



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 19, 2018 – 10:19 pm GMT

PDB ID : 2C20  
Title : CRYSTAL STRUCTURE OF UDP-GLUCOSE 4-EPIMERASE  
Authors : Lebedev, A.A.; Moroz, O.V.; Blagova, E.V.; Levnikov, V.M.; Fogg, M.J.;  
Brannigan, J.A.; Wilkinson, A.J.; Wilson, K.S.  
Deposited on : 2005-09-22  
Resolution : 2.70 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 : 1.7.3 (157068), CSD as539be (2018)  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk30686

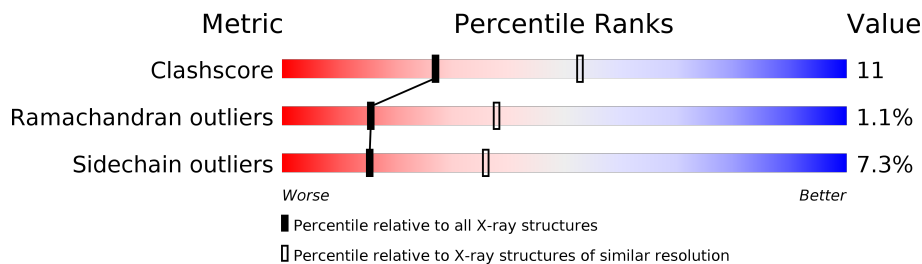
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.70 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	122078	2755 (2.70-2.70)
Ramachandran outliers	120005	2715 (2.70-2.70)
Sidechain outliers	119972	2715 (2.70-2.70)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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	330	<div> <div>70%</div> <div>28%</div> <div>.</div> </div>
1	B	330	<div> <div>71%</div> <div>25%</div> <div>.</div> </div>
1	C	330	<div> <div>77%</div> <div>21%</div> <div>..</div> </div>
1	D	330	<div> <div>70%</div> <div>26%</div> <div>.</div> </div>
1	E	330	<div> <div>74%</div> <div>23%</div> <div>.</div> </div>
1	F	330	<div> <div>75%</div> <div>21%</div> <div>.</div> </div>

## 2 Entry composition [i](#)

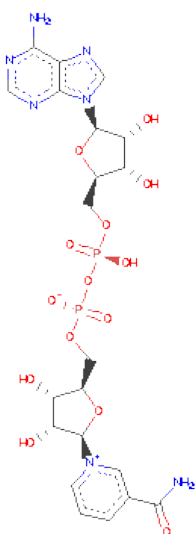
There are 4 unique types of molecules in this entry. The entry contains 16074 atoms, of which 0 are hydrogens and 0 are deuteriums.

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-GLUCOSE 4-EPIMERASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	329	Total	C	N	O	S	0	1	0
			2595	1645	440	499	11			
1	B	329	Total	C	N	O	S	0	1	0
			2595	1645	440	499	11			
1	C	329	Total	C	N	O	S	0	1	0
			2595	1645	440	499	11			
1	D	329	Total	C	N	O	S	0	1	0
			2595	1645	440	499	11			
1	E	329	Total	C	N	O	S	0	1	0
			2595	1645	440	499	11			
1	F	329	Total	C	N	O	S	0	1	0
			2595	1645	440	499	11			

- Molecule 2 is NICOTINAMIDE-ADENINE-DINUCLEOTIDE (three-letter code: NAD) (formula:  $C_{21}H_{27}N_7O_{14}P_2$ ).



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

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	D	1	Total	Zn	0	0
			1	1		
3	E	1	Total	Zn	0	0
			1	1		
3	B	1	Total	Zn	0	0
			1	1		
3	C	1	Total	Zn	0	0
			1	1		
3	A	1	Total	Zn	0	0
			1	1		
3	F	1	Total	Zn	0	0
			1	1		

- Molecule 4 is water.

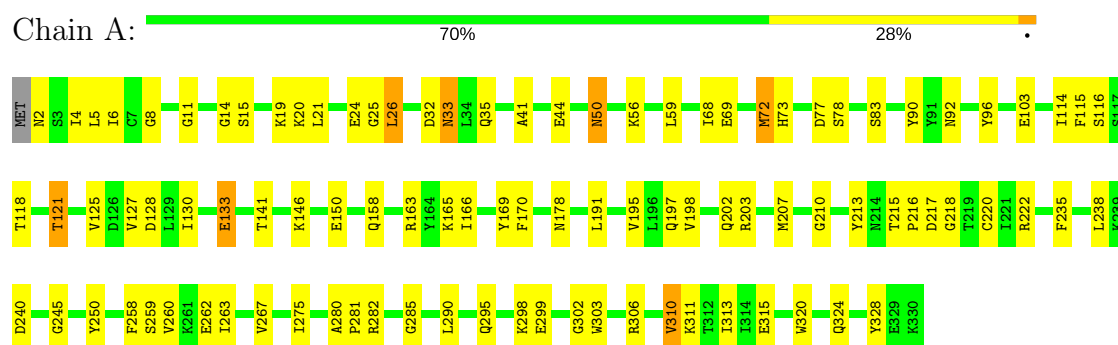
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	41	Total	O	0	0
			41	41		
4	B	41	Total	O	0	0
			41	41		
4	C	38	Total	O	0	0
			38	38		
4	D	36	Total	O	0	0
			36	36		
4	E	36	Total	O	0	0
			36	36		
4	F	42	Total	O	0	0
			42	42		

### 3 Residue-property plots

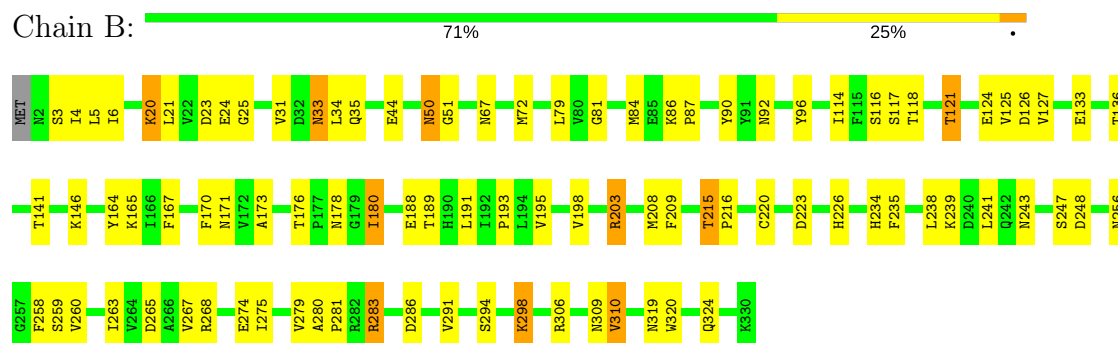
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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.

Note EDS was not executed.

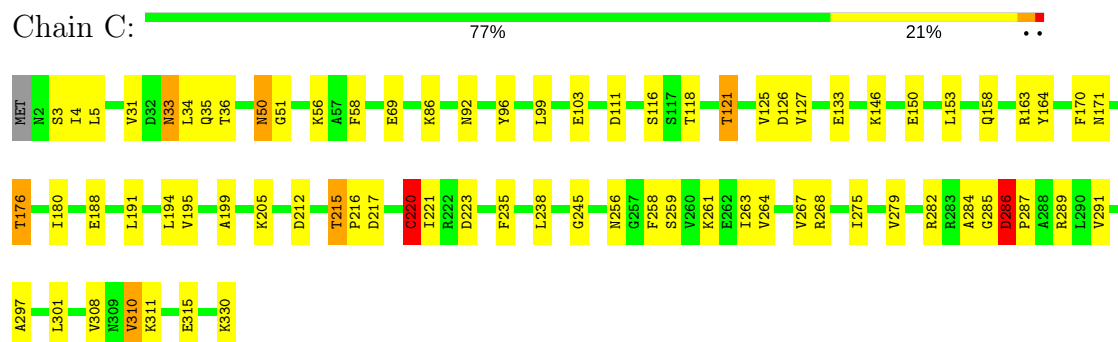
#### • Molecule 1: UDP-GLUCOSE 4-EPIMERASE



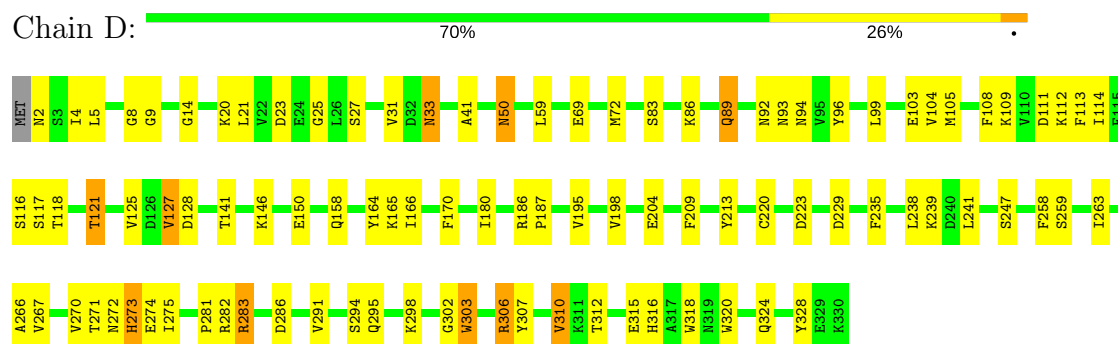
#### • Molecule 1: UDP-GLUCOSE 4-EPIMERASE



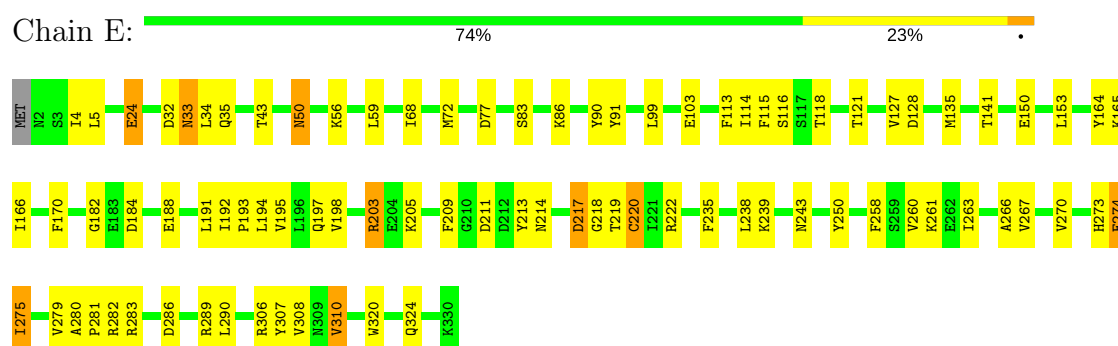
#### • Molecule 1: UDP-GLUCOSE 4-EPIMERASE



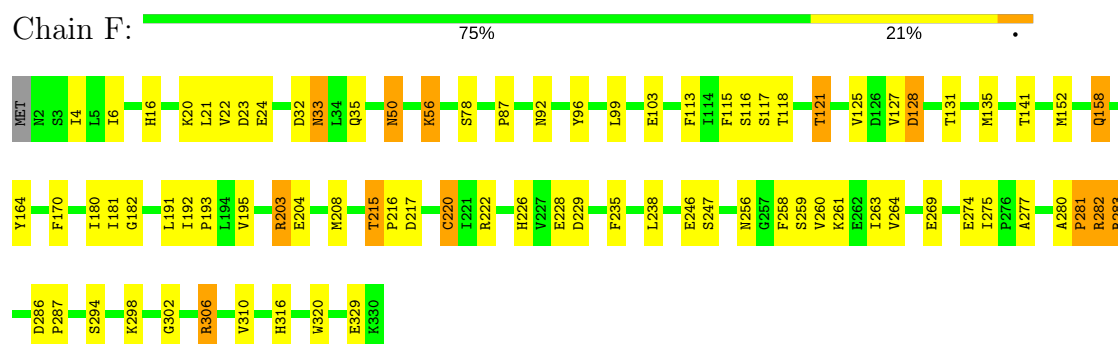
• Molecule 1: UDP-GLUCOSE 4-EPIMERASE



• Molecule 1: UDP-GLUCOSE 4-EPIMERASE



• Molecule 1: UDP-GLUCOSE 4-EPIMERASE



## 4 Data and refinement statistics

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

Property	Value	Source
Space group	P 32 1 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	136.16 Å   136.16 Å   248.65 Å 90.00°   90.00°   120.00°	Depositor
Resolution (Å)	25.02 – 2.70	Depositor
% Data completeness (in resolution range)	73.1 (25.02-2.70)	Depositor
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.205 , 0.276	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	16074	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	16.0	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: ZN, 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	0.67	1/2656 (0.0%)	0.72	0/3600
1	B	0.71	1/2656 (0.0%)	0.73	0/3600
1	C	0.69	1/2656 (0.0%)	0.72	0/3600
1	D	0.63	0/2656	0.71	0/3600
1	E	1.15	2/2656 (0.1%)	0.73	2/3600 (0.1%)
1	F	0.68	1/2656 (0.0%)	0.71	0/3600
All	All	0.78	6/15936 (0.0%)	0.72	2/21600 (0.0%)

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	274	GLU	CD-OE2	36.08	1.65	1.25
1	E	274	GLU	CD-OE1	30.32	1.59	1.25
1	B	44	GLU	CD-OE2	6.86	1.33	1.25
1	F	220	CYS	CB-SG	-6.49	1.71	1.82
1	C	220	CYS	CB-SG	-6.08	1.72	1.82

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	274	GLU	OE1-CD-OE2	9.22	134.37	123.30
1	E	184	ASP	CB-CG-OD1	5.31	123.08	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	2595	0	2513	62	0
1	B	2595	0	2513	57	0
1	C	2595	0	2513	52	0
1	D	2595	0	2513	59	0
1	E	2595	0	2513	48	0
1	F	2595	0	2513	57	0
2	A	44	0	26	4	0
2	B	44	0	26	1	0
2	C	44	0	26	2	0
2	D	44	0	26	1	0
2	E	44	0	26	2	0
2	F	44	0	26	3	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
3	E	1	0	0	0	0
3	F	1	0	0	0	0
4	A	41	0	0	0	0
4	B	41	0	0	0	0
4	C	38	0	0	1	0
4	D	36	0	0	0	0
4	E	36	0	0	1	0
4	F	42	0	0	1	0
All	All	16074	0	15234	337	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 11.

The worst 5 of 337 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:E:274:GLU:CD	1:E:274:GLU:OE2	1.65	1.35
1:C:212:ASP:C	1:C:282:ARG:HH22	1.59	1.03
1:F:158:GLN:HA	1:F:158:GLN:HE21	1.35	0.92

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:280:ALA:HB1	1:E:281:PRO:HD2	1.53	0.91
1:F:158:GLN:CA	1:F:158:GLN:HE21	1.84	0.90

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.

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	328/330 (99%)	297 (90%)	27 (8%)	4 (1%)	14	35
1	B	328/330 (99%)	307 (94%)	19 (6%)	2 (1%)	27	54
1	C	328/330 (99%)	309 (94%)	17 (5%)	2 (1%)	27	54
1	D	328/330 (99%)	302 (92%)	22 (7%)	4 (1%)	14	35
1	E	328/330 (99%)	291 (89%)	33 (10%)	4 (1%)	14	35
1	F	328/330 (99%)	301 (92%)	22 (7%)	5 (2%)	11	29
All	All	1968/1980 (99%)	1807 (92%)	140 (7%)	21 (1%)	16	38

5 of 21 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	286	ASP
1	A	25	GLY
1	D	272	ASN
1	F	281	PRO
1	F	282	ARG

### 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	276/277 (100%)	258 (94%)	18 (6%)	19	42
1	B	276/277 (100%)	250 (91%)	26 (9%)	9	22
1	C	276/277 (100%)	256 (93%)	20 (7%)	16	37
1	D	276/277 (100%)	258 (94%)	18 (6%)	19	42
1	E	276/277 (100%)	259 (94%)	17 (6%)	20	44
1	F	276/277 (100%)	254 (92%)	22 (8%)	13	30
All	All	1656/1662 (100%)	1535 (93%)	121 (7%)	15	36

5 of 121 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	205	LYS
1	D	50	ASN
1	F	203	ARG
1	C	220	CYS
1	C	308	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 32 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	65	GLN
1	D	33	ASN
1	F	158	GLN
1	C	161	ASN
1	D	50	ASN

### 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 [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 12 ligands modelled in this entry, 6 are monoatomic - leaving 6 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
2	NAD	A	401	1	40,48,48	1.83	3 (7%)	44,73,73	1.79	3 (6%)
2	NAD	B	401	-	40,48,48	1.78	3 (7%)	44,73,73	1.96	3 (6%)
2	NAD	C	401	-	40,48,48	1.79	3 (7%)	44,73,73	1.77	3 (6%)
2	NAD	D	401	-	40,48,48	1.83	3 (7%)	44,73,73	2.04	1 (2%)
2	NAD	E	401	-	40,48,48	1.78	3 (7%)	44,73,73	1.96	5 (11%)
2	NAD	F	401	-	40,48,48	1.73	3 (7%)	44,73,73	1.80	4 (9%)

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	A	401	1	-	0/22/62/62	0/5/5/5
2	NAD	B	401	-	-	0/22/62/62	0/5/5/5
2	NAD	C	401	-	-	0/22/62/62	0/5/5/5
2	NAD	D	401	-	-	0/22/62/62	0/5/5/5
2	NAD	E	401	-	-	0/22/62/62	0/5/5/5
2	NAD	F	401	-	-	0/22/62/62	0/5/5/5

The worst 5 of 18 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	401	NAD	C2A-N1A	2.17	1.38	1.33
2	C	401	NAD	C2A-N1A	2.21	1.38	1.33

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	401	NAD	C2A-N1A	2.28	1.38	1.33
2	F	401	NAD	C2A-N1A	2.34	1.38	1.33
2	D	401	NAD	C2A-N1A	2.35	1.38	1.33

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	401	NAD	N3A-C2A-N1A	-12.37	118.27	128.86
2	B	401	NAD	N3A-C2A-N1A	-11.00	119.44	128.86
2	E	401	NAD	N3A-C2A-N1A	-10.46	119.91	128.86
2	F	401	NAD	N3A-C2A-N1A	-10.20	120.13	128.86
2	A	401	NAD	N3A-C2A-N1A	-10.05	120.26	128.86

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6 monomers are involved in 13 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	401	NAD	4	0
2	B	401	NAD	1	0
2	C	401	NAD	2	0
2	D	401	NAD	1	0
2	E	401	NAD	2	0
2	F	401	NAD	3	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

### 6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

### 6.4 Ligands

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers

EDS was not executed - this section is therefore empty.