



# Full wwPDB X-ray Structure Validation Report ⓘ

Jan 31, 2016 – 09:03 PM GMT

PDB ID : 1NAH  
Title : UDP-GALACTOSE 4-EPIMERASE FROM ESCHERICHIA COLI, REDUCED  
Authors : Thoden, J.B.; Frey, P.A.; Holden, H.M.  
Deposited on : 1995-11-22  
Resolution : 1.80 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  
<http://wwpdb.org/validation/2016/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 : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

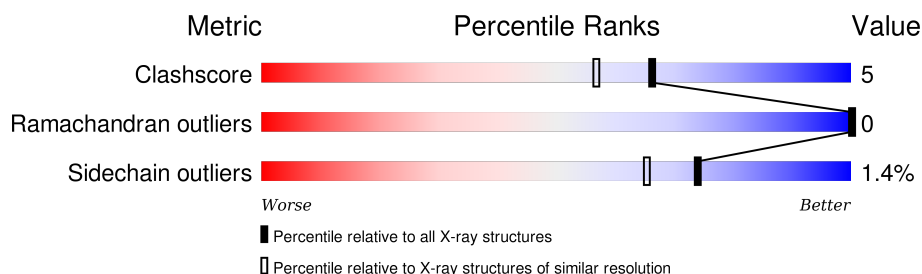
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.80 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))
Clashscore	102246	5383 (1.80-1.80)
Ramachandran outliers	100387	5320 (1.80-1.80)
Sidechain outliers	100360	5319 (1.80-1.80)

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. 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	338	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
6	PEG	A	344	-	-	X	-

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 3199 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 UDP-GALACTOSE 4-EPIMERASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	338	2625	1655	463	495	12	0	0	0

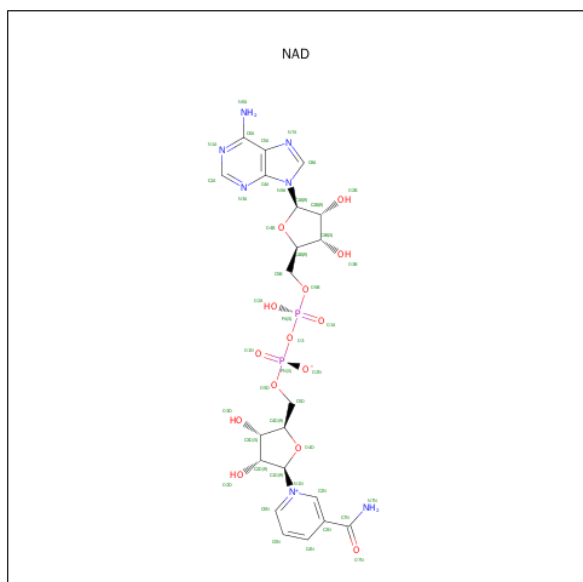
There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	131	ASN	GLN	CONFLICT	UNP P09147

- Molecule 2 is SODIUM ION (three-letter code: NA) (formula: Na).

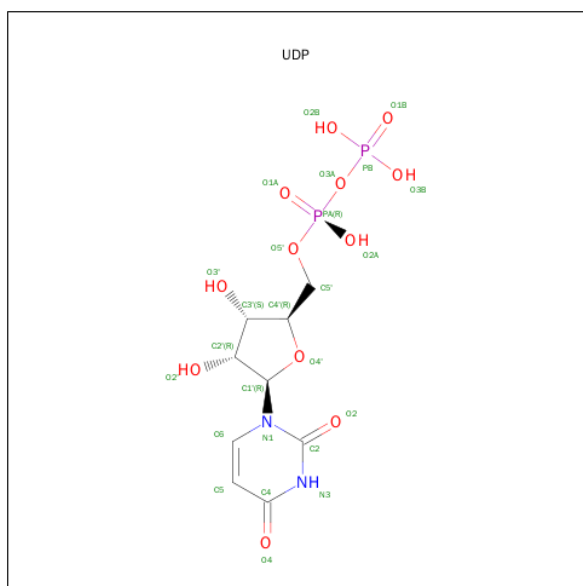
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	2	Total	Na	0	0
			2	2		

- 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		

- Molecule 4 is URIDINE-5'-DIPHOSPHATE (three-letter code: UDP) (formula:  $C_9H_{14}N_2O_{12}P_2$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	N	O	P	0	0
			25	9	2	12	2		

- Molecule 5 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula:  $C_2H_6O_2$ ).



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

- Molecule 6 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: C<sub>4</sub>H<sub>10</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	C	O	0	0
			7	4	3		

- Molecule 7 is water.

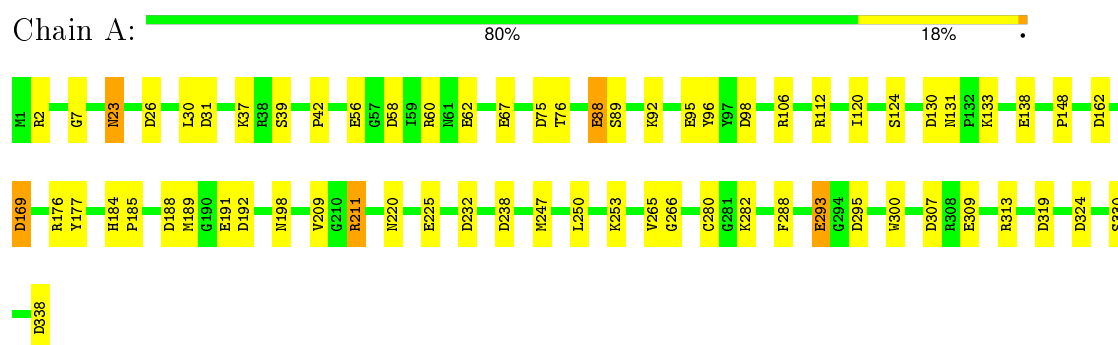
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	492	Total	O	0	0
			492	492		

### 3 Residue-property plots

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

Note EDS was not executed.

#### • Molecule 1: UDP-GALACTOSE 4-EPIMERASE



## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	83.50Å 83.50Å 108.40Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	30.00 – 1.80	Depositor
% Data completeness (in resolution range)	(Not available) (30.00-1.80)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	TNT	Depositor
R, $R_{free}$	0.165 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	3199	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	21.0	wwPDB-VP

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NA, UDP, PEG, EDO, 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	Bond lengths		Bond angles	
		RMSZ	$\# Z  > 5$	RMSZ	$\# Z  > 5$
1	A	1.06	12/2691 (0.4%)	1.36	37/3662 (1.0%)

All (12) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	88	GLU	CD-OE1	-8.85	1.16	1.25
1	A	138	GLU	CD-OE1	7.62	1.34	1.25
1	A	191	GLU	CD-OE1	7.22	1.33	1.25
1	A	225	GLU	CD-OE2	7.00	1.33	1.25
1	A	293	GLU	CD-OE1	6.87	1.33	1.25
1	A	309	GLU	CD-OE1	6.17	1.32	1.25
1	A	95	GLU	CD-OE1	5.97	1.32	1.25
1	A	62	GLU	CD-OE2	-5.95	1.19	1.25
1	A	67	GLU	CD-OE1	5.81	1.32	1.25
1	A	88	GLU	CD-OE2	5.73	1.31	1.25
1	A	62	GLU	CD-OE1	5.58	1.31	1.25
1	A	56	GLU	CD-OE1	-5.28	1.19	1.25

All (37) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	2	ARG	NE-CZ-NH1	10.41	125.50	120.30
1	A	295	ASP	CB-CG-OD2	-10.34	108.99	118.30
1	A	295	ASP	CB-CG-OD1	9.41	126.77	118.30
1	A	2	ARG	NE-CZ-NH2	-8.95	115.83	120.30
1	A	313	ARG	NE-CZ-NH1	8.58	124.59	120.30
1	A	319	ASP	CB-CG-OD1	7.74	125.27	118.30
1	A	319	ASP	CB-CG-OD2	-7.34	111.69	118.30
1	A	124	SER	N-CA-CB	-7.32	99.52	110.50
1	A	192	ASP	CB-CG-OD1	7.32	124.89	118.30

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	112	ARG	NE-CZ-NH2	-7.28	116.66	120.30
1	A	307	ASP	CB-CG-OD1	-7.23	111.79	118.30
1	A	131	ASN	N-CA-CB	7.13	123.43	110.60
1	A	177	TYR	CB-CG-CD1	7.12	125.27	121.00
1	A	177	TYR	CB-CG-CD2	-7.07	116.76	121.00
1	A	26	ASP	CB-CG-OD1	-7.02	111.98	118.30
1	A	211	ARG	NE-CZ-NH1	6.97	123.78	120.30
1	A	188	ASP	CB-CG-OD1	-6.71	112.26	118.30
1	A	338	ASP	CB-CG-OD1	-6.65	112.31	118.30
1	A	76	THR	CA-CB-CG2	-6.63	103.11	112.40
1	A	307	ASP	CB-CG-OD2	6.53	124.18	118.30
1	A	238	ASP	CB-CG-OD2	-6.46	112.48	118.30
1	A	58	ASP	CB-CG-OD2	-6.12	112.79	118.30
1	A	26	ASP	CB-CG-OD2	5.97	123.67	118.30
1	A	330	SER	CB-CA-C	5.88	121.27	110.10
1	A	169	ASP	CB-CG-OD1	-5.88	113.01	118.30
1	A	324	ASP	CB-CG-OD1	5.78	123.50	118.30
1	A	192	ASP	CB-CG-OD2	-5.77	113.11	118.30
1	A	176	ARG	NE-CZ-NH2	-5.72	117.44	120.30
1	A	75	ASP	CB-CG-OD2	5.70	123.43	118.30
1	A	106	ARG	NE-CZ-NH2	-5.64	117.48	120.30
1	A	58	ASP	CB-CG-OD1	5.56	123.31	118.30
1	A	162	ASP	CB-CG-OD2	5.47	123.22	118.30
1	A	96	TYR	CA-CB-CG	-5.37	103.19	113.40
1	A	60	ARG	NE-CZ-NH1	5.26	122.93	120.30
1	A	130	ASP	CB-CG-OD1	-5.19	113.63	118.30
1	A	313	ARG	NE-CZ-NH2	-5.15	117.72	120.30
1	A	98	ASP	CB-CG-OD2	5.14	122.93	118.30

There are no chirality outliers.

There are no planarity outliers.

## 5.2 Too-close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2625	0	2555	26	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	2	0	0	0	0
3	A	44	0	26	0	0
4	A	25	0	11	0	0
5	A	4	0	6	0	0
6	A	7	0	10	6	0
7	A	492	0	0	5	5
All	All	3199	0	2608	26	5

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 5.

All (26) 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:253:LYS:HZ3	6:A:344:PEG:H12	1.54	0.72
1:A:37:LYS:HE3	7:A:402:HOH:O	1.89	0.72
1:A:88:GLU:HG3	1:A:92:LYS:HE2	1.70	0.72
1:A:253:LYS:NZ	6:A:344:PEG:H12	2.05	0.70
1:A:211:ARG:HD2	7:A:568:HOH:O	1.99	0.61
1:A:184:HIS:CG	1:A:185:PRO:HD2	2.38	0.59
1:A:250:LEU:CD2	6:A:344:PEG:H32	2.34	0.58
1:A:250:LEU:HD21	6:A:344:PEG:H32	1.87	0.57
1:A:120:ILE:HD11	1:A:247:MET:HA	1.87	0.56
1:A:265:VAL:HG22	1:A:266:GLY:N	2.20	0.56
1:A:39:SER:O	1:A:42:PRO:HD2	2.07	0.55
1:A:23:ASN:ND2	7:A:376:HOH:O	2.40	0.54
1:A:7:GLY:HA2	1:A:31:ASP:OD2	2.09	0.52
1:A:253:LYS:HZ3	6:A:344:PEG:C1	2.21	0.52
1:A:220:ASN:HB3	1:A:288:PHE:CD1	2.48	0.48
1:A:232:ASP:HB2	1:A:300:TRP:HA	1.97	0.47
1:A:253:LYS:NZ	6:A:344:PEG:C1	2.78	0.46
1:A:30:LEU:HD13	1:A:30:LEU:C	2.36	0.46
1:A:133:LYS:NZ	7:A:747:HOH:O	2.49	0.45
1:A:89:SER:O	1:A:148:PRO:HG2	2.20	0.42
1:A:280:CYS:SG	1:A:282:LYS:HG2	2.60	0.42
1:A:169:ASP:N	1:A:169:ASP:OD1	2.52	0.42
1:A:88:GLU:CG	1:A:92:LYS:HE2	2.45	0.42
1:A:120:ILE:HD11	1:A:247:MET:CA	2.49	0.41
1:A:265:VAL:CG2	1:A:266:GLY:N	2.82	0.41
1:A:198:ASN:ND2	7:A:753:HOH:O	2.51	0.41

All (5) 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)
7:A:666:HOH:O	7:A:666:HOH:O[6_555]	0.39	1.81
7:A:665:HOH:O	7:A:665:HOH:O[6_555]	0.57	1.63
7:A:661:HOH:O	7:A:661:HOH:O[4_556]	0.63	1.57
7:A:664:HOH:O	7:A:664:HOH:O[5_555]	1.24	0.96
7:A:827:HOH:O	7:A:827:HOH:O[6_555]	1.55	0.65

## 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	336/338 (99%)	329 (98%)	7 (2%)	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	282/282 (100%)	278 (99%)	4 (1%)	74	65

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	23	ASN
1	A	189	MET

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	209	VAL
1	A	293	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	23	ASN
1	A	146	GLN
1	A	158	GLN
1	A	274	ASN
1	A	329	GLN

### 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](#)

Of 6 ligands modelled in this entry, 2 are monoatomic - leaving 4 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
3	NAD	A	341	-	38,48,48	2.13	5 (13%)	47,73,73	1.37	4 (8%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	UDP	A	342	-	18,26,26	1.98	4 (22%)	26,40,40	2.61	4 (15%)
5	EDO	A	343	-	3,3,3	0.80	0	2,2,2	0.50	0
6	PEG	A	344	-	6,6,6	1.16	0	5,5,5	1.67	1 (20%)

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	NAD	A	341	-	-	0/22/62/62	0/5/5/5
4	UDP	A	342	-	-	0/12/32/32	0/2/2/2
5	EDO	A	343	-	-	0/1/1/1	0/0/0/0
6	PEG	A	344	-	-	0/4/4/4	0/0/0/0

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	342	UDP	C6-C5	-2.51	1.32	1.38
3	A	341	NAD	C6N-C5N	-2.09	1.33	1.38
4	A	342	UDP	PB-O1B	2.42	1.59	1.51
3	A	341	NAD	C6N-N1N	3.35	1.44	1.35
4	A	342	UDP	C6-N1	4.15	1.41	1.35
4	A	342	UDP	C4-N3	5.11	1.42	1.33
3	A	341	NAD	C2N-C3N	6.19	1.48	1.39
3	A	341	NAD	C5N-C4N	6.46	1.52	1.38
3	A	341	NAD	C4N-C3N	7.10	1.51	1.39

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	341	NAD	C5N-C4N-C3N	-5.57	113.34	120.33
4	A	342	UDP	C5-C4-N3	-2.76	116.04	123.12
3	A	341	NAD	C3N-C2N-N1N	2.95	123.76	120.36
4	A	342	UDP	C4'-O4'-C1'	3.03	113.04	109.72
3	A	341	NAD	C3N-C7N-N7N	3.08	121.19	117.82
4	A	342	UDP	O3A-PA-O5'	3.39	111.94	102.94
3	A	341	NAD	C6N-C5N-C4N	3.46	124.67	119.44
6	A	344	PEG	C3-O2-C2	3.50	128.37	113.31
4	A	342	UDP	C4-N3-C2	11.63	125.66	114.14

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	344	PEG	6	0

## 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](#)

EDS was not executed - this section will therefore be empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

EDS was not executed - this section will therefore be empty.

### 6.3 Carbohydrates [i](#)

EDS was not executed - this section will therefore be empty.

### 6.4 Ligands [i](#)

EDS was not executed - this section will therefore be empty.

### 6.5 Other polymers [i](#)

EDS was not executed - this section will therefore be empty.