



# Full wwPDB X-ray Structure Validation Report ⓘ

May 23, 2020 – 06:51 pm BST

PDB ID : 2V7G  
Title : Crystal Structure of an Engineered Urocanase Tetramer  
Authors : Treiber, N.; Schulz, G.E.  
Deposited on : 2007-07-30  
Resolution : 2.00 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 : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.11  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
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) : 2.11

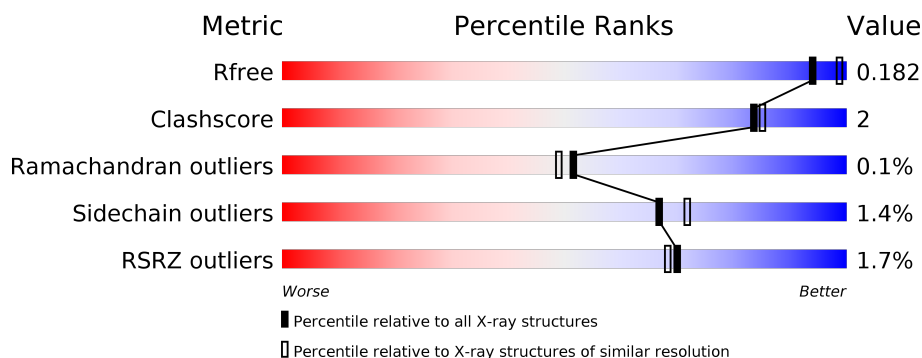
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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))
$R_{free}$	130704	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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	557	<div> <div>92%</div> <div>7%</div> </div>
1	B	557	<div> <div>94%</div> <div>6%</div> </div>
1	C	557	<div> <div>2%</div> <div>91%</div> <div>8%</div> </div>
1	D	557	<div> <div>4%</div> <div>92%</div> <div>7%</div> </div>

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 18528 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 UROCANATE HYDRATASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	553	Total	C	N	O	S	0	0	0
			4240	2664	757	797	22			
1	B	553	Total	C	N	O	S	0	0	0
			4240	2664	757	797	22			
1	C	553	Total	C	N	O	S	0	0	0
			4240	2664	757	797	22			
1	D	553	Total	C	N	O	S	0	0	0
			4240	2664	757	797	22			

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	198	SER	CYS	engineered mutation	UNP P25080
A	277	ILE	ALA	engineered mutation	UNP P25080
A	291	ILE	GLU	engineered mutation	UNP P25080
A	301	ALA	GLN	engineered mutation	UNP P25080
B	198	SER	CYS	engineered mutation	UNP P25080
B	277	ILE	ALA	engineered mutation	UNP P25080
B	291	ILE	GLU	engineered mutation	UNP P25080
B	301	ALA	GLN	engineered mutation	UNP P25080
C	198	SER	CYS	engineered mutation	UNP P25080
C	277	ILE	ALA	engineered mutation	UNP P25080
C	291	ILE	GLU	engineered mutation	UNP P25080
C	301	ALA	GLN	engineered mutation	UNP P25080
D	198	SER	CYS	engineered mutation	UNP P25080
D	277	ILE	ALA	engineered mutation	UNP P25080
D	291	ILE	GLU	engineered mutation	UNP P25080
D	301	ALA	GLN	engineered mutation	UNP P25080
A	164	SER	THR	conflict	UNP P25080
B	164	SER	THR	conflict	UNP P25080
C	164	SER	THR	conflict	UNP P25080
D	164	SER	THR	conflict	UNP P25080
A	165	LEU	VAL	conflict	UNP P25080

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	165	LEU	VAL	conflict	UNP P25080
C	165	LEU	VAL	conflict	UNP P25080
D	165	LEU	VAL	conflict	UNP P25080
A	167	GLY	ALA	conflict	UNP P25080
B	167	GLY	ALA	conflict	UNP P25080
C	167	GLY	ALA	conflict	UNP P25080
D	167	GLY	ALA	conflict	UNP P25080

- Molecule 2 is ACETATE ION (three-letter code: ACT) (formula:  $C_2H_3O_2$ ).



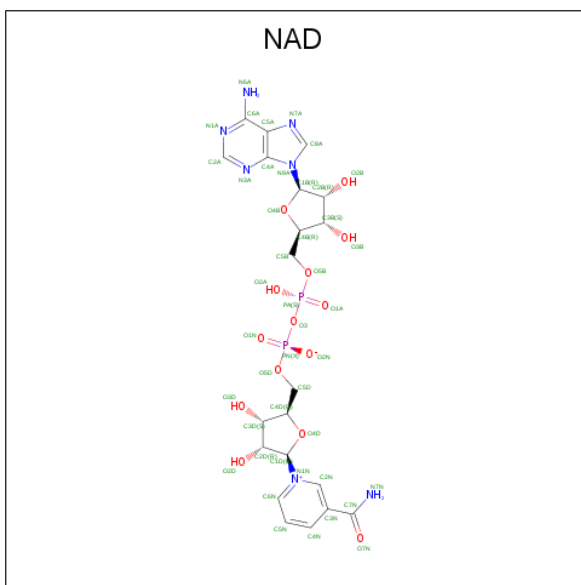
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			4	2	2		
2	A	1	Total	C	O	0	0
			4	2	2		
2	B	1	Total	C	O	0	0
			4	2	2		
2	B	1	Total	C	O	0	0
			4	2	2		
2	B	1	Total	C	O	0	0
			4	2	2		
2	C	1	Total	C	O	0	0
			4	2	2		
2	C	1	Total	C	O	0	0
			4	2	2		

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			6	3	3		
3	A	1	Total	C	O	0	0
			6	3	3		
3	B	1	Total	C	O	0	0
			6	3	3		
3	B	1	Total	C	O	0	0
			6	3	3		
3	B	1	Total	C	O	0	0
			6	3	3		
3	D	1	Total	C	O	0	0
			6	3	3		
3	D	1	Total	C	O	0	0
			6	3	3		
3	D	1	Total	C	O	0	0
			6	3	3		

- Molecule 4 is NICOTINAMIDE-ADENINE-DINUCLEOTIDE (three-letter code: NAD) (formula:  $C_{21}H_{27}N_7O_{14}P_2$ ).



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

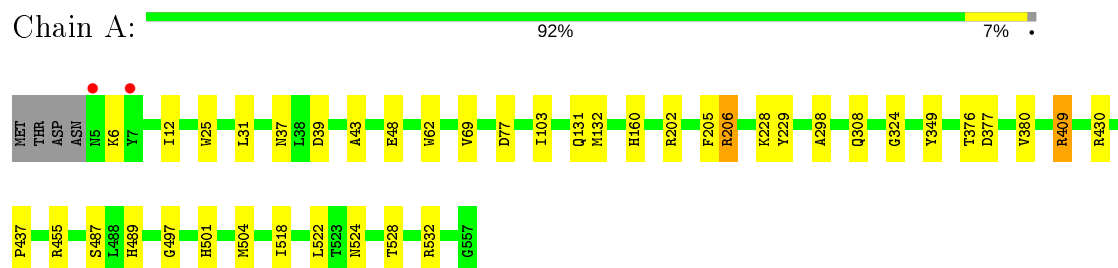
- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	347	Total O 347 347	0	0
5	B	341	Total O 341 341	0	0
5	C	330	Total O 330 330	0	0
5	D	292	Total O 292 292	0	0

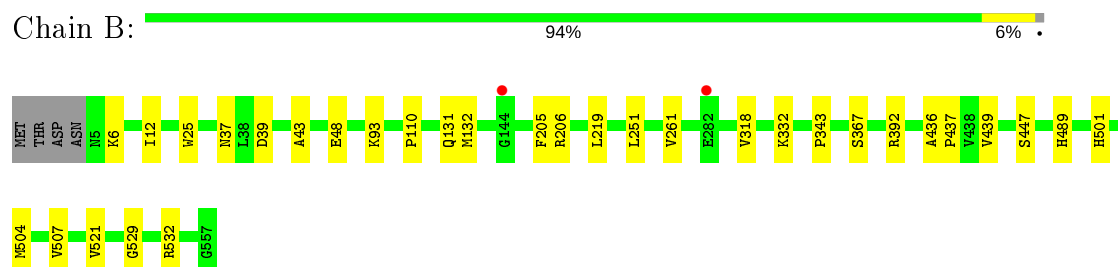
### 3 Residue-property plots [i](#)

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.

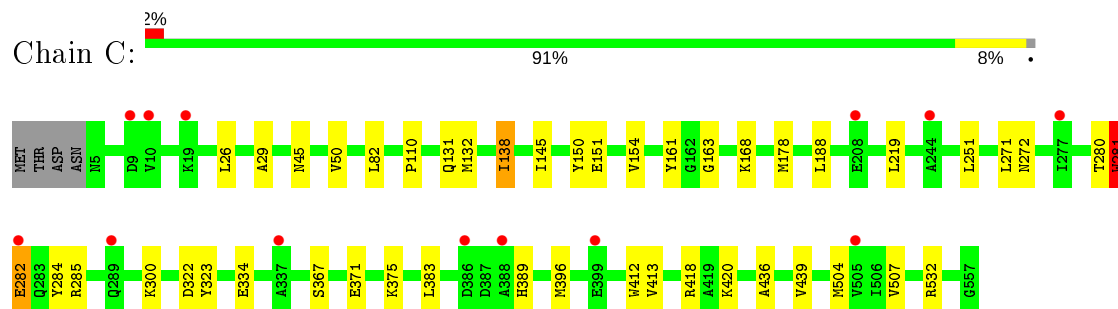
#### • Molecule 1: UROCANATE HYDRATASE



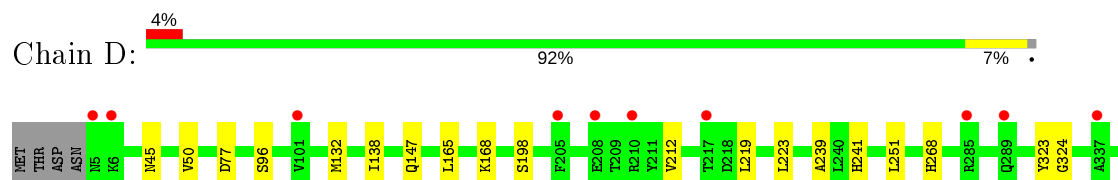
#### • Molecule 1: UROCANATE HYDRATASE

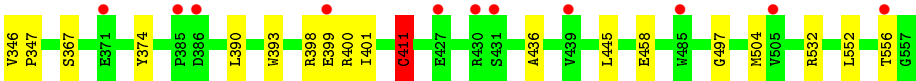


#### • Molecule 1: UROCANATE HYDRATASE



#### • Molecule 1: UROCANATE HYDRATASE







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	68.07Å 154.32Å 216.39Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	28.81 – 2.00 28.81 – 2.00	Depositor EDS
% Data completeness (in resolution range)	100.0 (28.81-2.00) 97.0 (28.81-2.00)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	6.19 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.156 , 0.181 0.157 , 0.182	Depositor DCC
$R_{free}$ test set	1498 reflections (1.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	23.3	Xtriage
Anisotropy	0.037	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 49.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	18528	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.52% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, NAD, ACT

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	0.43	0/4333	0.57	0/5881
1	B	0.42	0/4333	0.55	0/5881
1	C	0.40	0/4333	0.55	1/5881 (0.0%)
1	D	0.40	1/4333 (0.0%)	0.53	0/5881
All	All	0.41	1/17332 (0.0%)	0.55	1/23524 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	C	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	411	CYS	CB-SG	-5.74	1.72	1.81

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	281	TRP	N-CA-C	5.08	124.73	111.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	281	TRP	Peptide

## 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	4240	0	4160	22	0
1	B	4240	0	4160	18	0
1	C	4240	0	4160	27	0
1	D	4240	0	4160	20	0
2	A	8	0	6	0	0
2	B	12	0	9	0	0
2	C	8	0	6	0	0
3	A	12	0	16	1	0
3	B	18	0	24	1	0
3	D	24	0	32	0	0
4	A	44	0	26	0	0
4	B	44	0	26	0	0
4	C	44	0	26	0	0
4	D	44	0	26	2	0
5	A	347	0	0	0	0
5	B	341	0	0	0	0
5	C	330	0	0	0	0
5	D	292	0	0	1	0
All	All	18528	0	16837	83	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 (83) 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:377:ASP:OD1	1:A:409:ARG:HD2	1.97	0.64
1:C:413:VAL:HG23	1:C:418:ARG:HG2	1.82	0.60
1:C:281:TRP:N	1:C:282:GLU:HB2	2.17	0.60
1:C:300:LYS:HE3	1:C:334:GLU:OE2	2.03	0.58
1:C:280:THR:OG1	1:C:282:GLU:HB3	2.05	0.56
1:D:165:LEU:O	1:D:168:LYS:HB2	2.06	0.55
1:C:281:TRP:O	1:C:284:TYR:HB3	2.07	0.55
1:A:430:ARG:HH12	1:B:392:ARG:HD3	1.72	0.55
1:D:556:THR:HG23	5:D:2284:HOH:O	2.07	0.54

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:439:VAL:HG12	1:C:507:VAL:HG12	1.91	0.53
1:C:45:ASN:HB3	1:C:50:VAL:HB	1.89	0.53
1:B:6:LYS:O	1:B:48:GLU:HA	2.09	0.52
1:A:6:LYS:O	1:A:48:GLU:HA	2.08	0.52
1:A:31:LEU:HD11	1:A:69:VAL:HG13	1.92	0.52
1:D:147:GLN:HG2	1:D:393:TRP:CZ2	2.45	0.52
1:D:219:LEU:HD11	1:D:251:LEU:HD21	1.92	0.52
1:C:145:ILE:CD1	1:C:178:MET:HA	2.41	0.50
1:D:212:VAL:HG11	1:D:239:ALA:HB2	1.93	0.50
1:C:145:ILE:HD11	1:C:178:MET:HA	1.95	0.49
1:D:399:GLU:HG3	1:D:400:ARG:HG2	1.94	0.49
1:D:346:VAL:HB	1:D:347:PRO:HD3	1.93	0.49
1:B:219:LEU:HD11	1:B:251:LEU:HD21	1.95	0.49
1:D:324:GLY:HA3	4:D:3001:NAD:H71N	1.78	0.49
1:B:447:SER:OG	1:B:529:GLY:HA3	2.12	0.48
1:C:271:LEU:O	1:C:285:ARG:NH2	2.42	0.48
1:C:29:ALA:CB	1:C:507:VAL:HG11	2.43	0.48
1:D:390:LEU:HD11	1:D:411:CYS:SG	2.53	0.48
1:B:489:HIS:HB2	1:B:501:HIS:CE1	2.49	0.48
1:D:552:LEU:H	1:D:556:THR:CG2	2.27	0.48
1:A:298:ALA:HB1	3:A:1560:GOL:H31	1.95	0.48
1:A:160:HIS:HE1	1:A:349:TYR:OH	1.96	0.47
1:A:25:TRP:CD2	1:A:437:PRO:HG2	2.50	0.47
1:A:430:ARG:HH12	1:B:392:ARG:CD	2.28	0.46
1:B:504:MET:CE	1:B:521:VAL:CG1	2.93	0.46
1:A:37:ASN:O	1:A:43:ALA:HB2	2.15	0.46
1:A:12:ILE:HD11	1:A:62:TRP:CZ3	2.51	0.46
1:B:37:ASN:O	1:B:43:ALA:HB2	2.16	0.46
1:C:26:LEU:HD23	1:C:507:VAL:HG21	1.98	0.46
1:A:12:ILE:O	1:A:39:ASP:HA	2.16	0.46
1:B:367:SER:HB3	1:B:436:ALA:HB3	1.98	0.46
1:B:110:PRO:HB2	1:D:497:GLY:HA2	1.97	0.45
1:A:524:ASN:O	1:A:528:THR:HG23	2.17	0.44
1:D:367:SER:HB3	1:D:436:ALA:HB3	1.98	0.44
1:D:77:ASP:O	1:D:96:SER:HA	2.17	0.44
1:C:371:GLU:HG3	1:C:375:LYS:HE3	2.00	0.44
1:C:272:ASN:HA	1:C:285:ARG:NH2	2.33	0.44
1:D:45:ASN:HB3	1:D:50:VAL:HB	1.99	0.44
1:A:131:GLN:HB3	1:A:132:MET:H	1.68	0.44
1:C:219:LEU:HD11	1:C:251:LEU:HD21	2.00	0.44
1:B:131:GLN:HB3	1:B:132:MET:H	1.70	0.43

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:393:TRP:CH2	1:D:401:ILE:HD13	2.53	0.43
1:C:131:GLN:HB3	1:C:132:MET:H	1.68	0.43
1:C:396:MET:HB2	1:C:396:MET:HE2	1.88	0.43
1:C:150:TYR:CZ	1:C:154:VAL:HG21	2.53	0.43
1:A:489:HIS:HB2	1:A:501:HIS:CE1	2.53	0.43
1:C:138:ILE:O	1:C:138:ILE:HG13	2.18	0.43
1:A:228:LYS:HE3	1:A:229:TYR:CE1	2.54	0.42
1:B:25:TRP:CD2	1:B:437:PRO:HG2	2.55	0.42
1:C:151:GLU:HG2	1:C:412:TRP:CD2	2.53	0.42
1:C:29:ALA:CB	1:C:507:VAL:CG1	2.98	0.42
1:A:376:THR:O	1:A:380:VAL:HG23	2.19	0.42
1:C:188:LEU:O	1:C:396:MET:HE1	2.19	0.42
1:B:392:ARG:HA	1:B:392:ARG:HD2	1.85	0.42
1:C:367:SER:HB3	1:C:436:ALA:HB3	2.02	0.42
1:A:324:GLY:HA3	1:A:455:ARG:HD2	2.02	0.41
1:B:504:MET:HE3	1:B:521:VAL:CG1	2.50	0.41
1:A:103:ILE:O	1:A:487:SER:HA	2.21	0.41
1:A:518:ILE:HG23	1:A:522:LEU:HD12	2.03	0.41
1:C:383:LEU:HD22	1:C:420:LYS:HE2	2.02	0.41
1:D:268:HIS:CD2	4:D:3001:NAD:O3D	2.73	0.41
1:A:12:ILE:HD11	1:A:62:TRP:HZ3	1.85	0.41
1:C:161:TYR:CD1	1:C:168:LYS:HE2	2.56	0.41
1:D:198:SER:O	1:D:241:HIS:HE1	2.04	0.41
1:D:374:TYR:CE1	1:D:398:ARG:HG3	2.55	0.41
1:D:132:MET:CE	1:D:445:LEU:HD21	2.50	0.41
1:B:439:VAL:HG12	1:B:507:VAL:HG22	2.02	0.41
1:A:497:GLY:HA2	1:C:110:PRO:HB2	2.03	0.41
1:C:163:GLY:HA3	1:C:389:HIS:CE1	2.55	0.41
1:D:132:MET:HE1	1:D:445:LEU:HD21	2.03	0.41
1:B:261:VAL:HG23	1:B:318:VAL:HG11	2.03	0.41
1:B:12:ILE:O	1:B:39:ASP:HA	2.21	0.41
1:B:343:PRO:HG3	3:B:1563:GOL:H32	2.04	0.40
1:A:202:ARG:O	1:A:206:ARG:HD2	2.22	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	551/557 (99%)	535 (97%)	16 (3%)	0	100	100
1	B	551/557 (99%)	531 (96%)	20 (4%)	0	100	100
1	C	551/557 (99%)	539 (98%)	11 (2%)	1 (0%)	47	44
1	D	551/557 (99%)	536 (97%)	14 (2%)	1 (0%)	47	44
All	All	2204/2228 (99%)	2141 (97%)	61 (3%)	2 (0%)	51	49

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	138	ILE
1	D	138	ILE

### 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	435/439 (99%)	428 (98%)	7 (2%)	62	67
1	B	435/439 (99%)	430 (99%)	5 (1%)	73	78
1	C	435/439 (99%)	429 (99%)	6 (1%)	67	72
1	D	435/439 (99%)	429 (99%)	6 (1%)	67	72
All	All	1740/1756 (99%)	1716 (99%)	24 (1%)	67	72

All (24) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	77	ASP
1	A	205	PHE
1	A	206	ARG
1	A	308	GLN
1	A	409	ARG
1	A	504	MET
1	A	532	ARG
1	B	93	LYS
1	B	205	PHE
1	B	206	ARG
1	B	332	LYS
1	B	532	ARG
1	C	82	LEU
1	C	282	GLU
1	C	322	ASP
1	C	323	TYR
1	C	504	MET
1	C	532	ARG
1	D	223	LEU
1	D	323	TYR
1	D	411	CYS
1	D	458	GLU
1	D	504	MET
1	D	532	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (13) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	18	ASN
1	A	160	HIS
1	A	308	GLN
1	B	18	ASN
1	B	241	HIS
1	C	5	ASN
1	C	119	ASN
1	C	268	HIS
1	C	389	HIS
1	C	454	ASN
1	D	18	ASN
1	D	241	HIS
1	D	268	HIS

### 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 ⓘ

20 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	ACT	B	1560	-	1,3,3	1.68	0	0,3,3	0.00	-
3	GOL	D	1560	-	5,5,5	0.36	0	5,5,5	0.19	0
2	ACT	A	1558	-	1,3,3	1.64	0	0,3,3	0.00	-
3	GOL	D	1558	-	5,5,5	0.41	0	5,5,5	0.29	0
4	NAD	B	3001	-	42,48,48	1.22	4 (9%)	50,73,73	1.25	3 (6%)
3	GOL	D	1561	-	5,5,5	0.41	0	5,5,5	0.35	0
3	GOL	A	1561	-	5,5,5	0.35	0	5,5,5	0.44	0
2	ACT	A	1559	-	1,3,3	1.48	0	0,3,3	0.00	-
3	GOL	B	1561	-	5,5,5	0.32	0	5,5,5	0.36	0
4	NAD	A	3001	-	42,48,48	1.33	4 (9%)	50,73,73	1.25	4 (8%)
4	NAD	C	3001	-	42,48,48	1.27	4 (9%)	50,73,73	1.26	6 (12%)
4	NAD	D	3001	-	42,48,48	1.16	4 (9%)	50,73,73	1.28	5 (10%)
3	GOL	B	1562	-	5,5,5	0.42	0	5,5,5	0.43	0
3	GOL	A	1560	-	5,5,5	0.51	0	5,5,5	0.53	0
3	GOL	D	1559	-	5,5,5	0.39	0	5,5,5	0.22	0
2	ACT	B	1558	-	1,3,3	1.52	0	0,3,3	0.00	-
2	ACT	C	1558	-	1,3,3	1.47	0	0,3,3	0.00	-



Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	ACT	C	1559	-	1,3,3	1.69	0	0,3,3	0.00	-
2	ACT	B	1559	-	1,3,3	1.46	0	0,3,3	0.00	-
3	GOL	B	1563	-	5,5,5	0.30	0	5,5,5	0.52	0

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	GOL	D	1560	-	-	4/4/4/4	-
3	GOL	D	1559	-	-	2/4/4/4	-
4	NAD	B	3001	-	-	5/26/62/62	0/5/5/5
3	GOL	D	1561	-	-	2/4/4/4	-
3	GOL	A	1561	-	-	4/4/4/4	-
3	GOL	B	1561	-	-	1/4/4/4	-
4	NAD	A	3001	-	-	5/26/62/62	0/5/5/5
4	NAD	C	3001	-	-	2/26/62/62	0/5/5/5
4	NAD	D	3001	-	-	1/26/62/62	0/5/5/5
3	GOL	B	1562	-	-	0/4/4/4	-
3	GOL	A	1560	-	-	2/4/4/4	-
3	GOL	D	1558	-	-	2/4/4/4	-
3	GOL	B	1563	-	-	3/4/4/4	-

All (16) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	3001	NAD	C2N-N1N	5.72	1.41	1.35
4	B	3001	NAD	C2N-N1N	4.94	1.41	1.35
4	D	3001	NAD	C2N-N1N	4.50	1.40	1.35
4	C	3001	NAD	C2N-N1N	4.48	1.40	1.35
4	C	3001	NAD	O4D-C1D	3.80	1.46	1.41
4	A	3001	NAD	O4D-C1D	3.61	1.46	1.41
4	C	3001	NAD	O4B-C1B	3.44	1.45	1.41
4	D	3001	NAD	O4D-C1D	2.96	1.45	1.41
4	B	3001	NAD	O4B-C1B	2.90	1.45	1.41
4	C	3001	NAD	C6N-N1N	2.62	1.41	1.35
4	B	3001	NAD	C6N-N1N	2.37	1.41	1.35
4	A	3001	NAD	C6N-N1N	2.30	1.41	1.35
4	D	3001	NAD	O4B-C1B	2.26	1.44	1.41

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	3001	NAD	C3N-C7N	2.11	1.53	1.50
4	D	3001	NAD	C6N-N1N	2.07	1.40	1.35
4	B	3001	NAD	O4D-C1D	2.01	1.43	1.41

All (18) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	3001	NAD	N3A-C2A-N1A	-4.71	121.31	128.68
4	D	3001	NAD	N3A-C2A-N1A	-4.69	121.34	128.68
4	C	3001	NAD	N3A-C2A-N1A	-4.65	121.41	128.68
4	B	3001	NAD	N3A-C2A-N1A	-4.64	121.42	128.68
4	B	3001	NAD	O4B-C1B-C2B	-3.36	102.02	106.93
4	D	3001	NAD	C3N-C7N-N7N	-3.13	113.99	117.75
4	D	3001	NAD	O4B-C1B-C2B	-3.08	102.43	106.93
4	A	3001	NAD	O4D-C1D-C2D	-3.06	102.45	106.93
4	C	3001	NAD	O4B-C1B-C2B	-3.03	102.50	106.93
4	B	3001	NAD	O4D-C1D-C2D	-3.01	102.53	106.93
4	C	3001	NAD	C5B-C4B-C3B	-2.33	106.45	115.18
4	D	3001	NAD	C6N-N1N-C2N	-2.29	119.89	121.97
4	A	3001	NAD	O4B-C1B-C2B	-2.28	103.60	106.93
4	A	3001	NAD	C5B-C4B-C3B	-2.21	106.90	115.18
4	C	3001	NAD	C6N-N1N-C2N	-2.14	120.03	121.97
4	C	3001	NAD	C3N-C7N-N7N	-2.13	115.19	117.75
4	D	3001	NAD	PN-O3-PA	-2.06	125.77	132.83
4	C	3001	NAD	C2N-C3N-C4N	2.03	120.56	118.26

There are no chirality outliers.

All (33) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	D	1558	GOL	C1-C2-C3-O3
4	B	3001	NAD	O4D-C1D-N1N-C2N
4	B	3001	NAD	O4D-C1D-N1N-C6N
4	B	3001	NAD	C2D-C1D-N1N-C2N
4	B	3001	NAD	C2D-C1D-N1N-C6N
4	A	3001	NAD	O4D-C1D-N1N-C2N
4	A	3001	NAD	O4D-C1D-N1N-C6N
4	A	3001	NAD	C2D-C1D-N1N-C2N
4	A	3001	NAD	C2D-C1D-N1N-C6N
3	A	1560	GOL	O1-C1-C2-C3
3	A	1561	GOL	O1-C1-C2-C3
3	B	1563	GOL	O1-C1-C2-C3

*Continued on next page...*

*Continued from previous page...*

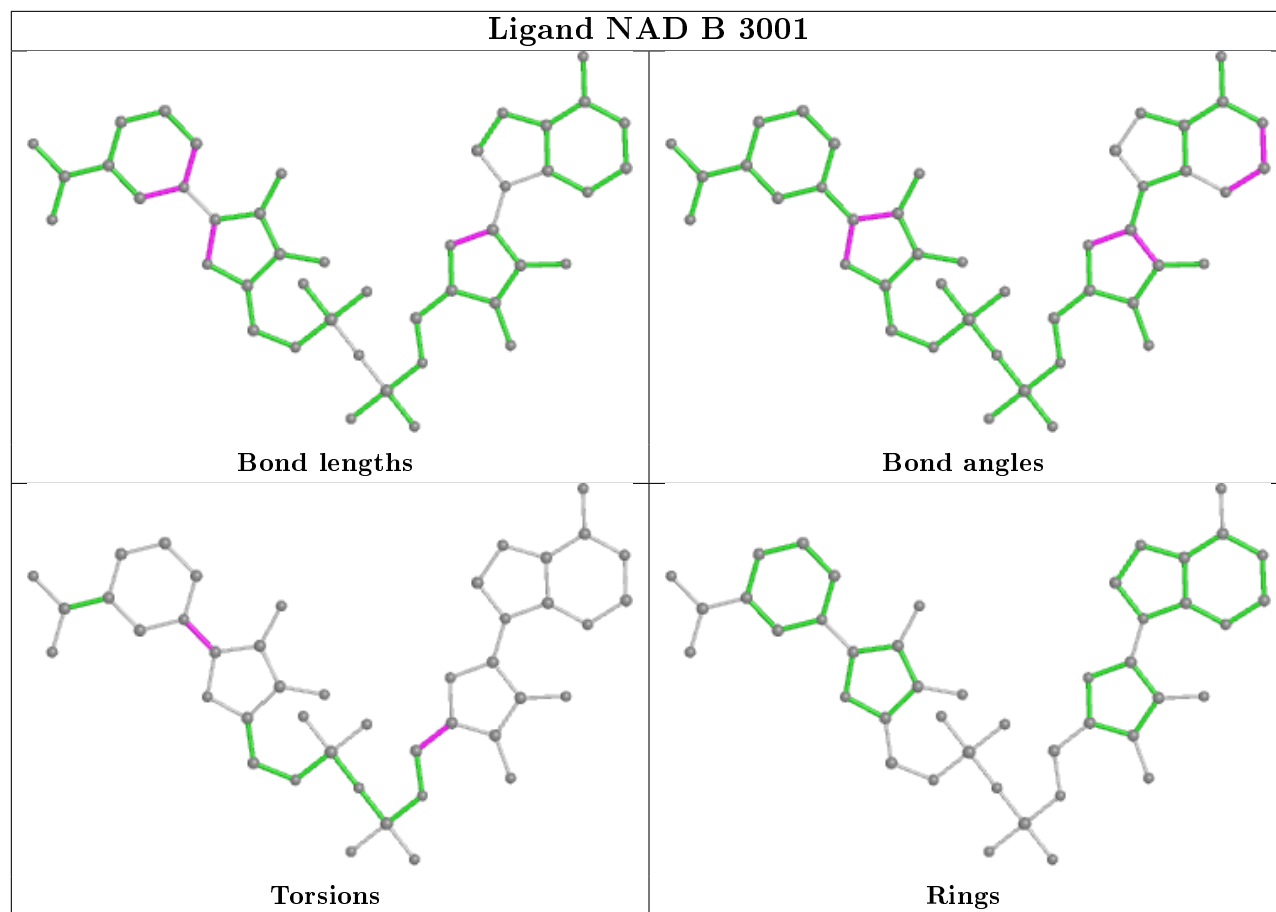
Mol	Chain	Res	Type	Atoms
3	D	1558	GOL	O2-C2-C3-O3
3	D	1560	GOL	O2-C2-C3-O3
3	A	1560	GOL	O1-C1-C2-O2
3	D	1559	GOL	O1-C1-C2-C3
3	D	1561	GOL	C1-C2-C3-O3
3	D	1560	GOL	O1-C1-C2-C3
3	D	1560	GOL	C1-C2-C3-O3
3	A	1561	GOL	O1-C1-C2-O2
3	B	1563	GOL	O1-C1-C2-O2
3	D	1559	GOL	O1-C1-C2-O2
3	A	1561	GOL	O2-C2-C3-O3
4	B	3001	NAD	O4B-C4B-C5B-O5B
3	B	1563	GOL	C1-C2-C3-O3
3	D	1561	GOL	O2-C2-C3-O3
3	D	1560	GOL	O1-C1-C2-O2
4	A	3001	NAD	O4B-C4B-C5B-O5B
4	D	3001	NAD	O4B-C4B-C5B-O5B
3	A	1561	GOL	C1-C2-C3-O3
3	B	1561	GOL	O1-C1-C2-C3
4	C	3001	NAD	C5B-O5B-PA-O1A
4	C	3001	NAD	O4B-C4B-C5B-O5B

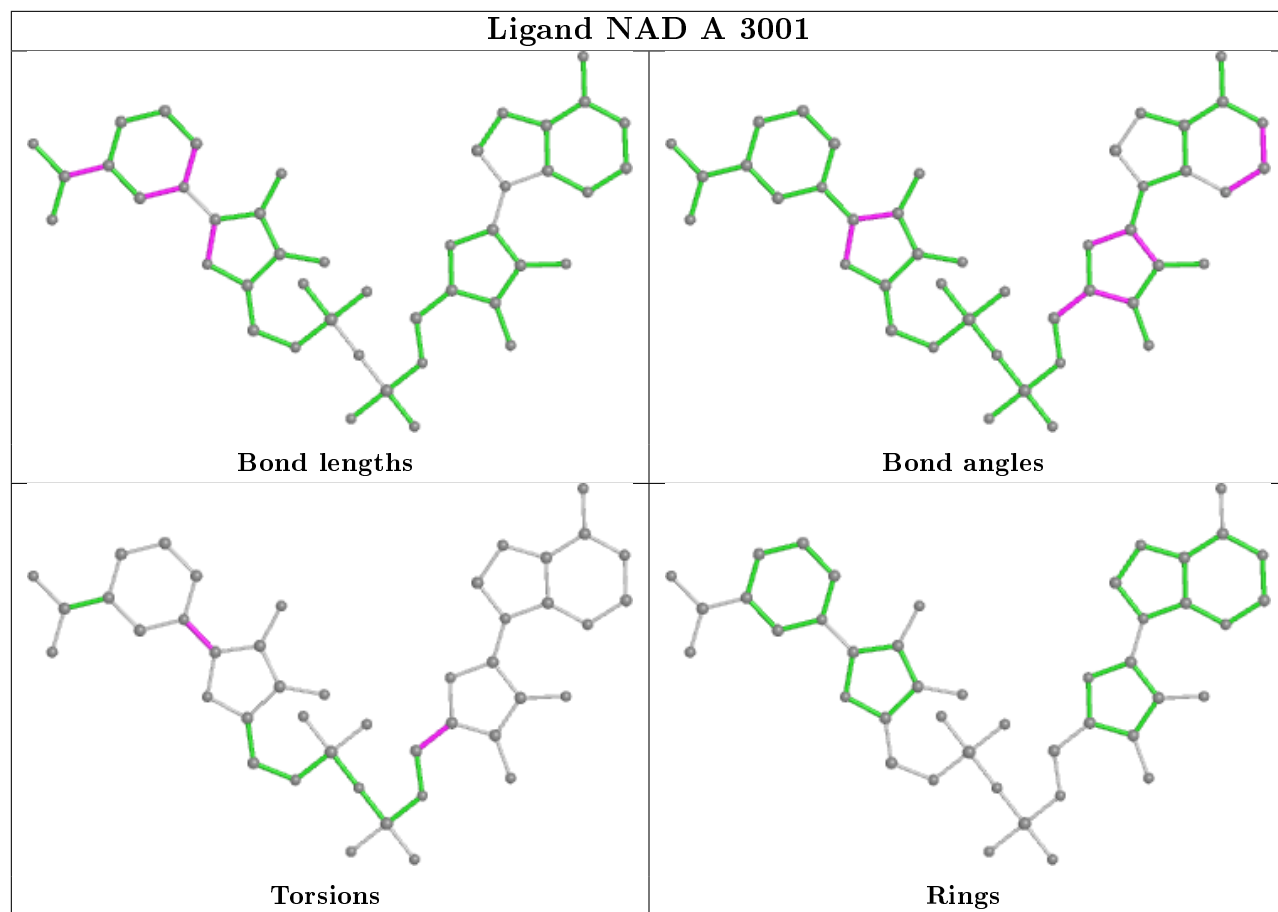
There are no ring outliers.

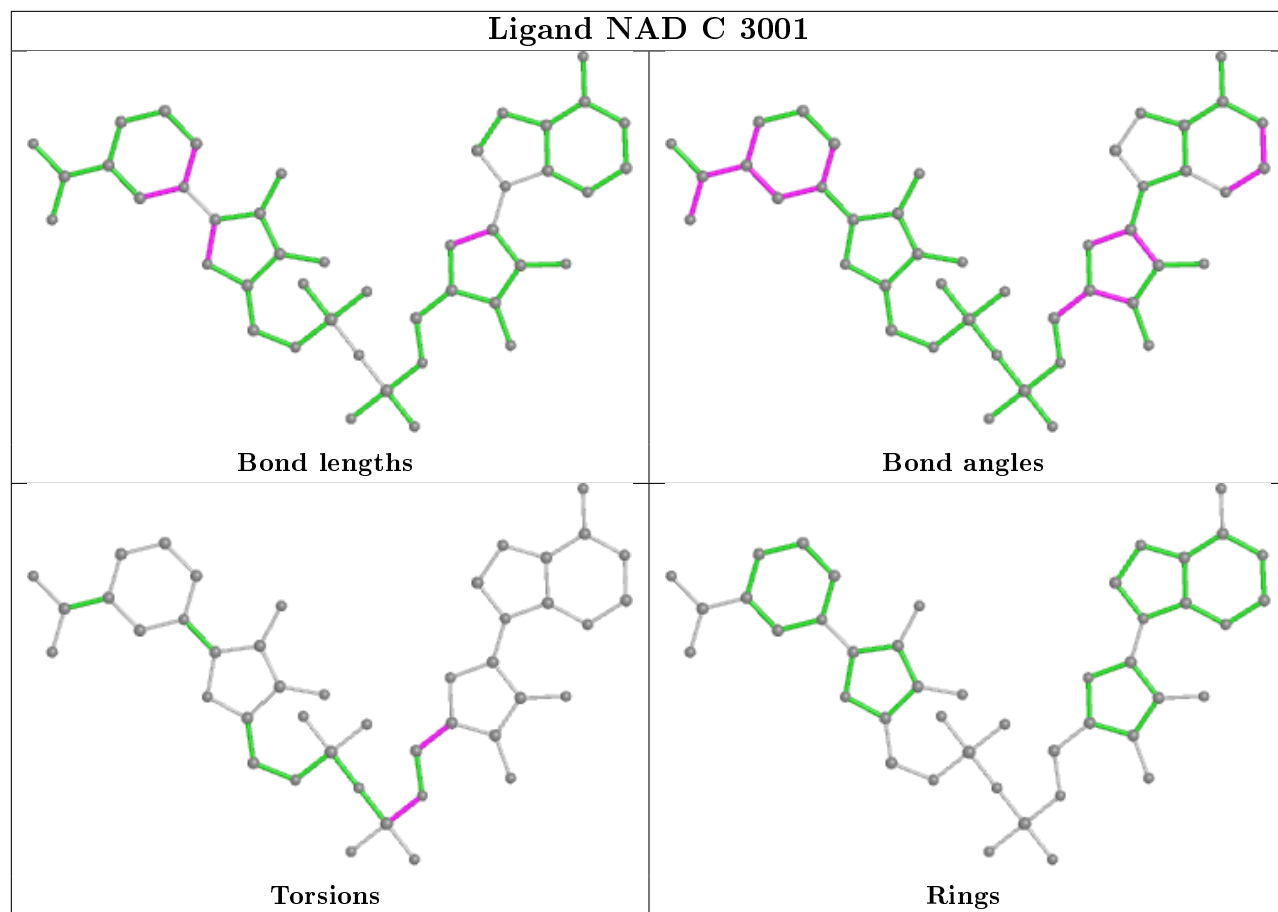
3 monomers are involved in 4 short contacts:

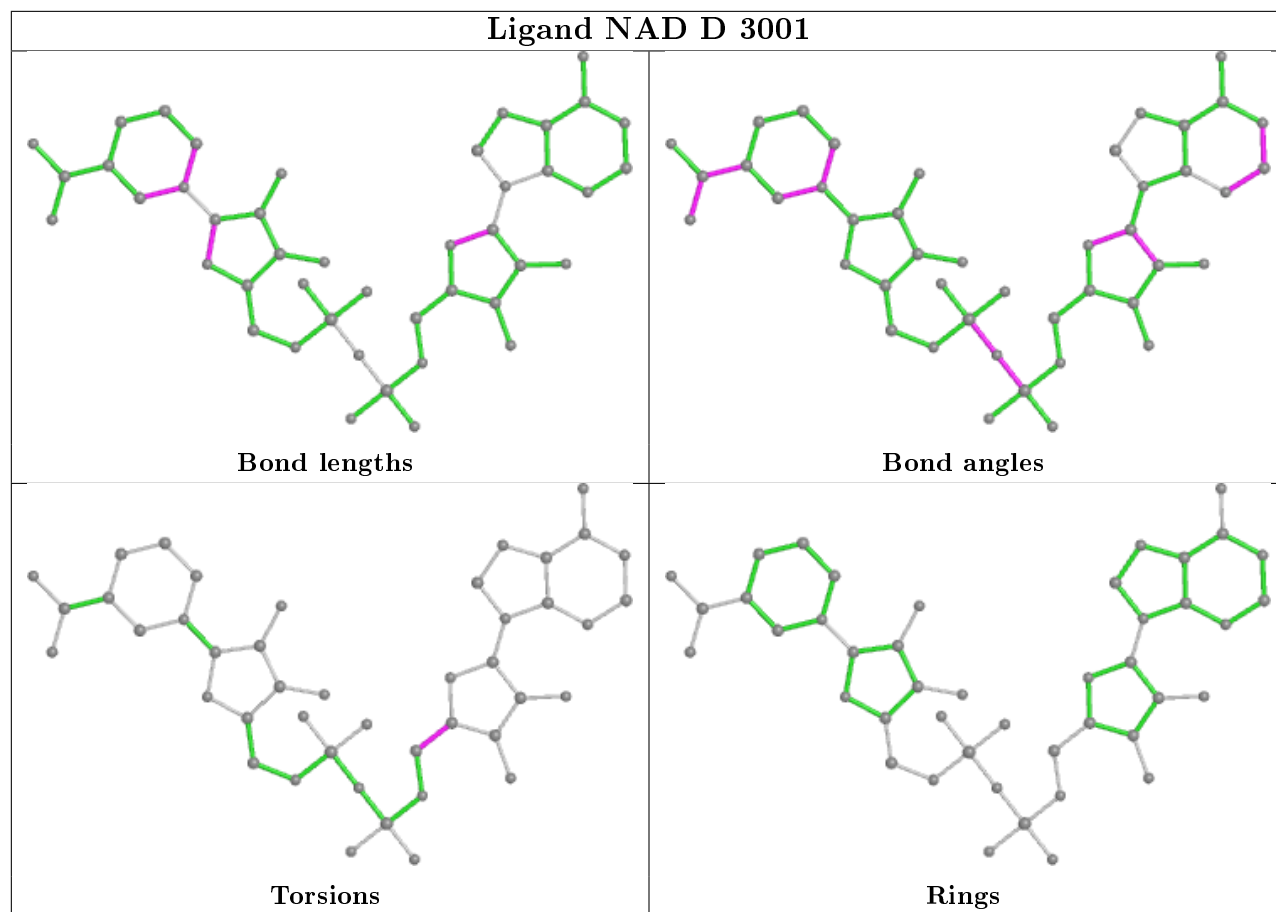
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	D	3001	NAD	2	0
3	A	1560	GOL	1	0
3	B	1563	GOL	1	0

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	553/557 (99%)	-0.09	2 (0%) 92 92	21, 25, 30, 40	0
1	B	553/557 (99%)	-0.10	2 (0%) 92 92	21, 25, 29, 35	0
1	C	553/557 (99%)	0.05	13 (2%) 59 57	19, 24, 32, 44	0
1	D	553/557 (99%)	0.12	21 (3%) 40 39	18, 24, 33, 40	0
All	All	2212/2228 (99%)	-0.01	38 (1%) 70 68	18, 24, 31, 44	0

All (38) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	386	ASP	4.8
1	D	337	ALA	4.6
1	D	430	ARG	3.6
1	D	505	VAL	3.3
1	D	5	ASN	3.1
1	D	386	ASP	3.1
1	D	285	ARG	3.1
1	C	337	ALA	2.9
1	A	5	ASN	2.9
1	B	282	GLU	2.9
1	C	399	GLU	2.7
1	D	210	ARG	2.7
1	C	10	VAL	2.6
1	D	289	GLN	2.6
1	C	282	GLU	2.6
1	D	371	GLU	2.5
1	D	217	THR	2.4
1	C	9	ASP	2.4
1	D	427	GLU	2.4
1	D	556	THR	2.3
1	D	205	PHE	2.3

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	388	ALA	2.3
1	C	289	GLN	2.3
1	A	7	TYR	2.2
1	D	431	SER	2.2
1	C	505	VAL	2.2
1	D	208	GLU	2.2
1	D	6	LYS	2.2
1	C	277	ILE	2.1
1	B	144	GLY	2.1
1	D	385	PRO	2.1
1	C	19	LYS	2.0
1	C	244	ALA	2.0
1	D	399	GLU	2.0
1	C	208	GLU	2.0
1	D	485	TRP	2.0
1	D	101	VAL	2.0
1	D	439	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
2	ACT	B	1558	4/4	0.79	0.28	53,54,54,54	0
3	GOL	B	1563	6/6	0.79	0.33	46,47,47,48	0
3	GOL	B	1561	6/6	0.81	0.30	44,46,46,48	0
2	ACT	B	1559	4/4	0.81	0.19	64,64,64,64	0
3	GOL	D	1559	6/6	0.82	0.20	58,59,59,59	0
3	GOL	A	1560	6/6	0.84	0.27	36,36,38,39	0

*Continued on next page...*

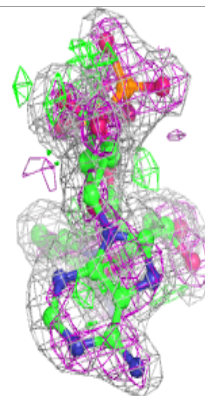
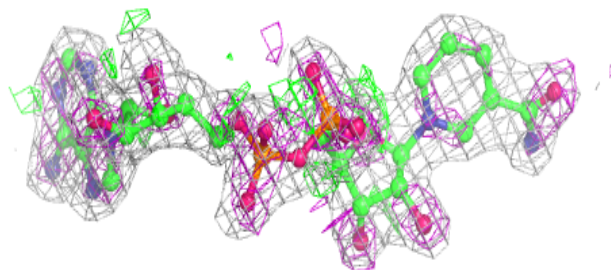
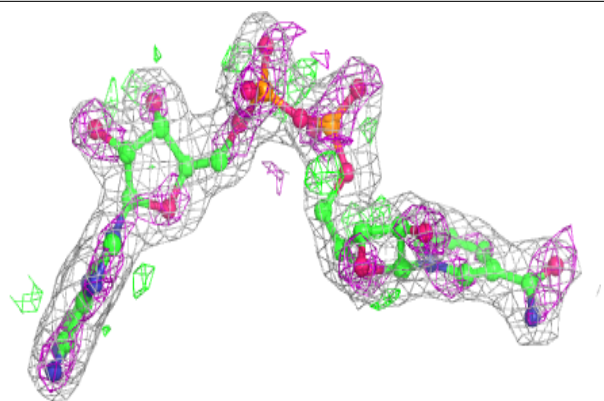
*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	GOL	D	1561	6/6	0.85	0.19	46,46,47,47	0
3	GOL	A	1561	6/6	0.86	0.27	42,43,43,43	0
3	GOL	B	1562	6/6	0.86	0.24	40,43,44,44	0
3	GOL	D	1558	6/6	0.91	0.21	37,39,40,40	0
3	GOL	D	1560	6/6	0.92	0.26	38,42,43,44	0
2	ACT	C	1558	4/4	0.93	0.14	33,33,34,34	0
2	ACT	A	1559	4/4	0.94	0.20	53,53,53,53	0
2	ACT	C	1559	4/4	0.96	0.11	24,24,24,24	0
4	NAD	C	3001	44/44	0.97	0.10	19,20,22,24	0
4	NAD	A	3001	44/44	0.98	0.09	14,15,20,22	0
4	NAD	D	3001	44/44	0.98	0.10	18,21,23,24	0
4	NAD	B	3001	44/44	0.98	0.09	13,16,21,25	0
2	ACT	A	1558	4/4	0.98	0.10	16,16,16,16	0
2	ACT	B	1560	4/4	0.99	0.08	17,17,17,18	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.

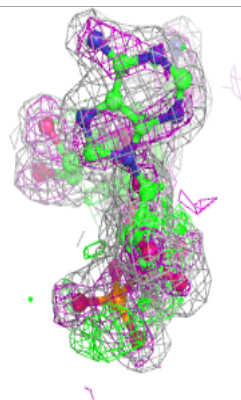
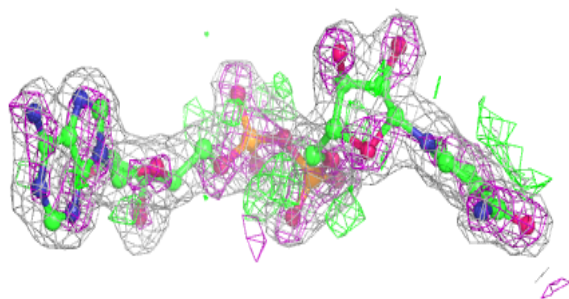
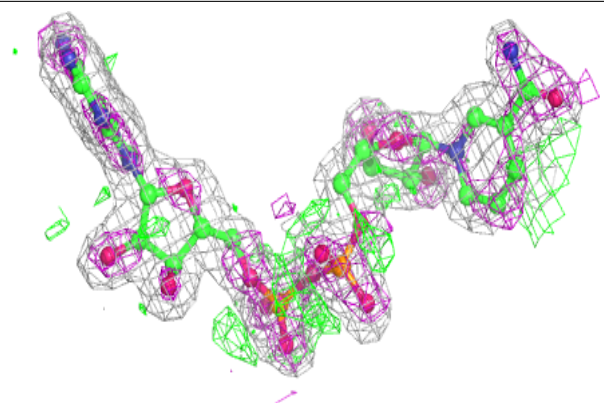
**Electron density around NAD C 3001:**

2mF<sub>o</sub>-DF<sub>c</sub> (at 0.7 rmsd) in gray  
mF<sub>o</sub>-DF<sub>c</sub> (at 3 rmsd) in purple (negative)  
and green (positive)

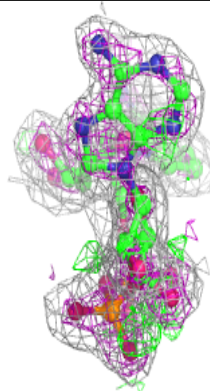
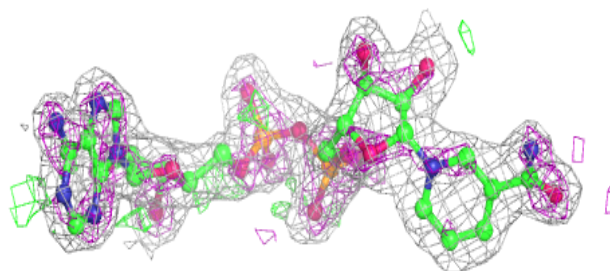
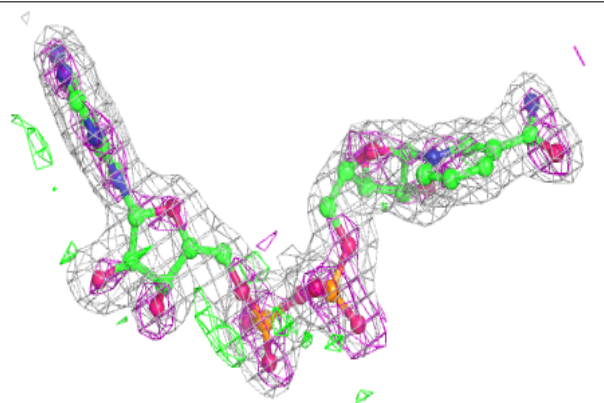


**Electron density around NAD A 3001:**

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

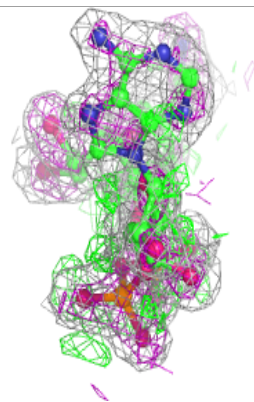
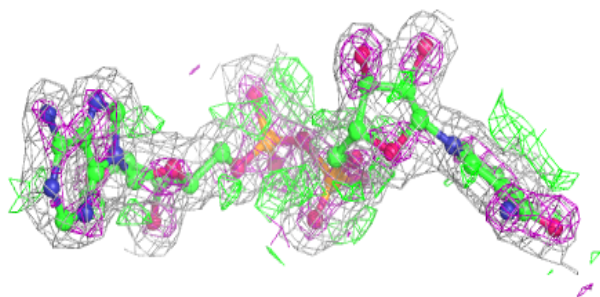
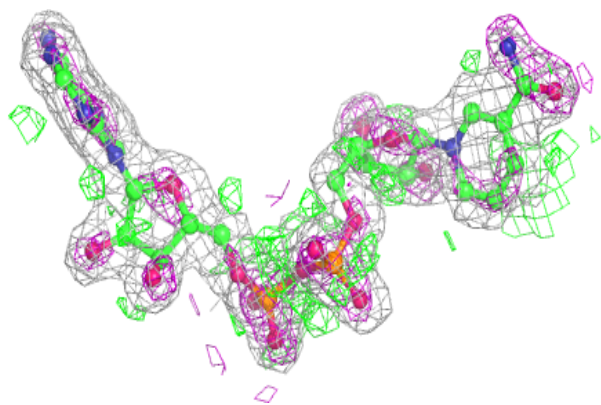
**Electron density around NAD D 3001:**

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



**Electron density around NAD B 3001:**

$2mF_o-DF_c$  (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.