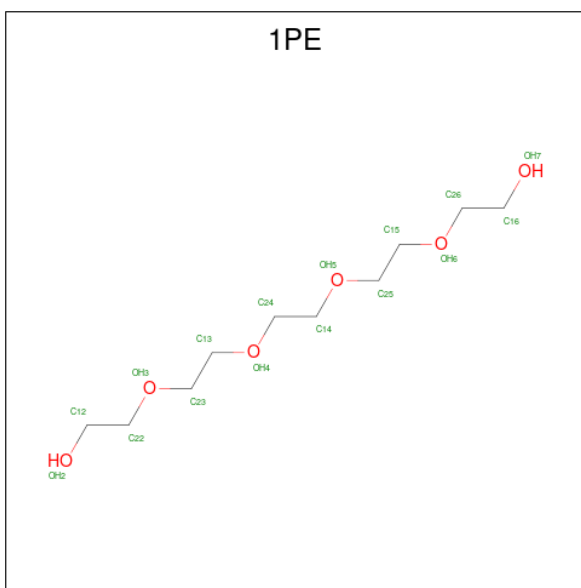
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			44	21	7	14	2		
3	B	1	Total	C	N	O	P	0	0
			44	21	7	14	2		
3	C	1	Total	C	N	O	P	0	0
			44	21	7	14	2		
3	D	1	Total	C	N	O	P	0	0
			44	21	7	14	2		

- Molecule 4 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: C<sub>6</sub>H<sub>14</sub>O<sub>4</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			10	6	4		

- Molecule 5 is PENTAETHYLENE GLYCOL (three-letter code: 1PE) (formula:  $C_{10}H_{22}O_6$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	C	1	Total	C	O	0	0
			16	10	6		

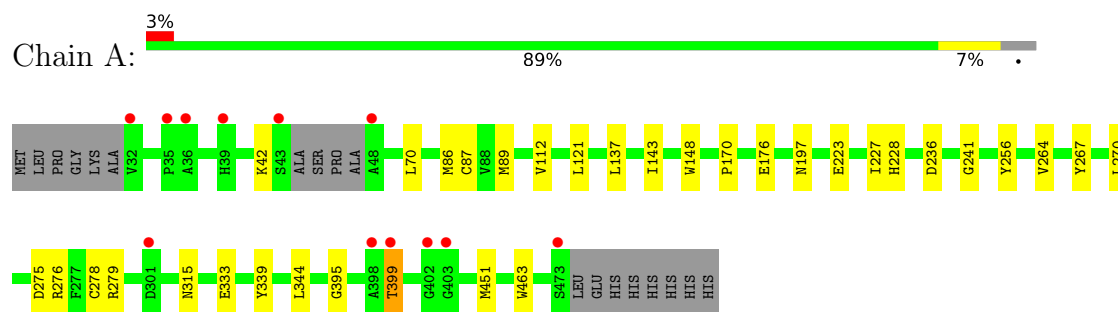
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	276	Total 276	O 276	0	0
6	B	327	Total 327	O 327	0	0
6	C	330	Total 330	O 330	0	0
6	D	313	Total 313	O 313	0	0

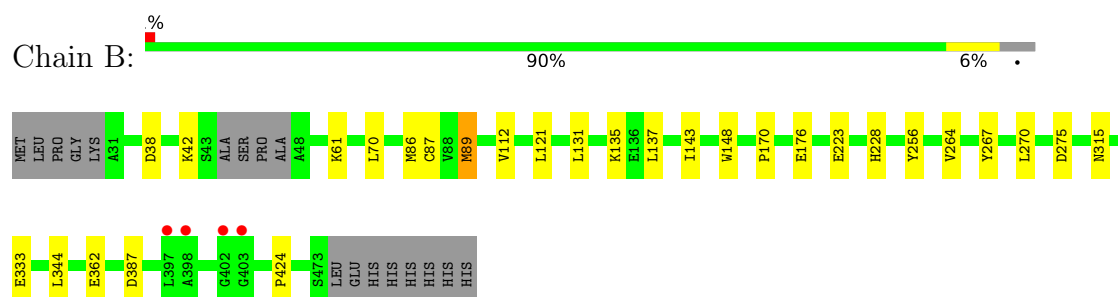
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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.

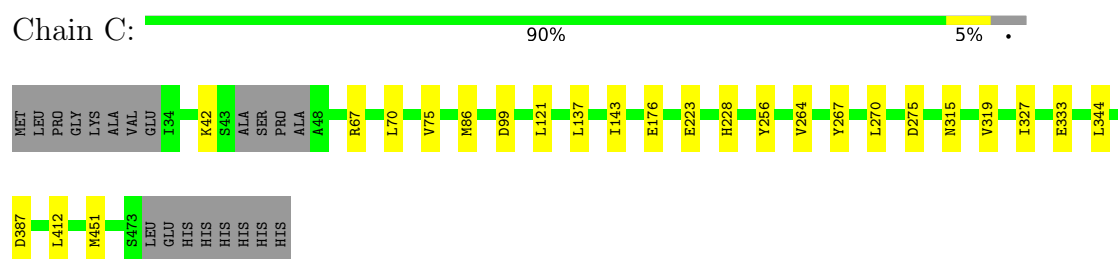
- Molecule 1: Glycosyl hydrolase family 109 protein 2



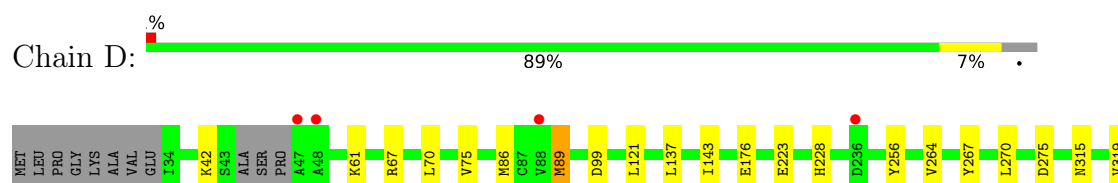
- Molecule 1: Glycosyl hydrolase family 109 protein 2



- Molecule 1: Glycosyl hydrolase family 109 protein 2



- Molecule 1: Glycosyl hydrolase family 109 protein 2





I327	E333	L344	Y369	D387	H388	P389	K400	L412	P424	S473	LEU	GLU	HIS	HIS	HIS	HIS	HIS	HIS	HIS
------	------	------	------	------	------	------	------	------	------	------	-----	-----	-----	-----	-----	-----	-----	-----	-----

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	66.11Å 182.64Å 70.87Å 90.00° 90.14° 90.00°	Depositor
Resolution (Å)	70.87 – 2.13 91.32 – 2.13	Depositor EDS
% Data completeness (in resolution range)	93.6 (70.87-2.13) 93.5 (91.32-2.13)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.12 (at 2.12Å)	Xtriage
Refinement program	PHENIX 1.16_3549	Depositor
R, $R_{free}$	0.158 , 0.194 0.158 , 0.194	Depositor DCC
$R_{free}$ test set	1896 reflections (2.05%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	28.9	Xtriage
Anisotropy	0.198	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 43.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.024 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	15130	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.11% of the height of the origin peak. No significant pseudotranslation is detected.*

<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, 0.333 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: PGE, 1PE, NAD, A2G

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	0.28	0/3501	0.49	0/4733
1	B	0.29	0/3498	0.50	0/4729
1	C	0.29	0/3479	0.49	0/4703
1	D	0.28	0/3484	0.49	0/4709
All	All	0.28	0/13962	0.49	0/18874

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	A	3413	0	3374	18	0
1	B	3413	0	3370	13	0
1	C	3394	0	3354	12	0
1	D	3402	0	3358	15	0
2	A	15	0	15	1	0
2	B	15	0	15	1	0
2	C	15	0	15	1	0
2	D	15	0	15	1	0
3	A	44	0	26	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	44	0	26	1	0
3	C	44	0	26	1	0
3	D	44	0	26	1	0
4	A	10	0	14	0	0
5	C	16	0	22	0	0
6	A	276	0	0	0	0
6	B	327	0	0	0	0
6	C	330	0	0	0	0
6	D	313	0	0	1	0
All	All	15130	0	13656	60	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 2.

All (60) 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:86:MET:HA	1:A:89:MET:HG3	1.67	0.75
1:B:121:LEU:HD22	1:B:137:LEU:HD21	1.72	0.71
1:A:42:LYS:NZ	1:A:176:GLU:OE2	2.28	0.63
1:D:86:MET:HE2	1:D:412:LEU:HD22	1.78	0.63
1:A:121:LEU:HD12	1:A:137:LEU:HD21	1.79	0.63
1:A:451[A]:MET:HE1	1:C:451[A]:MET:HG2	1.83	0.59
1:C:86:MET:HE2	1:C:412:LEU:HD22	1.84	0.59
2:A:501:A2G:H3	3:A:502:NAD:C4N	2.33	0.59
2:D:501:A2G:H3	3:D:502:NAD:C4N	2.34	0.58
2:B:501:A2G:H3	3:B:502:NAD:C4N	2.34	0.57
1:B:86:MET:HA	1:B:89:MET:HG3	1.86	0.57
1:C:121:LEU:HD12	1:C:137:LEU:HD21	1.86	0.57
1:D:86:MET:HA	1:D:89:MET:HG3	1.87	0.56
1:B:70:LEU:HD11	1:B:143:ILE:HG13	1.88	0.55
2:C:501:A2G:H3	3:C:502:NAD:C4N	2.37	0.54
1:C:223:GLU:HB2	1:C:344:LEU:HB3	1.92	0.52
1:D:223:GLU:HB2	1:D:344:LEU:HB3	1.93	0.51
1:D:121:LEU:HD12	1:D:137:LEU:HD21	1.93	0.51
1:D:228:HIS:O	1:D:333:GLU:HA	2.11	0.51
1:D:61:LYS:HG2	1:D:424:PRO:HD3	1.92	0.51
1:A:70:LEU:HD11	1:A:143:ILE:HG13	1.92	0.50
1:B:223:GLU:HB2	1:B:344:LEU:HB3	1.94	0.50
1:C:70:LEU:HD11	1:C:143:ILE:HG13	1.93	0.49
1:A:223:GLU:HB2	1:A:344:LEU:HB3	1.93	0.49

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:42:LYS:NZ	1:B:176:GLU:OE2	2.35	0.48
1:C:228:HIS:O	1:C:333:GLU:HA	2.13	0.48
1:C:75:VAL:HG22	1:C:99:ASP:HB2	1.96	0.47
1:A:270:LEU:HD23	1:A:275:ASP:HB3	1.96	0.47
1:D:70:LEU:HD11	1:D:143:ILE:HG13	1.96	0.47
1:A:228:HIS:O	1:A:333:GLU:HA	2.15	0.47
1:B:61:LYS:HG2	1:B:424:PRO:HD3	1.96	0.46
1:C:42:LYS:NZ	1:C:176:GLU:OE2	2.42	0.46
1:B:148:TRP:CD1	1:B:170:PRO:HD2	2.50	0.45
1:C:264:VAL:HA	1:C:267:TYR:CD2	2.51	0.45
1:D:75:VAL:HG22	1:D:99:ASP:HB2	1.98	0.45
1:A:87[A]:CYS:SG	1:A:112:VAL:HG22	2.56	0.45
1:B:270:LEU:HD23	1:B:275:ASP:HB3	1.98	0.45
1:D:270:LEU:HD23	1:D:275:ASP:HB3	1.98	0.45
1:A:148:TRP:CD1	1:A:170:PRO:HD2	2.52	0.44
1:B:228:HIS:O	1:B:333:GLU:HA	2.17	0.44
1:D:389:PRO:HB2	1:D:473:SER:HA	1.99	0.44
1:C:319:VAL:HB	1:C:327:ILE:HB	2.00	0.43
1:B:131:LEU:O	1:B:135:LYS:HG3	2.18	0.43
1:D:264:VAL:HA	1:D:267:TYR:CD2	2.54	0.42
1:D:400:LYS:HB2	1:D:400:LYS:HE3	1.80	0.42
1:B:87[A]:CYS:SG	1:B:112:VAL:HG22	2.59	0.42
1:D:369:TYR:OH	6:D:601:HOH:O	2.19	0.42
1:A:278:CYS:SG	1:A:279:ARG:NH1	2.93	0.42
1:A:451[A]:MET:CE	1:C:451[A]:MET:HG2	2.49	0.42
1:A:236:ASP:HB2	1:A:241:GLY:HA3	2.02	0.41
1:C:270:LEU:HD23	1:C:275:ASP:HB3	2.02	0.41
1:A:264:VAL:HA	1:A:267:TYR:CD2	2.55	0.41
1:B:264:VAL:HA	1:B:267:TYR:CD2	2.55	0.41
1:A:395:GLY:O	1:A:399:THR:OG1	2.36	0.41
1:A:227:ILE:O	1:A:339:TYR:HA	2.21	0.41
1:A:276:ARG:HD2	1:A:463:TRP:CE2	2.55	0.41
1:B:362:GLU:HG3	1:B:362:GLU:H	1.65	0.41
1:D:319:VAL:HB	1:D:327:ILE:HB	2.03	0.40
1:D:42:LYS:NZ	1:D:176:GLU:OE2	2.46	0.40
1:A:197:ASN:N	1:A:197:ASN:OD1	2.50	0.40

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	437/456 (96%)	425 (97%)	12 (3%)	0	100	100
1	B	437/456 (96%)	425 (97%)	12 (3%)	0	100	100
1	C	434/456 (95%)	423 (98%)	11 (2%)	0	100	100
1	D	435/456 (95%)	423 (97%)	12 (3%)	0	100	100
All	All	1743/1824 (96%)	1696 (97%)	47 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 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	359/370 (97%)	356 (99%)	3 (1%)	83	86
1	B	358/370 (97%)	353 (99%)	5 (1%)	69	73
1	C	356/370 (96%)	352 (99%)	4 (1%)	76	78
1	D	356/370 (96%)	351 (99%)	5 (1%)	69	73
All	All	1429/1480 (97%)	1412 (99%)	17 (1%)	74	76

All (17) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	256	TYR
1	A	315	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	399	THR
1	B	38	ASP
1	B	89	MET
1	B	256	TYR
1	B	315	ASN
1	B	387	ASP
1	C	67	ARG
1	C	256	TYR
1	C	315	ASN
1	C	387	ASP
1	D	67	ARG
1	D	89	MET
1	D	256	TYR
1	D	315	ASN
1	D	387	ASP

Some 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 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	A2G	A	501	-	15,15,15	1.70	4 (26%)	21,21,21	1.25	2 (9%)
3	NAD	A	502	-	39,48,48	2.59	16 (41%)	44,73,73	1.93	9 (20%)
4	PGE	A	503	-	9,9,9	0.52	0	8,8,8	0.26	0
2	A2G	B	501	-	15,15,15	1.69	3 (20%)	21,21,21	1.18	2 (9%)
3	NAD	B	502	-	39,48,48	2.54	16 (41%)	44,73,73	1.96	9 (20%)
2	A2G	C	501	-	15,15,15	1.72	4 (26%)	21,21,21	1.17	2 (9%)
3	NAD	C	502	-	39,48,48	2.55	16 (41%)	44,73,73	2.03	9 (20%)
5	1PE	C	503	-	15,15,15	0.52	0	14,14,14	0.31	0
2	A2G	D	501	-	15,15,15	1.75	4 (26%)	21,21,21	1.27	2 (9%)
3	NAD	D	502	-	39,48,48	2.56	16 (41%)	44,73,73	2.06	8 (18%)

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	A2G	A	501	-	-	0/6/26/26	0/1/1/1
3	NAD	A	502	-	-	2/22/62/62	0/5/5/5
4	PGE	A	503	-	-	5/7/7/7	-
2	A2G	B	501	-	-	0/6/26/26	0/1/1/1
3	NAD	B	502	-	-	2/22/62/62	0/5/5/5
2	A2G	C	501	-	-	0/6/26/26	0/1/1/1
3	NAD	C	502	-	-	2/22/62/62	0/5/5/5
5	1PE	C	503	-	-	7/13/13/13	-
2	A2G	D	501	-	-	0/6/26/26	0/1/1/1
3	NAD	D	502	-	-	2/22/62/62	0/5/5/5

All (79) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	502	NAD	O4D-C1D	8.05	1.52	1.41
3	C	502	NAD	O4D-C1D	7.92	1.52	1.41
3	B	502	NAD	O4D-C1D	7.82	1.52	1.41
3	D	502	NAD	O4D-C1D	7.50	1.51	1.41
3	A	502	NAD	O4B-C1B	6.16	1.49	1.41

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	502	NAD	O4B-C1B	5.74	1.49	1.41
3	D	502	NAD	O4B-C1B	5.74	1.49	1.41
3	C	502	NAD	C7N-N7N	5.69	1.43	1.33
3	D	502	NAD	C7N-N7N	5.47	1.43	1.33
3	C	502	NAD	O4B-C1B	5.46	1.48	1.41
3	B	502	NAD	C7N-N7N	5.43	1.43	1.33
3	A	502	NAD	C7N-N7N	5.36	1.43	1.33
3	D	502	NAD	C4N-C3N	4.87	1.47	1.39
3	A	502	NAD	C4N-C3N	4.80	1.47	1.39
3	B	502	NAD	C4N-C3N	4.78	1.47	1.39
3	C	502	NAD	C4N-C3N	4.77	1.47	1.39
2	D	501	A2G	C7-N2	3.24	1.45	1.34
2	C	501	A2G	C3-C2	-3.23	1.46	1.53
3	D	502	NAD	C3N-C7N	3.22	1.55	1.50
3	C	502	NAD	C5N-C4N	3.22	1.45	1.38
3	D	502	NAD	C5N-C4N	3.20	1.45	1.38
2	C	501	A2G	C7-N2	3.18	1.45	1.34
2	B	501	A2G	C7-N2	3.18	1.45	1.34
2	A	501	A2G	C7-N2	3.17	1.45	1.34
3	A	502	NAD	C5N-C4N	3.11	1.45	1.38
3	D	502	NAD	O7N-C7N	-3.00	1.18	1.24
3	B	502	NAD	O7N-C7N	-2.92	1.18	1.24
2	D	501	A2G	C2-N2	2.92	1.50	1.45
3	C	502	NAD	C3N-C7N	2.90	1.55	1.50
3	B	502	NAD	C2D-C3D	-2.89	1.45	1.53
3	B	502	NAD	C5N-C4N	2.89	1.45	1.38
3	D	502	NAD	C2D-C3D	-2.88	1.46	1.53
2	A	501	A2G	C3-C2	-2.86	1.47	1.53
3	A	502	NAD	O2D-C2D	2.81	1.49	1.43
3	A	502	NAD	O7N-C7N	-2.80	1.18	1.24
3	D	502	NAD	C4A-N3A	-2.76	1.31	1.35
3	B	502	NAD	C4A-N3A	-2.75	1.31	1.35
3	A	502	NAD	C6A-N6A	2.73	1.44	1.34
3	A	502	NAD	C2D-C3D	-2.73	1.46	1.53
3	B	502	NAD	C6A-N6A	2.72	1.44	1.34
3	D	502	NAD	C6A-N6A	2.70	1.44	1.34
3	C	502	NAD	C6A-N6A	2.68	1.44	1.34
2	B	501	A2G	C3-C2	-2.68	1.48	1.53
3	C	502	NAD	O2D-C2D	2.67	1.49	1.43
3	C	502	NAD	O7N-C7N	-2.65	1.19	1.24
3	C	502	NAD	C2D-C3D	-2.64	1.46	1.53
3	A	502	NAD	C3N-C7N	2.62	1.54	1.50

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	502	NAD	O2D-C2D	2.61	1.49	1.43
3	A	502	NAD	C4A-N3A	-2.56	1.32	1.35
3	C	502	NAD	C4A-N3A	-2.54	1.32	1.35
3	D	502	NAD	O2D-C2D	2.50	1.48	1.43
2	D	501	A2G	C3-C2	-2.49	1.48	1.53
3	C	502	NAD	O4D-C4D	2.46	1.50	1.45
3	B	502	NAD	O2B-C2B	2.44	1.48	1.43
3	C	502	NAD	C2B-C3B	-2.44	1.47	1.53
3	B	502	NAD	C2B-C3B	-2.41	1.47	1.53
2	B	501	A2G	C2-N2	2.41	1.49	1.45
2	C	501	A2G	C2-N2	2.36	1.49	1.45
3	A	502	NAD	O4D-C4D	2.36	1.50	1.45
3	B	502	NAD	C3N-C7N	2.34	1.54	1.50
2	D	501	A2G	O3-C3	2.32	1.48	1.43
3	D	502	NAD	O2B-C2B	2.32	1.48	1.43
3	D	502	NAD	C3D-C4D	-2.32	1.47	1.53
3	D	502	NAD	C2B-C3B	-2.31	1.47	1.53
3	D	502	NAD	O4D-C4D	2.31	1.50	1.45
3	B	502	NAD	C3D-C4D	-2.27	1.47	1.53
3	C	502	NAD	C3D-C4D	-2.24	1.47	1.53
3	A	502	NAD	C3D-C4D	-2.24	1.47	1.53
2	A	501	A2G	O3-C3	2.23	1.48	1.43
3	A	502	NAD	O2B-C2B	2.22	1.48	1.43
3	A	502	NAD	C2B-C3B	-2.21	1.47	1.53
3	C	502	NAD	O2B-C2B	2.20	1.48	1.43
3	D	502	NAD	C5A-N7A	2.16	1.47	1.39
2	A	501	A2G	C2-N2	2.16	1.49	1.45
3	B	502	NAD	O4D-C4D	2.15	1.49	1.45
3	B	502	NAD	C5A-N7A	2.06	1.47	1.39
3	A	502	NAD	C5A-N7A	2.04	1.47	1.39
2	C	501	A2G	O3-C3	2.04	1.47	1.43
3	C	502	NAD	C5A-N7A	2.03	1.47	1.39

All (43) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	502	NAD	C4B-O4B-C1B	-7.73	101.77	109.83
3	D	502	NAD	C4B-O4B-C1B	-6.86	102.67	109.83
3	B	502	NAD	C4B-O4B-C1B	-6.86	102.67	109.83
3	D	502	NAD	C4D-O4D-C1D	-6.05	103.52	109.83
3	A	502	NAD	C4B-O4B-C1B	-6.00	103.57	109.83
3	B	502	NAD	C4D-O4D-C1D	-5.39	104.21	109.83

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	502	NAD	C4D-O4D-C1D	-5.24	104.36	109.83
3	A	502	NAD	N3A-C2A-N1A	-4.45	121.50	128.68
3	C	502	NAD	C4D-O4D-C1D	-4.37	105.27	109.83
3	C	502	NAD	N3A-C2A-N1A	-4.32	121.72	128.68
3	B	502	NAD	N3A-C2A-N1A	-4.12	122.03	128.68
3	D	502	NAD	N3A-C2A-N1A	-4.10	122.07	128.68
3	C	502	NAD	C1B-N9A-C4A	-4.01	119.71	126.64
3	B	502	NAD	C4A-C5A-N7A	-3.99	105.25	109.40
3	D	502	NAD	C4A-C5A-N7A	-3.96	105.28	109.40
3	C	502	NAD	C4A-C5A-N7A	-3.94	105.30	109.40
3	D	502	NAD	C1B-N9A-C4A	-3.84	120.01	126.64
3	A	502	NAD	C1B-N9A-C4A	-3.82	120.04	126.64
3	A	502	NAD	C4A-C5A-N7A	-3.58	105.66	109.40
3	D	502	NAD	PN-O3-PA	-3.57	121.24	132.57
3	C	502	NAD	PN-O3-PA	-3.32	122.01	132.57
3	A	502	NAD	PN-O3-PA	-3.28	122.13	132.57
3	B	502	NAD	PN-O3-PA	-3.25	122.25	132.57
2	D	501	A2G	C4-C3-C2	3.23	115.14	110.35
3	B	502	NAD	C1B-N9A-C4A	-3.20	121.11	126.64
2	B	501	A2G	C4-C3-C2	2.67	114.31	110.35
3	A	502	NAD	C2D-C3D-C4D	2.50	107.39	102.60
3	A	502	NAD	C2B-C3B-C4B	2.48	107.35	102.60
3	C	502	NAD	C5N-C4N-C3N	-2.46	117.39	120.34
2	A	501	A2G	O-C1-C2	2.44	111.97	109.52
3	C	502	NAD	C2D-C3D-C4D	2.33	107.06	102.60
2	C	501	A2G	C4-C3-C2	2.31	113.78	110.35
3	D	502	NAD	C2B-C3B-C4B	2.28	106.97	102.60
3	D	502	NAD	C2D-C3D-C4D	2.28	106.97	102.60
2	D	501	A2G	O-C1-C2	2.27	111.80	109.52
3	C	502	NAD	O3D-C3D-C4D	-2.27	104.52	111.07
2	C	501	A2G	O-C1-C2	2.27	111.79	109.52
3	B	502	NAD	C2B-C3B-C4B	2.22	106.86	102.60
3	A	502	NAD	O3D-C3D-C4D	-2.16	104.83	111.07
2	A	501	A2G	C4-C3-C2	2.14	113.52	110.35
3	B	502	NAD	C5N-C4N-C3N	-2.10	117.82	120.34
3	B	502	NAD	C2D-C3D-C4D	2.06	106.54	102.60
2	B	501	A2G	C8-C7-N2	2.00	119.56	116.10

There are no chirality outliers.

All (20) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	503	PGE	O1-C1-C2-O2
3	A	502	NAD	O4B-C4B-C5B-O5B
5	C	503	1PE	OH4-C13-C23-OH3
5	C	503	1PE	OH5-C14-C24-OH4
3	A	502	NAD	C3B-C4B-C5B-O5B
4	A	503	PGE	O2-C3-C4-O3
3	D	502	NAD	O4B-C4B-C5B-O5B
3	D	502	NAD	C3B-C4B-C5B-O5B
5	C	503	1PE	OH7-C16-C26-OH6
5	C	503	1PE	OH6-C15-C25-OH5
4	A	503	PGE	O3-C5-C6-O4
3	B	502	NAD	O4B-C4B-C5B-O5B
3	C	502	NAD	O4B-C4B-C5B-O5B
5	C	503	1PE	C13-C23-OH3-C22
5	C	503	1PE	C24-C14-OH5-C25
4	A	503	PGE	C1-C2-O2-C3
4	A	503	PGE	C3-C4-O3-C5
3	B	502	NAD	C3B-C4B-C5B-O5B
3	C	502	NAD	C3B-C4B-C5B-O5B
5	C	503	1PE	C25-C15-OH6-C26

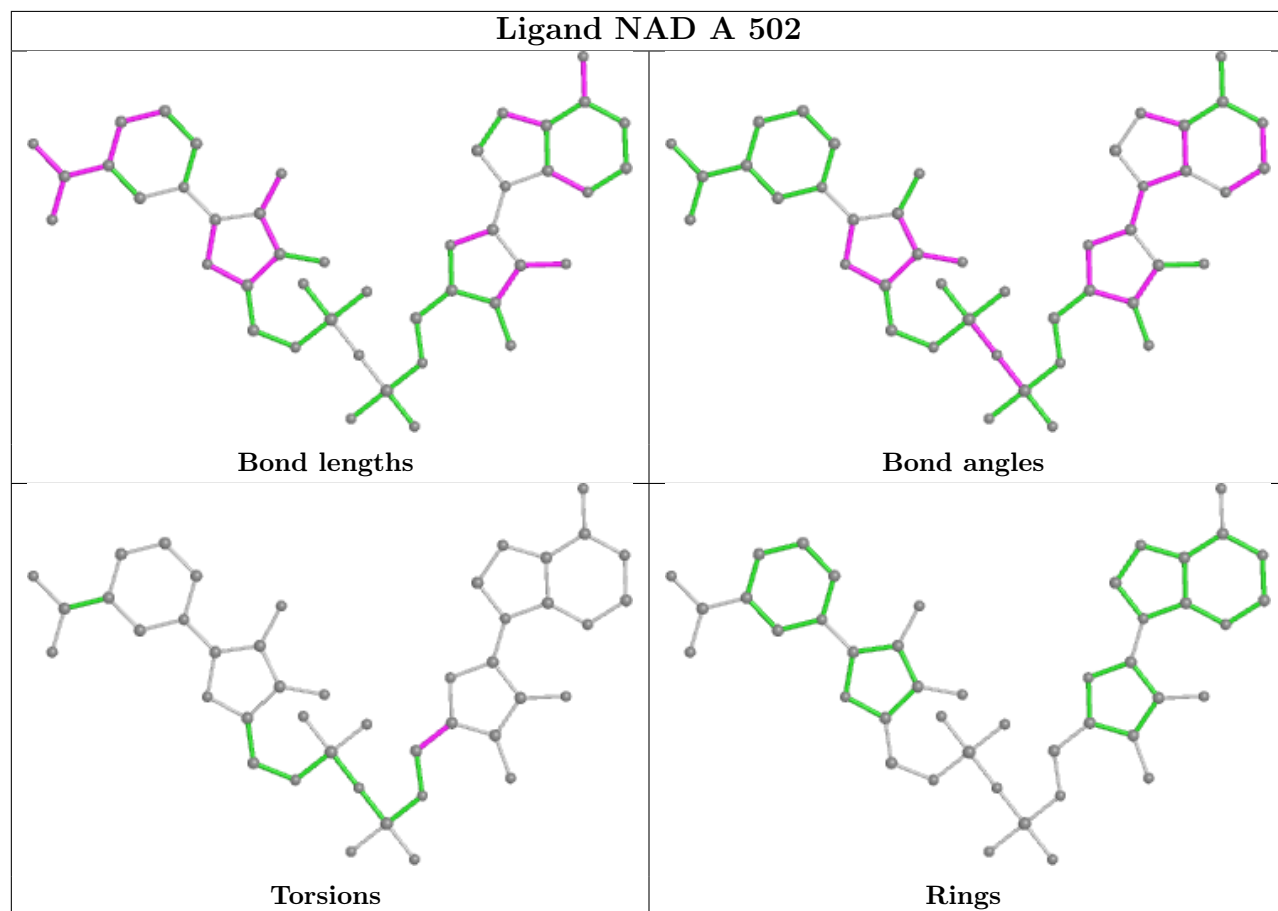
There are no ring outliers.

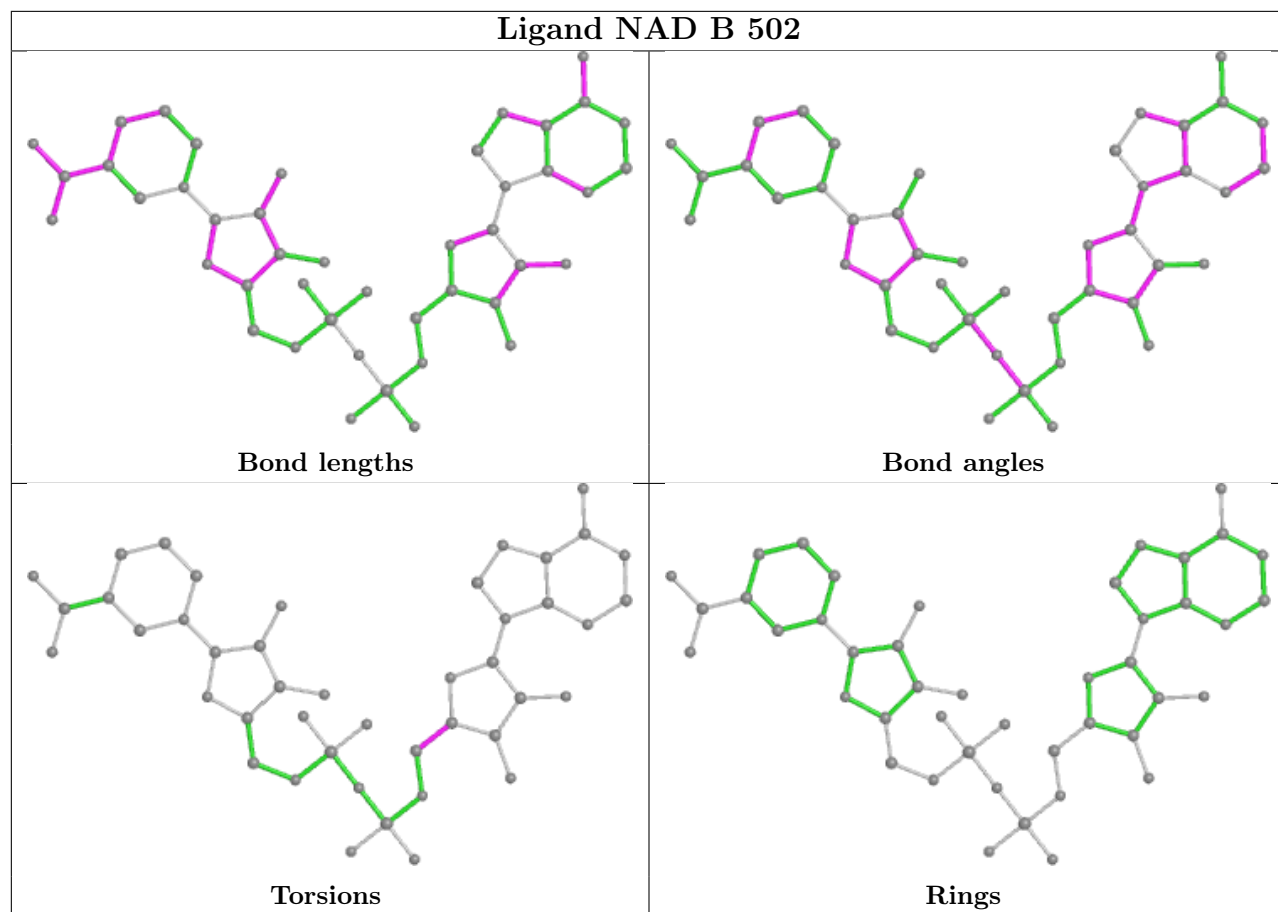
8 monomers are involved in 4 short contacts:

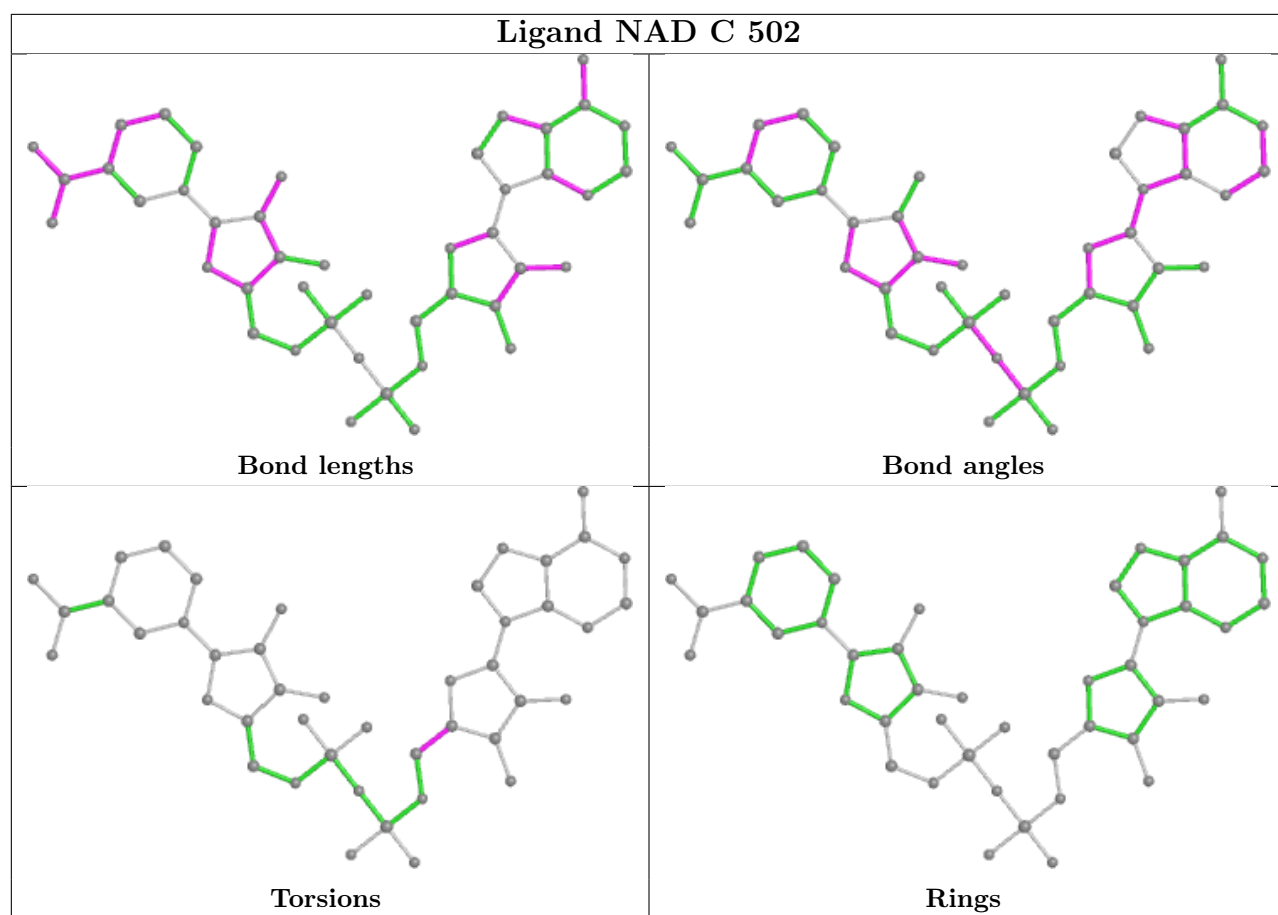
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	501	A2G	1	0
3	A	502	NAD	1	0
2	B	501	A2G	1	0
3	B	502	NAD	1	0
2	C	501	A2G	1	0
3	C	502	NAD	1	0
2	D	501	A2G	1	0
3	D	502	NAD	1	0

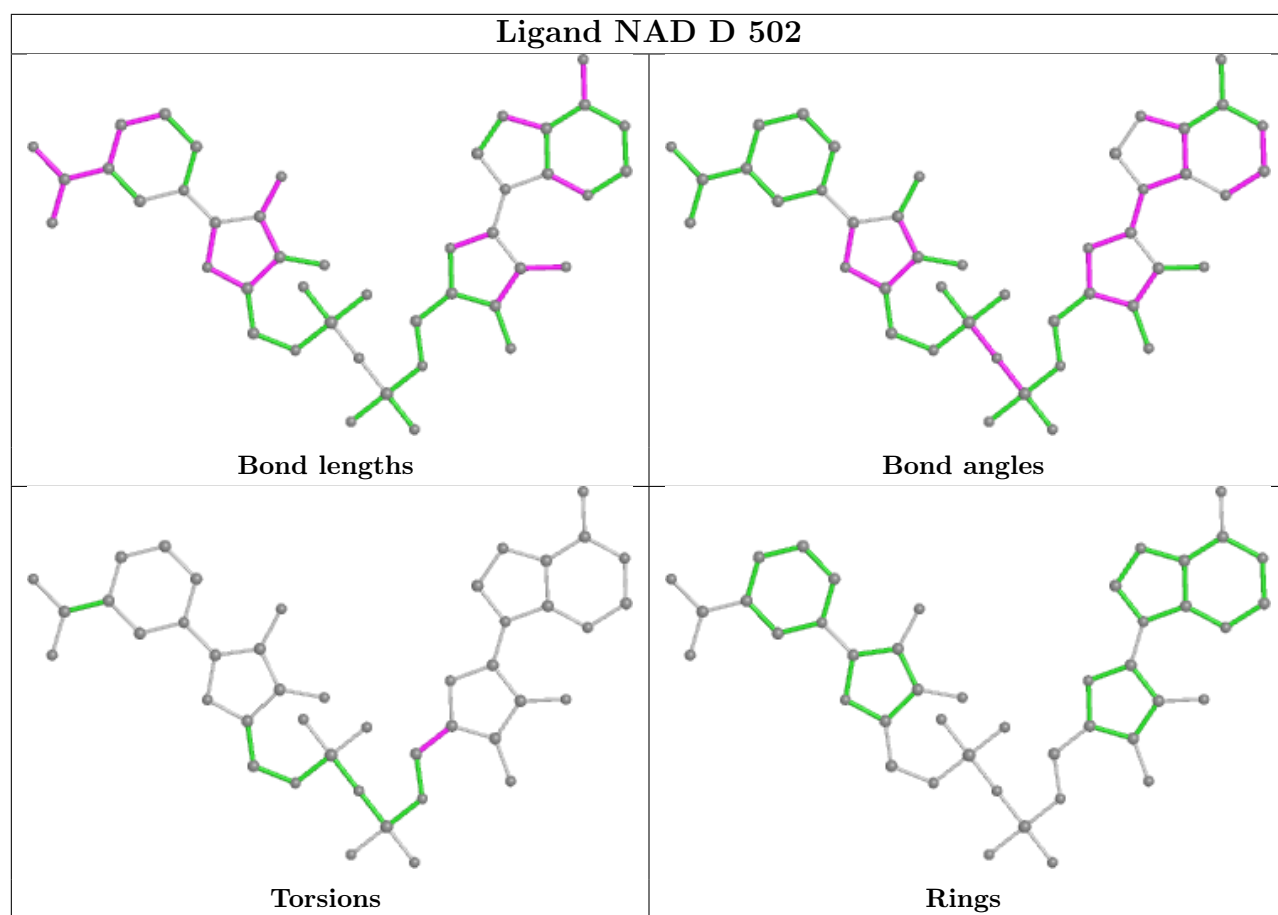
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 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 ⓘ

### 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<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	438/456 (96%)	-0.46	12 (2%) 54 61	22, 32, 69, 93	0
1	B	439/456 (96%)	-0.57	4 (0%) 84 87	19, 28, 54, 82	0
1	C	436/456 (95%)	-0.60	0 100 100	20, 28, 52, 71	0
1	D	437/456 (95%)	-0.60	4 (0%) 84 87	18, 29, 53, 82	0
All	All	1750/1824 (95%)	-0.56	20 (1%) 80 84	18, 30, 55, 93	0

All (20) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	397	LEU	3.6
1	A	473	SER	3.6
1	A	48	ALA	3.4
1	B	398	ALA	3.3
1	A	35	PRO	3.0
1	A	39	HIS	3.0
1	B	403	GLY	2.9
1	A	36	ALA	2.7
1	D	47	ALA	2.7
1	A	301	ASP	2.6
1	D	48	ALA	2.6
1	A	398	ALA	2.5
1	A	403	GLY	2.3
1	A	32	VAL	2.2
1	A	399	THR	2.1
1	B	402	GLY	2.1
1	A	43	SER	2.1
1	D	236	ASP	2.0
1	A	402	GLY	2.0
1	D	88	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 ⓘ

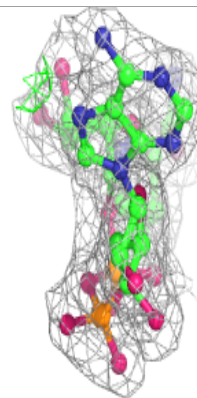
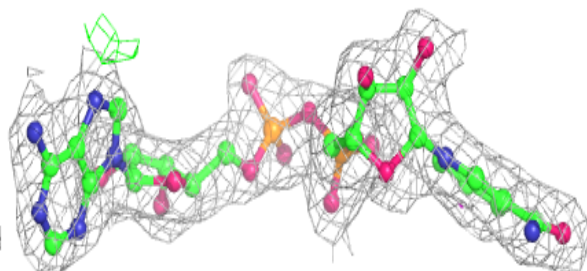
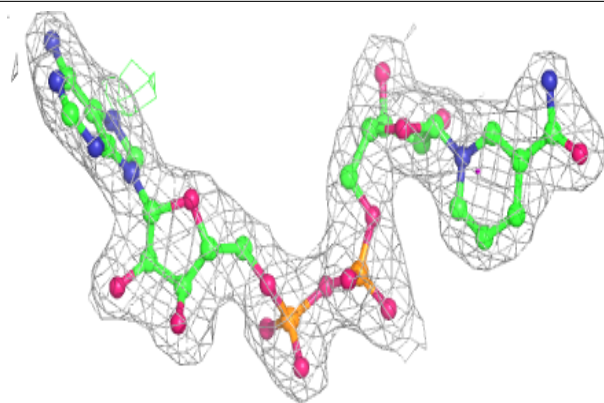
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	PGE	A	503	10/10	0.83	0.14	58,59,61,61	0
5	1PE	C	503	16/16	0.93	0.14	37,42,45,47	0
2	A2G	B	501	15/15	0.95	0.08	26,28,30,31	0
2	A2G	D	501	15/15	0.96	0.08	22,25,29,29	0
2	A2G	A	501	15/15	0.96	0.08	29,31,33,33	0
2	A2G	C	501	15/15	0.97	0.08	25,27,31,32	0
3	NAD	A	502	44/44	0.97	0.08	27,31,34,35	0
3	NAD	D	502	44/44	0.98	0.08	18,26,30,31	0
3	NAD	B	502	44/44	0.98	0.08	19,25,27,27	0
3	NAD	C	502	44/44	0.98	0.08	20,25,28,28	0

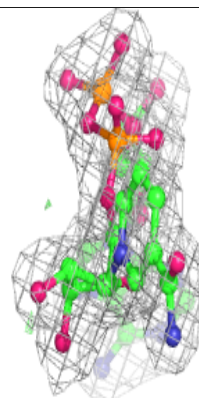
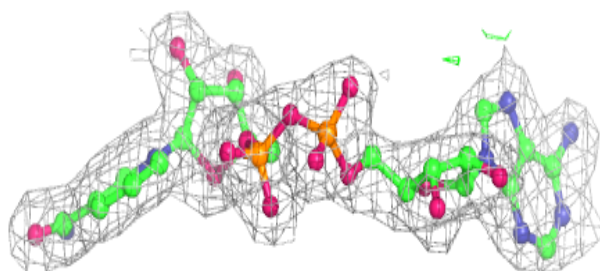
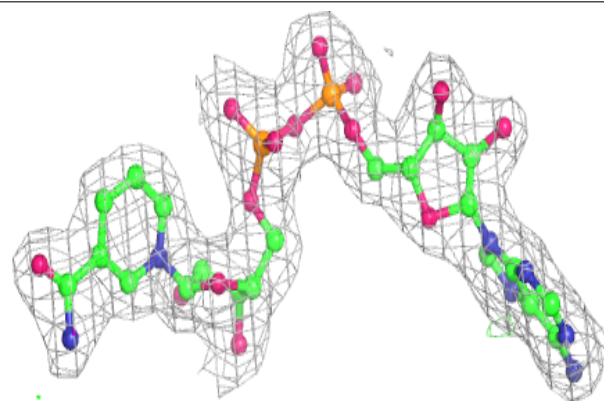
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

**Electron density around NAD A 502:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)

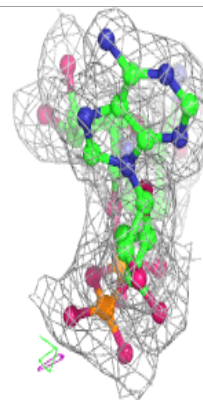
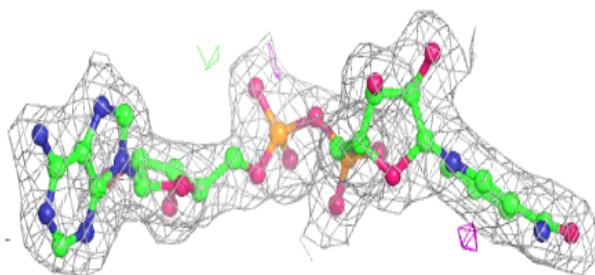
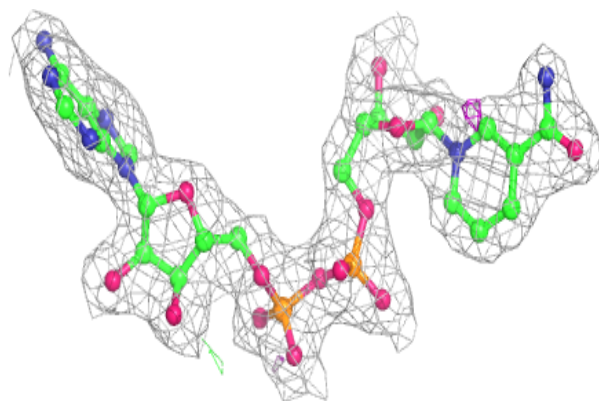
**Electron density around NAD D 502:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)

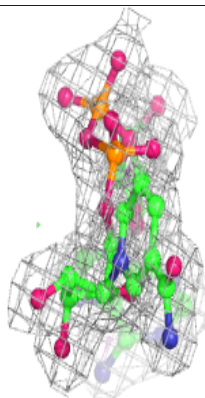
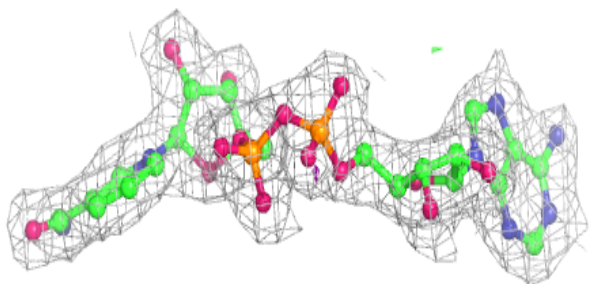
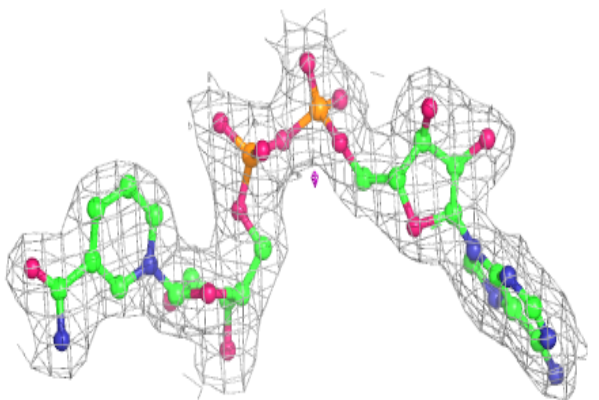


**Electron density around NAD B 502:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)

**Electron density around NAD C 502:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



## 6.5 Other polymers [i](#)

There are no such residues in this entry.