



# wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 19, 2018 – 05:00 PM EST

PDB ID : 6C4J  
Title : Ligand bound full length hUGDH with A104L substitution  
Authors : Beattie, N.R.; Pioso, B.J.; Wood, Z.A.; Sidlo, A.M.  
Deposited on : 2018-01-12  
Resolution : 2.53 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) : 1.13  
EDS : rb-20031633  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Refmac : 5.8.0158  
CCP4 : 7.0 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20031633

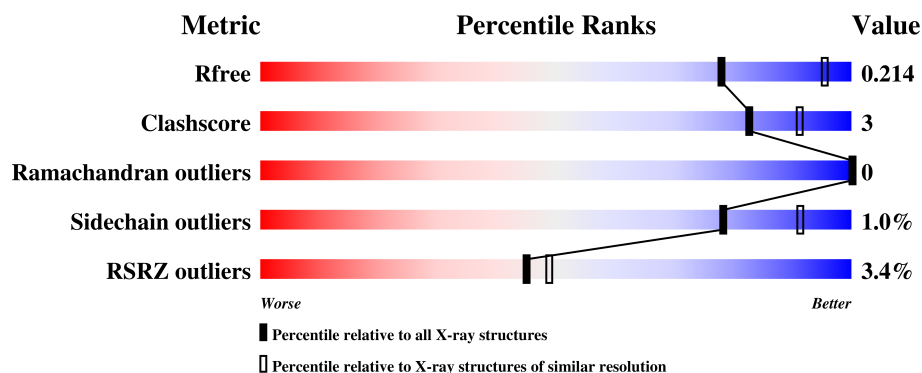
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.53 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	5045 (2.54-2.50)
Clashscore	122126	5751 (2.54-2.50)
Ramachandran outliers	120053	5650 (2.54-2.50)
Sidechain outliers	120020	5652 (2.54-2.50)
RSRZ outliers	108989	4938 (2.54-2.50)

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.

Mol	Chain	Length	Quality of chain
1	A	494	<div> <div>7%</div> <div>80% 5% 14%</div> </div>
1	B	494	<div> <div>6%</div> <div>86% 7% 7%</div> </div>
1	C	494	<div> <div>2%</div> <div>87% 6% 7%</div> </div>
1	D	494	<div> <div>88% 5% 7%</div> </div>
1	E	494	<div> <div>86% 7% 7%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	494	
1	G	494	
1	H	494	
1	I	494	
1	J	494	
1	K	494	
1	L	494	

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
5	SO4	C	503	-	-	X	-

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 44722 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 6-dehydrogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	426	Total	C	N	O	S	0	3	0
			3347	2117	580	633	17			
1	B	460	Total	C	N	O	S	0	0	0
			3609	2285	623	681	20			
1	C	460	Total	C	N	O	S	0	2	0
			3618	2291	623	684	20			
1	D	460	Total	C	N	O	S	0	1	0
			3612	2287	623	682	20			
1	E	460	Total	C	N	O	S	0	2	0
			3617	2290	624	683	20			
1	F	460	Total	C	N	O	S	0	1	0
			3612	2287	623	682	20			
1	G	460	Total	C	N	O	S	0	3	0
			3623	2293	627	683	20			
1	H	459	Total	C	N	O	S	0	0	0
			3601	2280	622	680	19			
1	I	460	Total	C	N	O	S	0	3	0
			3623	2294	624	685	20			
1	J	458	Total	C	N	O	S	0	2	0
			3598	2279	619	681	19			
1	K	460	Total	C	N	O	S	0	1	0
			3612	2287	623	682	20			
1	L	460	Total	C	N	O	S	0	1	0
			3612	2287	623	682	20			

There are 12 discrepancies between the modelled and reference sequences:

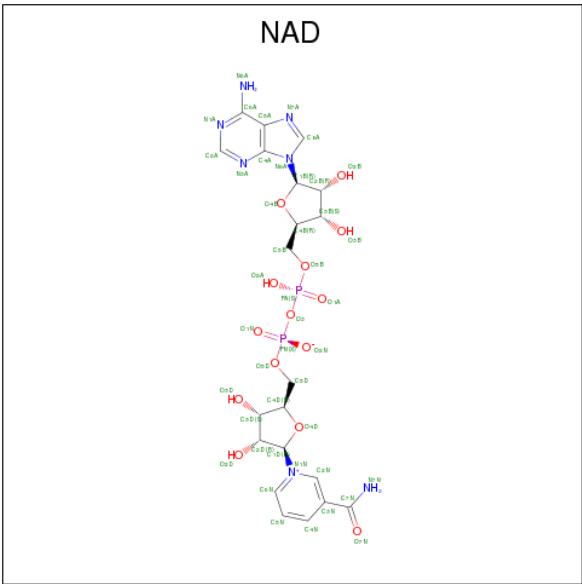
Chain	Residue	Modelled	Actual	Comment	Reference
A	104	LEU	ALA	engineered mutation	UNP O60701
B	104	LEU	ALA	engineered mutation	UNP O60701
C	104	LEU	ALA	engineered mutation	UNP O60701
D	104	LEU	ALA	engineered mutation	UNP O60701
E	104	LEU	ALA	engineered mutation	UNP O60701

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	104	LEU	ALA	engineered mutation	UNP O60701
G	104	LEU	ALA	engineered mutation	UNP O60701
H	104	LEU	ALA	engineered mutation	UNP O60701
I	104	LEU	ALA	engineered mutation	UNP O60701
J	104	LEU	ALA	engineered mutation	UNP O60701
K	104	LEU	ALA	engineered mutation	UNP O60701
L	104	LEU	ALA	engineered mutation	UNP O60701

- Molecule 2 is NICOTINAMIDE-ADENINE-DINUCLEOTIDE (three-letter code: NAD) (formula: C<sub>21</sub>H<sub>27</sub>N<sub>7</sub>O<sub>14</sub>P<sub>2</sub>).



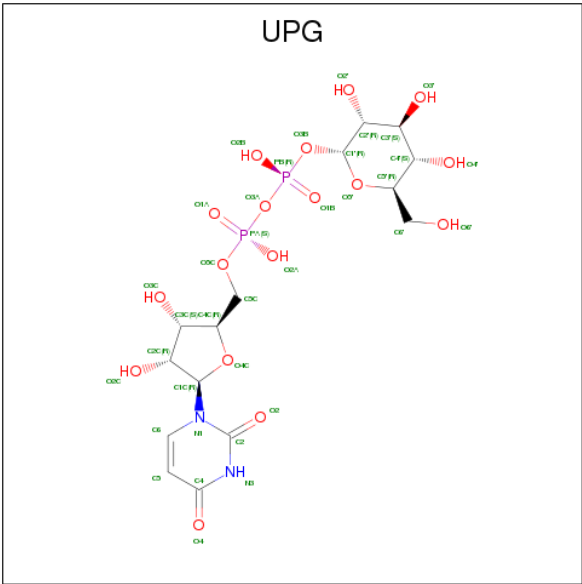
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			44	21	7	14	2		
2	B	1	Total	C	N	O	P	0	0
			44	21	7	14	2		
2	C	1	Total	C	N	O	P	0	0
			44	21	7	14	2		
2	D	1	Total	C	N	O	P	0	0
			44	21	7	14	2		
2	E	1	Total	C	N	O	P	0	0
			44	21	7	14	2		
2	F	1	Total	C	N	O	P	0	0
			44	21	7	14	2		
2	G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

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	H	1	Total	C	N	O	P	0	0
			44	21	7	14	2		
2	I	1	Total	C	N	O	P	0	0
			44	21	7	14	2		
2	J	1	Total	C	N	O	P	0	0
			44	21	7	14	2		
2	K	1	Total	C	N	O	P	0	0
			44	21	7	14	2		
2	L	1	Total	C	N	O	P	0	0
			44	21	7	14	2		

- Molecule 3 is URIDINE-5'-DIPHOSPHATE-GLUCOSE (three-letter code: UPG) (formula: C<sub>15</sub>H<sub>24</sub>N<sub>2</sub>O<sub>17</sub>P<sub>2</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			36	15	2	17	2		
3	B	1	Total	C	N	O	P	0	0
			36	15	2	17	2		
3	C	1	Total	C	N	O	P	0	0
			36	15	2	17	2		
3	D	1	Total	C	N	O	P	0	0
			36	15	2	17	2		
3	E	1	Total	C	N	O	P	0	0
			36	15	2	17	2		
3	F	1	Total	C	N	O	P	0	0
			36	15	2	17	2		

Continued on next page...

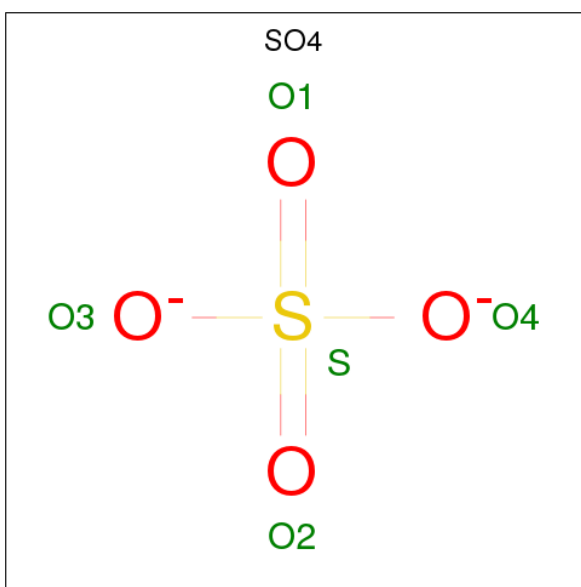
*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	G	1	Total	C	N	O	P	0	0
			36	15	2	17	2		
3	H	1	Total	C	N	O	P	0	0
			36	15	2	17	2		
3	I	1	Total	C	N	O	P	0	0
			36	15	2	17	2		
3	J	1	Total	C	N	O	P	0	0
			36	15	2	17	2		
3	K	1	Total	C	N	O	P	0	0
			36	15	2	17	2		
3	L	1	Total	C	N	O	P	0	0
			36	15	2	17	2		

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	G	1	Total	Cl	0	0
			1	1		
4	J	1	Total	Cl	0	0
			1	1		
4	D	1	Total	Cl	0	0
			1	1		
4	K	2	Total	Cl	0	0
			2	2		
4	H	1	Total	Cl	0	0
			1	1		
4	B	2	Total	Cl	0	0
			2	2		
4	I	2	Total	Cl	0	0
			2	2		
4	A	1	Total	Cl	0	0
			1	1		
4	L	2	Total	Cl	0	0
			2	2		

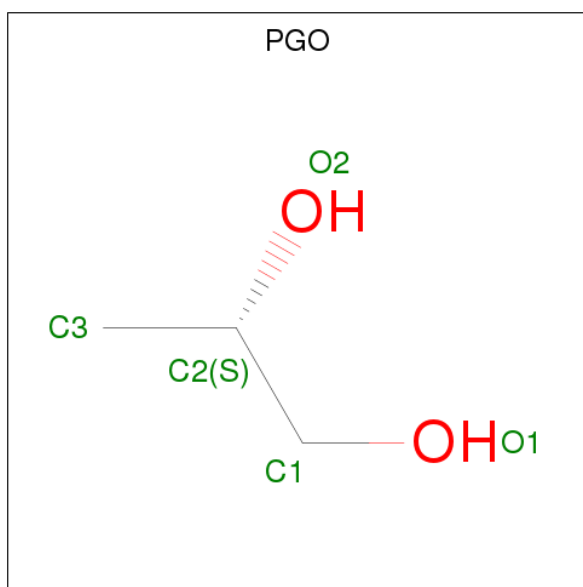
- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	C	1	Total	O	S	0	0
			5	4	1		
5	D	1	Total	O	S	0	0
			5	4	1		
5	D	1	Total	O	S	0	0
			5	4	1		
5	E	1	Total	O	S	0	0
			5	4	1		
5	F	1	Total	O	S	0	0
			5	4	1		
5	F	1	Total	O	S	0	0
			5	4	1		
5	G	1	Total	O	S	0	0
			5	4	1		
5	K	1	Total	O	S	0	0
			5	4	1		

- Molecule 6 is S-1,2-PROPANEDIOL (three-letter code: PGO) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>2</sub>).





Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	D	1	Total	C	O	0	0
			5	3	2		
6	E	1	Total	C	O	0	0
			5	3	2		
6	F	1	Total	C	O	0	0
			5	3	2		
6	G	1	Total	C	O	0	0
			5	3	2		
6	K	1	Total	C	O	0	0
			5	3	2		
6	L	1	Total	C	O	0	0
			5	3	2		

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	33	Total	O	0	0
			33	33		
7	B	29	Total	O	0	0
			29	29		
7	C	53	Total	O	0	0
			53	53		
7	D	64	Total	O	0	0
			64	64		
7	E	85	Total	O	0	0
			85	85		
7	F	66	Total	O	0	0
			66	66		

*Continued on next page...*

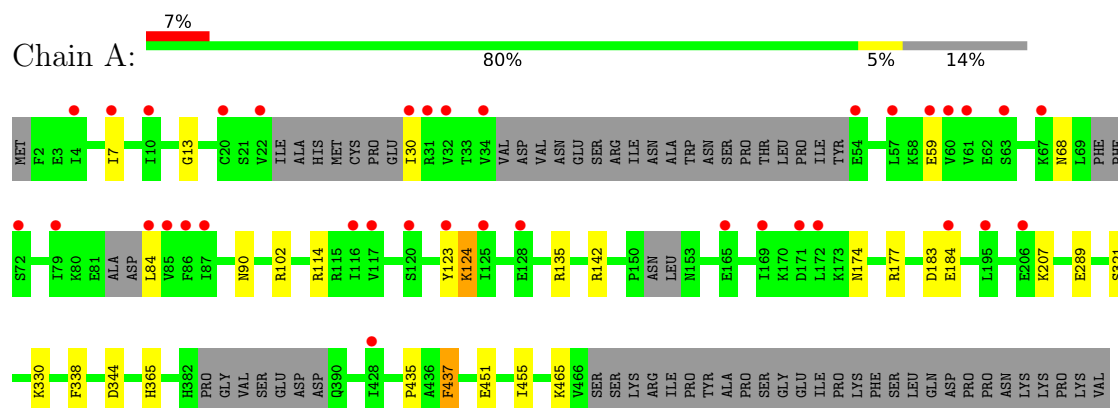
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	G	97	Total 97	O 97	0	0
7	H	25	Total 25	O 25	0	0
7	I	20	Total 20	O 20	0	0
7	J	29	Total 29	O 29	0	0
7	K	60	Total 60	O 60	0	0
7	L	34	Total 34	O 34	0	0

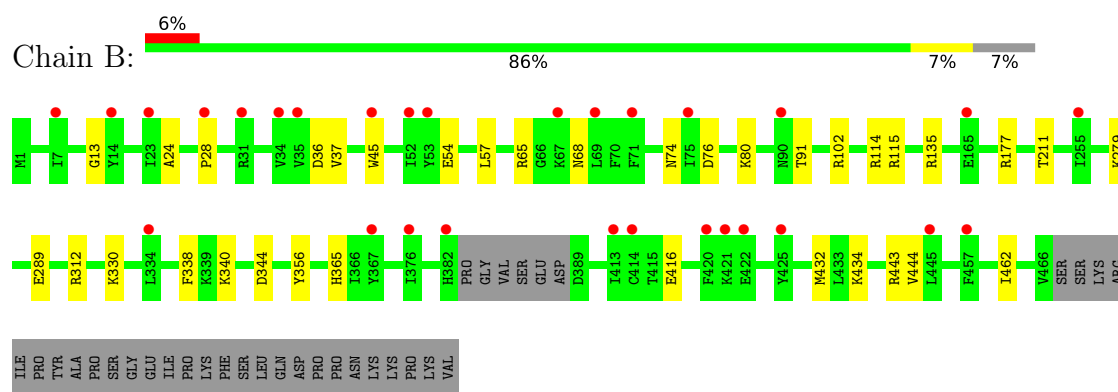
### 3 Residue-property plots [i](#)

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.

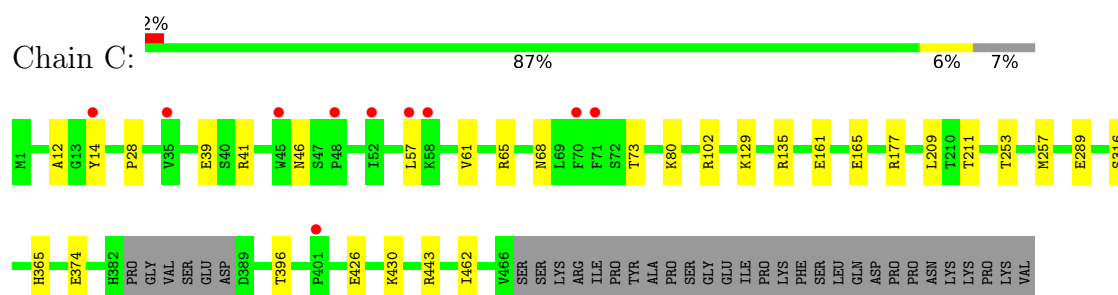
#### • Molecule 1: UDP-glucose 6-dehydrogenase




#### • Molecule 1: UDP-glucose 6-dehydrogenase

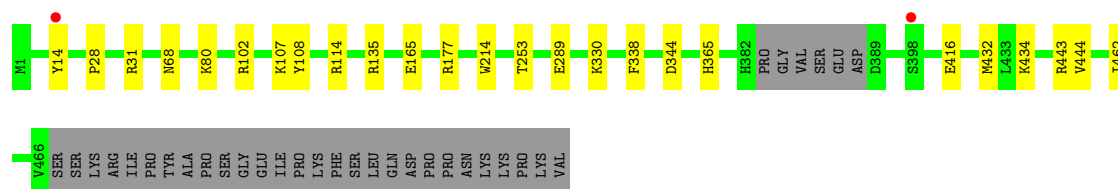


#### • Molecule 1: UDP-glucose 6-dehydrogenase




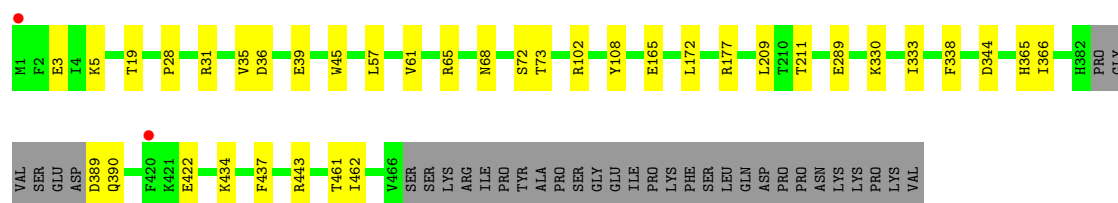
#### • Molecule 1: UDP-glucose 6-dehydrogenase

Chain D:  88% 5% 7%




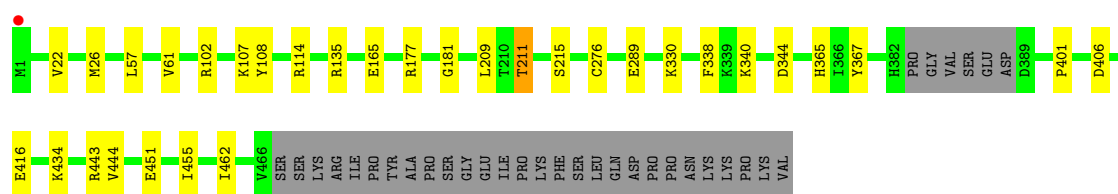
• Molecule 1: UDP-glucose 6-dehydrogenase

Chain E:  86% 7% 7%




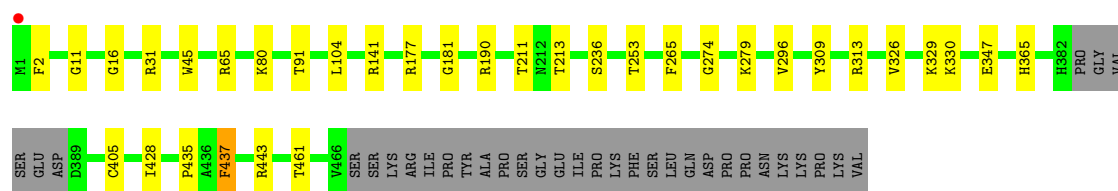
• Molecule 1: UDP-glucose 6-dehydrogenase

Chain F:  87% 6% 7%




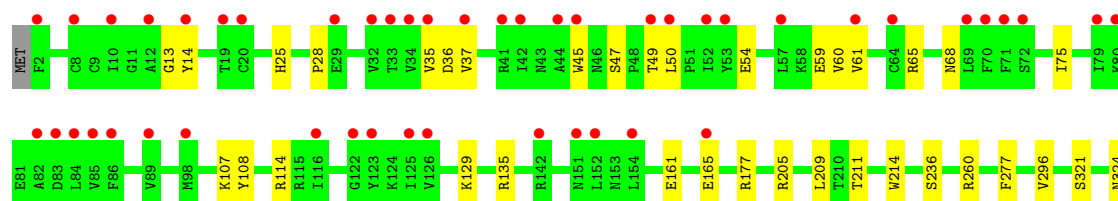
• Molecule 1: UDP-glucose 6-dehydrogenase

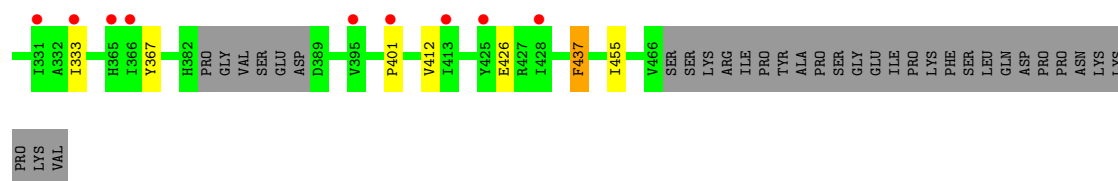
Chain G:  86% 7% 7%



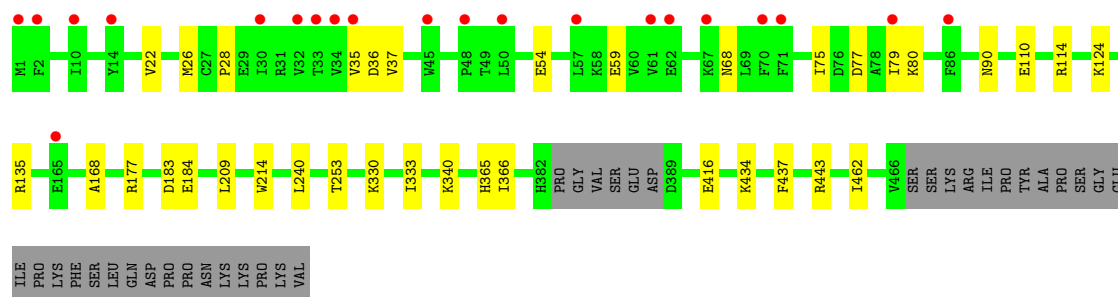
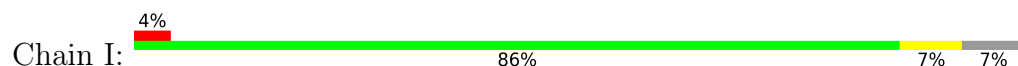
• Molecule 1: UDP-glucose 6-dehydrogenase

Chain H:  11% 84% 9% 7%

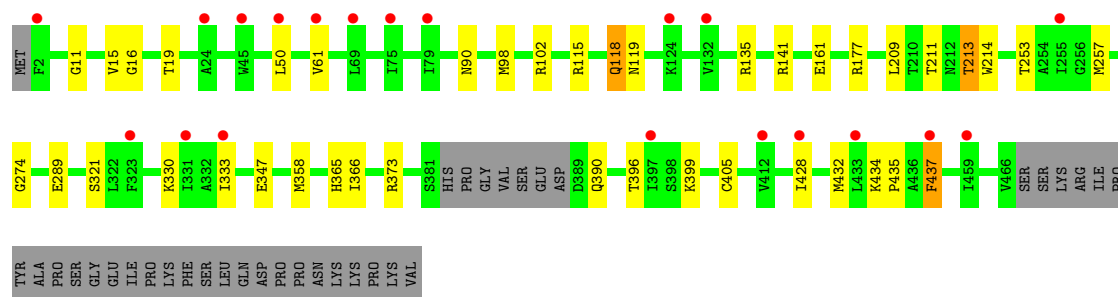
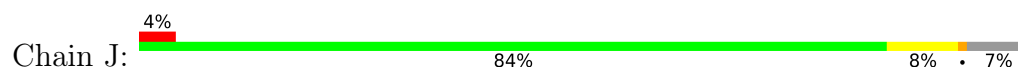




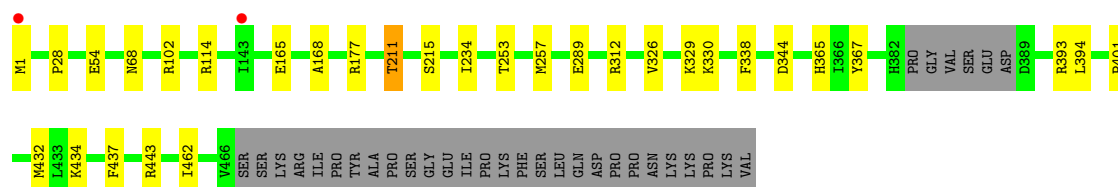
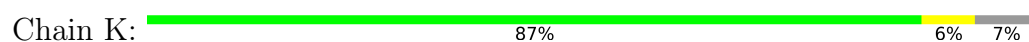
• Molecule 1: UDP-glucose 6-dehydrogenase



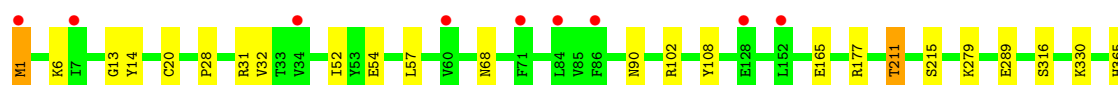
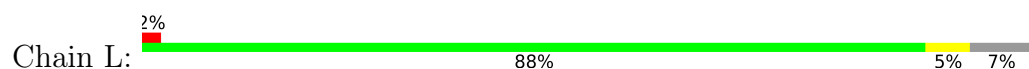
• Molecule 1: UDP-glucose 6-dehydrogenase



• Molecule 1: UDP-glucose 6-dehydrogenase



• Molecule 1: UDP-glucose 6-dehydrogenase



H382	PRO	GLY	VAL	SER	GLU	ASP	D389	E416	R443	T462	V466	SER	SER	LYS	ARG	ILE	PRO	TYR	ALA	PRO	SER	GLY	GLU	ILE	PRO	LYS	PHE	SER	LEU	GLN	ASP	PRO	PRO	ASN	LYS	LYS	PRO	LYS	VAL
------	-----	-----	-----	-----	-----	-----	------	------	------	------	------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	403.55Å 112.59Å 184.67Å 90.00° 98.50° 90.00°	Depositor
Resolution (Å)	48.63 – 2.53 82.58 – 2.53	Depositor EDS
% Data completeness (in resolution range)	99.9 (48.63-2.53) 100.0 (82.58-2.53)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.35 (at 2.51Å)	Xtriage
Refinement program	PHENIX (1.12_2829: ???)	Depositor
R, $R_{free}$	0.180 , 0.214 0.181 , 0.214	Depositor DCC
$R_{free}$ test set	13641 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	62.1	Xtriage
Anisotropy	0.329	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 45.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	44722	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	73.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.12% 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: UPG, PGO, SO4, NAD, CL

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	0/3409	0.44	0/4598
1	B	0.27	0/3674	0.43	0/4970
1	C	0.28	0/3689	0.44	0/4990
1	D	0.28	0/3680	0.45	0/4978
1	E	0.29	0/3688	0.45	0/4989
1	F	0.28	0/3680	0.45	0/4978
1	G	0.28	0/3699	0.45	0/5004
1	H	0.26	0/3666	0.42	0/4960
1	I	0.26	0/3697	0.43	0/5002
1	J	0.27	0/3668	0.43	0/4963
1	K	0.28	0/3680	0.45	0/4978
1	L	0.27	0/3680	0.44	0/4978
All	All	0.28	0/43910	0.44	0/59388

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	3347	0	3380	17	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	3609	0	3633	21	0
1	C	3618	0	3644	16	0
1	D	3612	0	3638	14	0
1	E	3617	0	3644	21	0
1	F	3612	0	3638	21	0
1	G	3623	0	3649	20	0
1	H	3601	0	3621	26	0
1	I	3623	0	3650	19	0
1	J	3598	0	3626	25	0
1	K	3612	0	3638	16	0
1	L	3612	0	3638	19	0
2	A	44	0	26	4	0
2	B	44	0	26	2	0
2	C	44	0	26	1	0
2	D	44	0	26	1	0
2	E	44	0	26	7	0
2	F	44	0	26	2	0
2	G	44	0	26	1	0
2	H	44	0	26	2	0
2	I	44	0	26	3	0
2	J	44	0	26	3	0
2	K	44	0	26	2	0
2	L	44	0	26	5	0
3	A	36	0	22	1	0
3	B	36	0	22	0	0
3	C	36	0	22	0	0
3	D	36	0	22	0	0
3	E	36	0	22	1	0
3	F	36	0	22	2	0
3	G	36	0	22	1	0
3	H	36	0	22	1	0
3	I	36	0	22	1	0
3	J	36	0	22	1	0
3	K	36	0	22	1	0
3	L	36	0	22	1	0
4	A	1	0	0	1	0
4	B	2	0	0	1	0
4	D	1	0	0	0	0
4	G	1	0	0	0	0
4	H	1	0	0	0	0
4	I	2	0	0	0	0
4	J	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	K	2	0	0	0	0
4	L	2	0	0	0	0
5	C	5	0	0	2	0
5	D	10	0	0	0	0
5	E	5	0	0	0	0
5	F	10	0	0	2	0
5	G	5	0	0	0	0
5	K	5	0	0	0	0
6	D	5	0	8	0	0
6	E	5	0	8	0	0
6	F	5	0	8	0	0
6	G	5	0	8	0	0
6	K	5	0	8	0	0
6	L	5	0	8	0	0
7	A	33	0	0	1	0
7	B	29	0	0	0	0
7	C	53	0	0	0	0
7	D	64	0	0	1	0
7	E	85	0	0	0	0
7	F	66	0	0	0	0
7	G	97	0	0	0	0
7	H	25	0	0	1	0
7	I	20	0	0	0	0
7	J	29	0	0	0	0
7	K	60	0	0	1	0
7	L	34	0	0	1	0
All	All	44722	0	44023	238	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 3.

The worst 5 of 238 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:36:ASP:HA	2:E:501:NAD:H8A	1.67	0.74
1:E:39:GLU:HG3	1:E:73:THR:HG21	1.71	0.73
1:F:135:ARG:NH1	5:F:504:SO4:O2	2.26	0.68
1:G:91:THR:O	1:G:104:LEU:HD22	1.95	0.67
1:G:45:TRP:O	1:G:65:ARG:NH1	2.29	0.65

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	415/494 (84%)	402 (97%)	13 (3%)	0	100	100
1	B	456/494 (92%)	446 (98%)	10 (2%)	0	100	100
1	C	458/494 (93%)	446 (97%)	12 (3%)	0	100	100
1	D	457/494 (92%)	446 (98%)	11 (2%)	0	100	100
1	E	458/494 (93%)	446 (97%)	12 (3%)	0	100	100
1	F	457/494 (92%)	446 (98%)	11 (2%)	0	100	100
1	G	459/494 (93%)	450 (98%)	9 (2%)	0	100	100
1	H	455/494 (92%)	443 (97%)	12 (3%)	0	100	100
1	I	459/494 (93%)	447 (97%)	12 (3%)	0	100	100
1	J	456/494 (92%)	447 (98%)	9 (2%)	0	100	100
1	K	457/494 (92%)	447 (98%)	10 (2%)	0	100	100
1	L	457/494 (92%)	446 (98%)	11 (2%)	0	100	100
All	All	5444/5928 (92%)	5312 (98%)	132 (2%)	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	368/427 (86%)	364 (99%)	4 (1%)	76	90
1	B	396/427 (93%)	393 (99%)	3 (1%)	83	94

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	398/427 (93%)	393 (99%)	5 (1%)	71	88
1	D	397/427 (93%)	393 (99%)	4 (1%)	78	91
1	E	398/427 (93%)	391 (98%)	7 (2%)	62	83
1	F	397/427 (93%)	393 (99%)	4 (1%)	78	91
1	G	399/427 (93%)	396 (99%)	3 (1%)	83	94
1	H	395/427 (92%)	392 (99%)	3 (1%)	83	94
1	I	399/427 (93%)	396 (99%)	3 (1%)	83	94
1	J	396/427 (93%)	391 (99%)	5 (1%)	71	88
1	K	397/427 (93%)	392 (99%)	5 (1%)	71	88
1	L	397/427 (93%)	393 (99%)	4 (1%)	78	91
All	All	4737/5124 (92%)	4687 (99%)	50 (1%)	78	90

5 of 50 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	437	PHE
1	G	31	ARG
1	L	1	MET
1	F	165	GLU
1	F	209	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	J	74	ASN
1	L	365	HIS
1	J	390	GLN
1	D	429	HIS
1	K	365	HIS

### 5.3.3 RNA ⓘ

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 51 ligands modelled in this entry, 13 are monoatomic - leaving 38 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	501	-	40,48,48	1.34	3 (7%)	44,73,73	1.55	6 (13%)
3	UPG	A	502	-	30,38,38	0.78	0	41,58,58	1.22	1 (2%)
2	NAD	B	501	-	40,48,48	1.34	3 (7%)	44,73,73	1.58	5 (11%)
3	UPG	B	502	-	30,38,38	0.78	0	41,58,58	1.22	1 (2%)
2	NAD	C	501	-	40,48,48	1.31	3 (7%)	44,73,73	1.47	3 (6%)
3	UPG	C	502	-	30,38,38	0.80	0	41,58,58	1.23	2 (4%)
5	SO4	C	503	-	4,4,4	0.15	0	6,6,6	0.10	0
2	NAD	D	501	-	40,48,48	1.32	3 (7%)	44,73,73	1.56	5 (11%)
3	UPG	D	502	-	30,38,38	0.74	0	41,58,58	1.26	1 (2%)
5	SO4	D	504	-	4,4,4	0.19	0	6,6,6	0.08	0
5	SO4	D	505	-	4,4,4	0.17	0	6,6,6	0.16	0
6	PGO	D	506	-	4,4,4	0.57	0	4,4,4	0.65	0
2	NAD	E	501	-	40,48,48	1.35	3 (7%)	44,73,73	1.93	6 (13%)
3	UPG	E	502	-	30,38,38	0.81	0	41,58,58	1.25	3 (7%)
5	SO4	E	503	-	4,4,4	0.16	0	6,6,6	0.10	0
6	PGO	E	504	-	4,4,4	0.56	0	4,4,4	0.55	0
2	NAD	F	501	-	40,48,48	1.36	3 (7%)	44,73,73	1.55	5 (11%)
3	UPG	F	502	-	30,38,38	0.74	0	41,58,58	1.22	2 (4%)
5	SO4	F	503	-	4,4,4	0.17	0	6,6,6	0.08	0
5	SO4	F	504	-	4,4,4	0.17	0	6,6,6	0.10	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
6	PGO	F	505	-	4,4,4	0.57	0	4,4,4	0.59	0
2	NAD	G	501	-	40,48,48	1.28	3 (7%)	44,73,73	1.64	4 (9%)
3	UPG	G	502	-	30,38,38	0.73	0	41,58,58	1.18	1 (2%)
5	SO4	G	504	-	4,4,4	0.41	0	6,6,6	0.07	0
6	PGO	G	505	-	4,4,4	0.55	0	4,4,4	0.58	0
2	NAD	H	501	-	40,48,48	1.34	3 (7%)	44,73,73	1.53	5 (11%)
3	UPG	H	502	-	30,38,38	0.76	0	41,58,58	1.23	1 (2%)
2	NAD	I	501	-	40,48,48	1.34	3 (7%)	44,73,73	1.51	6 (13%)
3	UPG	I	502	-	30,38,38	0.77	0	41,58,58	1.25	1 (2%)
2	NAD	J	501	-	40,48,48	1.33	3 (7%)	44,73,73	1.56	5 (11%)
3	UPG	J	502	-	30,38,38	0.75	0	41,58,58	1.24	1 (2%)
2	NAD	K	501	-	40,48,48	1.30	3 (7%)	44,73,73	1.56	5 (11%)
3	UPG	K	502	-	30,38,38	0.75	0	41,58,58	1.23	1 (2%)
5	SO4	K	505	-	4,4,4	0.17	0	6,6,6	0.11	0
6	PGO	K	506	-	4,4,4	0.49	0	4,4,4	0.66	0
2	NAD	L	501	-	40,48,48	1.36	4 (10%)	44,73,73	1.74	7 (15%)
3	UPG	L	502	-	30,38,38	0.78	0	41,58,58	1.33	2 (4%)
6	PGO	L	505	-	4,4,4	0.57	0	4,4,4	0.32	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
2	NAD	A	501	-	-	0/22/62/62	0/5/5/5
3	UPG	A	502	-	-	0/19/59/59	0/3/3/3
2	NAD	B	501	-	-	0/22/62/62	0/5/5/5
3	UPG	B	502	-	-	0/19/59/59	0/3/3/3
2	NAD	C	501	-	-	0/22/62/62	0/5/5/5
3	UPG	C	502	-	-	0/19/59/59	0/3/3/3
5	SO4	C	503	-	-	0/0/0/0	0/0/0/0
2	NAD	D	501	-	-	0/22/62/62	0/5/5/5
3	UPG	D	502	-	-	0/19/59/59	0/3/3/3
5	SO4	D	504	-	-	0/0/0/0	0/0/0/0
5	SO4	D	505	-	-	0/0/0/0	0/0/0/0
6	PGO	D	506	-	-	0/2/2/2	0/0/0/0
2	NAD	E	501	-	-	0/22/62/62	0/5/5/5
3	UPG	E	502	-	-	0/19/59/59	0/3/3/3
5	SO4	E	503	-	-	0/0/0/0	0/0/0/0

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	PGO	E	504	-	-	0/2/2/2	0/0/0/0
2	NAD	F	501	-	-	0/22/62/62	0/5/5/5
3	UPG	F	502	-	-	0/19/59/59	0/3/3/3
5	SO4	F	503	-	-	0/0/0/0	0/0/0/0
5	SO4	F	504	-	-	0/0/0/0	0/0/0/0
6	PGO	F	505	-	-	0/2/2/2	0/0/0/0
2	NAD	G	501	-	-	0/22/62/62	0/5/5/5
3	UPG	G	502	-	-	0/19/59/59	0/3/3/3
5	SO4	G	504	-	-	0/0/0/0	0/0/0/0
6	PGO	G	505	-	-	0/2/2/2	0/0/0/0
2	NAD	H	501	-	-	0/22/62/62	0/5/5/5
3	UPG	H	502	-	-	0/19/59/59	0/3/3/3
2	NAD	I	501	-	-	0/22/62/62	0/5/5/5
3	UPG	I	502	-	-	0/19/59/59	0/3/3/3
2	NAD	J	501	-	-	0/22/62/62	0/5/5/5
3	UPG	J	502	-	-	0/19/59/59	0/3/3/3
2	NAD	K	501	-	-	0/22/62/62	0/5/5/5
3	UPG	K	502	-	-	0/19/59/59	0/3/3/3
5	SO4	K	505	-	-	0/0/0/0	0/0/0/0
6	PGO	K	506	-	-	0/2/2/2	0/0/0/0
2	NAD	L	501	-	-	0/22/62/62	0/5/5/5
3	UPG	L	502	-	-	0/19/59/59	0/3/3/3
6	PGO	L	505	-	-	0/2/2/2	0/0/0/0

The worst 5 of 37 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	L	501	NAD	C3N-C7N	2.10	1.53	1.50
2	A	501	NAD	C6N-N1N	2.42	1.41	1.35
2	H	501	NAD	C6N-N1N	2.43	1.41	1.35
2	K	501	NAD	C6N-N1N	2.43	1.41	1.35
2	I	501	NAD	C6N-N1N	2.47	1.41	1.35

The worst 5 of 79 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	E	501	NAD	C4B-O4B-C1B	-7.17	102.36	109.83
2	G	501	NAD	C5N-C4N-C3N	-6.20	113.07	120.35
2	D	501	NAD	C5N-C4N-C3N	-6.00	113.30	120.35
2	K	501	NAD	C5N-C4N-C3N	-5.92	113.39	120.35
2	C	501	NAD	C5N-C4N-C3N	-5.90	113.42	120.35

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

24 monomers are involved in 39 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	501	NAD	4	0
3	A	502	UPG	1	0
2	B	501	NAD	2	0
2	C	501	NAD	1	0
5	C	503	SO4	2	0
2	D	501	NAD	1	0
2	E	501	NAD	7	0
3	E	502	UPG	1	0
2	F	501	NAD	2	0
3	F	502	UPG	2	0
5	F	503	SO4	1	0
5	F	504	SO4	1	0
2	G	501	NAD	1	0
3	G	502	UPG	1	0
2	H	501	NAD	2	0
3	H	502	UPG	1	0
2	I	501	NAD	3	0
3	I	502	UPG	1	0
2	J	501	NAD	3	0
3	J	502	UPG	1	0
2	K	501	NAD	2	0
3	K	502	UPG	1	0
2	L	501	NAD	5	0
3	L	502	UPG	1	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](#)

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	426/494 (86%)	0.59	36 (8%) 11 11	46, 77, 134, 160	0
1	B	460/494 (93%)	0.43	29 (6%) 20 21	44, 74, 119, 148	0
1	C	460/494 (93%)	0.33	10 (2%) 62 65	39, 64, 102, 130	0
1	D	460/494 (93%)	0.25	2 (0%) 92 93	38, 60, 91, 128	0
1	E	460/494 (93%)	0.20	2 (0%) 92 93	44, 59, 92, 125	0
1	F	460/494 (93%)	0.13	1 (0%) 94 96	43, 62, 90, 133	0
1	G	460/494 (93%)	0.23	1 (0%) 94 96	43, 58, 85, 134	0
1	H	459/494 (92%)	0.79	56 (12%) 4 4	50, 87, 133, 149	0
1	I	460/494 (93%)	0.41	21 (4%) 32 35	49, 79, 116, 145	0
1	J	458/494 (92%)	0.42	20 (4%) 34 37	47, 77, 106, 126	0
1	K	460/494 (93%)	0.16	2 (0%) 92 93	46, 62, 88, 138	0
1	L	460/494 (93%)	0.31	9 (1%) 65 68	49, 75, 104, 137	0
All	All	5483/5928 (92%)	0.35	189 (3%) 45 49	38, 68, 113, 160	0

The worst 5 of 189 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	H	52	ILE	7.4
1	H	71	PHE	5.3
1	A	57	LEU	5.3
1	H	70	PHE	5.2
1	A	34	VAL	5.0

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

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.3 Carbohydrates ⓘ

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(Å <sup>2</sup> )	Q<0.9
4	CL	K	503	1/1	0.65	0.15	97,97,97,97	0
4	CL	L	503	1/1	0.73	0.15	93,93,93,93	0
4	CL	G	503	1/1	0.76	0.11	87,87,87,87	0
5	SO4	C	503	5/5	0.77	0.32	97,101,104,109	5
4	CL	I	503	1/1	0.80	0.22	94,94,94,94	0
4	CL	I	504	1/1	0.81	0.10	102,102,102,102	0
6	PGO	D	506	5/5	0.84	0.24	58,62,71,71	0
4	CL	B	504	1/1	0.84	0.35	96,96,96,96	0
4	CL	J	503	1/1	0.85	0.15	87,87,87,87	0
6	PGO	G	505	5/5	0.88	0.32	76,84,89,98	0
5	SO4	F	504	5/5	0.88	0.19	76,76,83,89	5
4	CL	H	503	1/1	0.88	0.15	90,90,90,90	0
5	SO4	F	503	5/5	0.89	0.38	65,66,71,72	5
4	CL	L	504	1/1	0.91	0.34	93,93,93,93	0
2	NAD	A	501	44/44	0.91	0.16	79,100,117,119	0
6	PGO	F	505	5/5	0.91	0.25	64,68,74,74	0
4	CL	A	503	1/1	0.91	0.07	96,96,96,96	0
6	PGO	K	506	5/5	0.92	0.23	64,72,74,82	0
3	UPG	H	502	36/36	0.92	0.16	63,85,100,110	0
4	CL	D	503	1/1	0.92	0.40	90,90,90,90	0
5	SO4	D	505	5/5	0.92	0.23	81,89,95,96	5
5	SO4	K	505	5/5	0.92	0.18	55,71,73,74	5
3	UPG	I	502	36/36	0.92	0.17	59,93,127,130	0
2	NAD	H	501	44/44	0.93	0.20	88,98,107,110	0
5	SO4	D	504	5/5	0.93	0.24	56,58,63,66	5
2	NAD	L	501	44/44	0.93	0.17	73,87,110,116	0
4	CL	K	504	1/1	0.94	0.39	93,93,93,93	0
5	SO4	E	503	5/5	0.94	0.24	91,96,97,99	5
6	PGO	L	505	5/5	0.94	0.34	78,80,83,84	0
5	SO4	G	504	5/5	0.94	0.18	88,90,98,98	0
2	NAD	B	501	44/44	0.95	0.17	75,95,107,108	0
4	CL	B	503	1/1	0.95	0.38	93,93,93,93	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	UPG	B	502	36/36	0.96	0.15	52,72,103,108	0
6	PGO	E	504	5/5	0.96	0.30	67,74,77,78	0
3	UPG	C	502	36/36	0.96	0.15	44,64,80,86	0
2	NAD	I	501	44/44	0.96	0.15	80,98,106,113	0
2	NAD	E	501	44/44	0.96	0.17	47,60,72,81	0
3	UPG	J	502	36/36	0.97	0.15	53,62,66,68	0
2	NAD	J	501	44/44	0.97	0.14	53,67,73,83	0
2	NAD	D	501	44/44	0.97	0.15	50,63,72,78	0
3	UPG	F	502	36/36	0.97	0.15	41,55,65,68	0
3	UPG	K	502	36/36	0.97	0.14	53,64,72,78	0
2	NAD	C	501	44/44	0.97	0.14	54,77,82,84	0
3	UPG	G	502	36/36	0.98	0.18	38,49,56,62	0
3	UPG	A	502	36/36	0.98	0.14	54,62,74,82	0
3	UPG	E	502	36/36	0.98	0.14	44,58,63,69	0
2	NAD	K	501	44/44	0.98	0.16	41,57,66,73	0
3	UPG	L	502	36/36	0.98	0.14	45,67,82,87	0
2	NAD	G	501	44/44	0.98	0.18	37,48,58,62	0
3	UPG	D	502	36/36	0.98	0.15	43,57,70,73	0
2	NAD	F	501	44/44	0.98	0.15	38,60,68,72	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.