



# wwPDB X-ray Structure Validation Summary Report

Nov 8, 2021 – 07:11 pm GMT

PDB ID : 6YBH  
Title : Deoxyribonucleoside Kinase  
Authors : Allouche-Arnon, H.; Bar-Shir, A.; Dym, O.  
Deposited on : 2020-03-17  
Resolution : 2.40 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 : 1.8.4 (270009), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.23.2  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0267  
CCP4 : 7.1.010 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.23.2

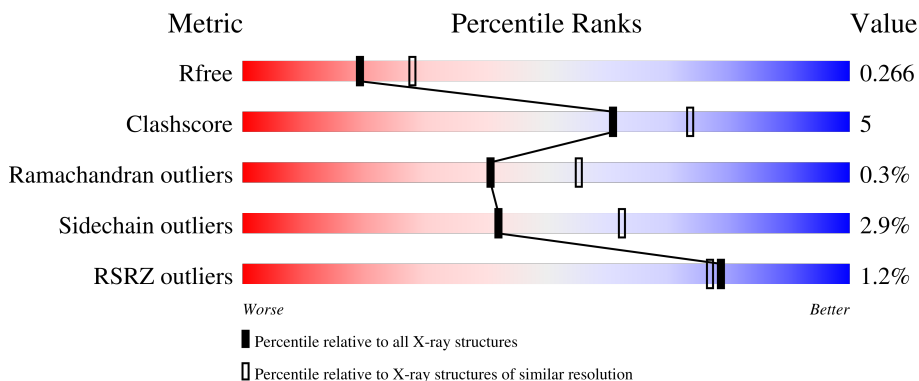
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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 of 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	224	
1	B	224	
1	C	224	
1	D	224	
1	E	224	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	224	<p>%</p> <p>82% 12% 5%</p>

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
4	GOL	B	306	-	-	X	-
4	GOL	F	304	-	-	X	-

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 10856 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 Deoxynucleoside kinase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	217	1788	1145	302	331	10	0	0	0
1	B	215	1777	1140	298	329	10	0	0	0
1	C	212	1736	1113	294	318	11	0	0	0
1	D	219	1820	1164	310	336	10	0	0	0
1	E	209	1684	1083	287	303	11	0	0	0
1	F	212	1670	1072	275	313	10	0	0	0

There are 108 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	12	MET	-	initiating methionine	UNP Q9XZT6
A	19	ARG	GLN	engineered mutation	UNP Q9XZT6
A	23	ILE	VAL	engineered mutation	UNP Q9XZT6
A	45	ASP	ASN	engineered mutation	UNP Q9XZT6
A	81	VAL	GLN	engineered mutation	UNP Q9XZT6
A	91	MET	SER	engineered mutation	UNP Q9XZT6
A	94	GLN	ALA	engineered mutation	UNP Q9XZT6
A	100	VAL	LEU	engineered mutation	UNP Q9XZT6
A	123	TRP	SER	engineered mutation	UNP Q9XZT6
A	145	ILE	VAL	engineered mutation	UNP Q9XZT6
A	166	LYS	GLN	engineered mutation	UNP Q9XZT6
A	168	GLY	ALA	engineered mutation	UNP Q9XZT6
A	170	PRO	SER	engineered mutation	UNP Q9XZT6
A	173	LYS	SER	engineered mutation	UNP Q9XZT6
A	174	ASN	CYS	engineered mutation	UNP Q9XZT6
A	178	GLU	LYS	engineered mutation	UNP Q9XZT6
A	182	GLN	GLU	engineered mutation	UNP Q9XZT6

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	185	GLN	GLU	engineered mutation	UNP Q9XZT6
B	12	MET	-	initiating methionine	UNP Q9XZT6
B	19	ARG	GLN	engineered mutation	UNP Q9XZT6
B	23	ILE	VAL	engineered mutation	UNP Q9XZT6
B	45	ASP	ASN	engineered mutation	UNP Q9XZT6
B	81	VAL	GLN	engineered mutation	UNP Q9XZT6
B	91	MET	SER	engineered mutation	UNP Q9XZT6
B	94	GLN	ALA	engineered mutation	UNP Q9XZT6
B	100	VAL	LEU	engineered mutation	UNP Q9XZT6
B	123	TRP	SER	engineered mutation	UNP Q9XZT6
B	145	ILE	VAL	engineered mutation	UNP Q9XZT6
B	166	LYS	GLN	engineered mutation	UNP Q9XZT6
B	168	GLY	ALA	engineered mutation	UNP Q9XZT6
B	170	PRO	SER	engineered mutation	UNP Q9XZT6
B	173	LYS	SER	engineered mutation	UNP Q9XZT6
B	174	ASN	CYS	engineered mutation	UNP Q9XZT6
B	178	GLU	LYS	engineered mutation	UNP Q9XZT6
B	182	GLN	GLU	engineered mutation	UNP Q9XZT6
B	185	GLN	GLU	engineered mutation	UNP Q9XZT6
C	12	MET	-	initiating methionine	UNP Q9XZT6
C	19	ARG	GLN	engineered mutation	UNP Q9XZT6
C	23	ILE	VAL	engineered mutation	UNP Q9XZT6
C	45	ASP	ASN	engineered mutation	UNP Q9XZT6
C	81	VAL	GLN	engineered mutation	UNP Q9XZT6
C	91	MET	SER	engineered mutation	UNP Q9XZT6
C	94	GLN	ALA	engineered mutation	UNP Q9XZT6
C	100	VAL	LEU	engineered mutation	UNP Q9XZT6
C	123	TRP	SER	engineered mutation	UNP Q9XZT6
C	145	ILE	VAL	engineered mutation	UNP Q9XZT6
C	166	LYS	GLN	engineered mutation	UNP Q9XZT6
C	168	GLY	ALA	engineered mutation	UNP Q9XZT6
C	170	PRO	SER	engineered mutation	UNP Q9XZT6
C	173	LYS	SER	engineered mutation	UNP Q9XZT6
C	174	ASN	CYS	engineered mutation	UNP Q9XZT6
C	178	GLU	LYS	engineered mutation	UNP Q9XZT6
C	182	GLN	GLU	engineered mutation	UNP Q9XZT6
C	185	GLN	GLU	engineered mutation	UNP Q9XZT6
D	12	MET	-	initiating methionine	UNP Q9XZT6
D	19	ARG	GLN	engineered mutation	UNP Q9XZT6
D	23	ILE	VAL	engineered mutation	UNP Q9XZT6
D	45	ASP	ASN	engineered mutation	UNP Q9XZT6
D	81	VAL	GLN	engineered mutation	UNP Q9XZT6

*Continued on next page...*

*Continued from previous page...*

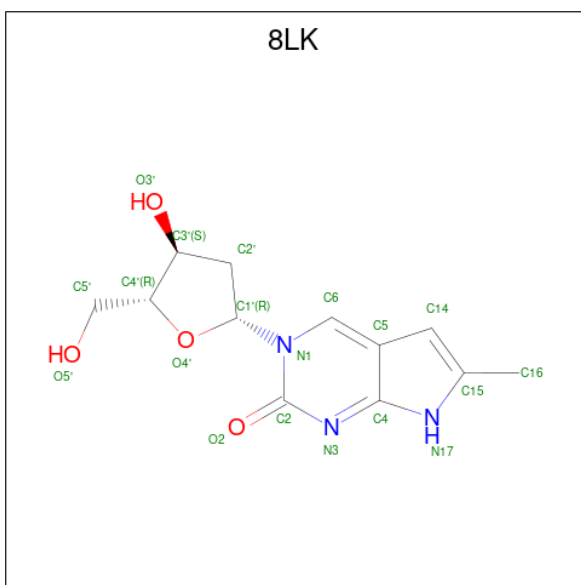
Chain	Residue	Modelled	Actual	Comment	Reference
D	91	MET	SER	engineered mutation	UNP Q9XZT6
D	94	GLN	ALA	engineered mutation	UNP Q9XZT6
D	100	VAL	LEU	engineered mutation	UNP Q9XZT6
D	123	TRP	SER	engineered mutation	UNP Q9XZT6
D	145	ILE	VAL	engineered mutation	UNP Q9XZT6
D	166	LYS	GLN	engineered mutation	UNP Q9XZT6
D	168	GLY	ALA	engineered mutation	UNP Q9XZT6
D	170	PRO	SER	engineered mutation	UNP Q9XZT6
D	173	LYS	SER	engineered mutation	UNP Q9XZT6
D	174	ASN	CYS	engineered mutation	UNP Q9XZT6
D	178	GLU	LYS	engineered mutation	UNP Q9XZT6
D	182	GLN	GLU	engineered mutation	UNP Q9XZT6
D	185	GLN	GLU	engineered mutation	UNP Q9XZT6
E	12	MET	-	initiating methionine	UNP Q9XZT6
E	19	ARG	GLN	engineered mutation	UNP Q9XZT6
E	23	ILE	VAL	engineered mutation	UNP Q9XZT6
E	45	ASP	ASN	engineered mutation	UNP Q9XZT6
E	81	VAL	GLN	engineered mutation	UNP Q9XZT6
E	91	MET	SER	engineered mutation	UNP Q9XZT6
E	94	GLN	ALA	engineered mutation	UNP Q9XZT6
E	100	VAL	LEU	engineered mutation	UNP Q9XZT6
E	123	TRP	SER	engineered mutation	UNP Q9XZT6
E	145	ILE	VAL	engineered mutation	UNP Q9XZT6
E	166	LYS	GLN	engineered mutation	UNP Q9XZT6
E	168	GLY	ALA	engineered mutation	UNP Q9XZT6
E	170	PRO	SER	engineered mutation	UNP Q9XZT6
E	173	LYS	SER	engineered mutation	UNP Q9XZT6
E	174	ASN	CYS	engineered mutation	UNP Q9XZT6
E	178	GLU	LYS	engineered mutation	UNP Q9XZT6
E	182	GLN	GLU	engineered mutation	UNP Q9XZT6
E	185	GLN	GLU	engineered mutation	UNP Q9XZT6
F	12	MET	-	initiating methionine	UNP Q9XZT6
F	19	ARG	GLN	engineered mutation	UNP Q9XZT6
F	23	ILE	VAL	engineered mutation	UNP Q9XZT6
F	45	ASP	ASN	engineered mutation	UNP Q9XZT6
F	81	VAL	GLN	engineered mutation	UNP Q9XZT6
F	91	MET	SER	engineered mutation	UNP Q9XZT6
F	94	GLN	ALA	engineered mutation	UNP Q9XZT6
F	100	VAL	LEU	engineered mutation	UNP Q9XZT6
F	123	TRP	SER	engineered mutation	UNP Q9XZT6
F	145	ILE	VAL	engineered mutation	UNP Q9XZT6
F	166	LYS	GLN	engineered mutation	UNP Q9XZT6

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	168	GLY	ALA	engineered mutation	UNP Q9XZT6
F	170	PRO	SER	engineered mutation	UNP Q9XZT6
F	173	LYS	SER	engineered mutation	UNP Q9XZT6
F	174	ASN	CYS	engineered mutation	UNP Q9XZT6
F	178	GLU	LYS	engineered mutation	UNP Q9XZT6
F	182	GLN	GLU	engineered mutation	UNP Q9XZT6
F	185	GLN	GLU	engineered mutation	UNP Q9XZT6

- Molecule 2 is Pyrrolo-dC (three-letter code: 8LK) (formula: C<sub>12</sub>H<sub>15</sub>N<sub>3</sub>O<sub>4</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	
			Total	C	N			O
2	A	1	Total	C	N	O	0	0
			19	12	3	4		
2	B	1	Total	C	N	O	0	0
			19	12	3	4		
2	C	1	Total	C	N	O	0	0
			19	12	3	4		
2	D	1	Total	C	N	O	0	0
			19	12	3	4		
2	E	1	Total	C	N	O	0	0
			19	12	3	4		
2	F	1	Total	C	N	O	0	0
			19	12	3	4		

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	S	0	0
			5	4	1		
3	A	1	Total	O	S	0	0
			5	4	1		
3	A	1	Total	O	S	0	0
			5	4	1		
3	A	1	Total	O	S	0	0
			5	4	1		
3	A	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		
3	C	1	Total	O	S	0	0
			5	4	1		
3	C	1	Total	O	S	0	0
			5	4	1		
3	C	1	Total	O	S	0	0
			5	4	1		
3	D	1	Total	O	S	0	0
			5	4	1		

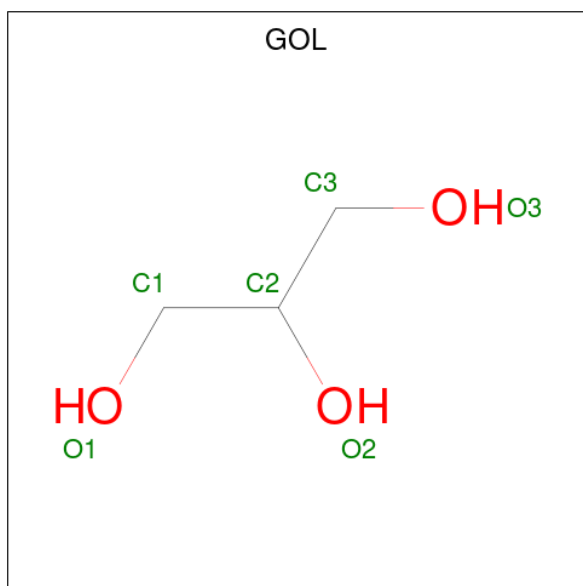
*Continued on next page...*



Continued from previous page...

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

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



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

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	D	1	Total	C	O	0	0
			6	3	3		
4	D	1	Total	C	O	0	0
			6	3	3		
4	D	1	Total	C	O	0	0
			6	3	3		
4	E	1	Total	C	O	0	0
			6	3	3		
4	E	1	Total	C	O	0	0
			6	3	3		
4	E	1	Total	C	O	0	0
			6	3	3		
4	F	1	Total	C	O	0	0
			6	3	3		

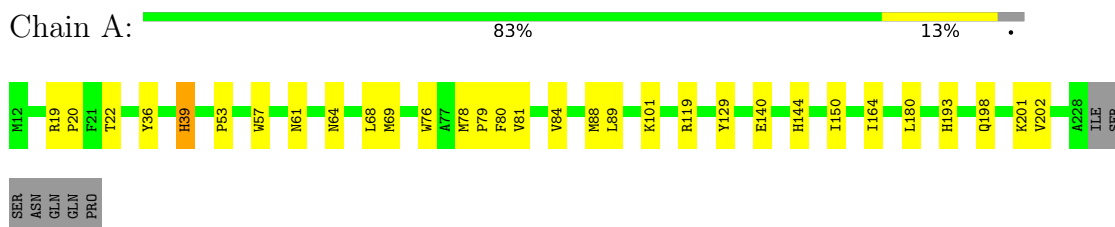
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	23	Total	O	0	0
			23	23		
5	B	20	Total	O	0	0
			20	20		
5	C	23	Total	O	0	0
			23	23		
5	D	15	Total	O	0	0
			15	15		
5	E	11	Total	O	0	0
			11	11		
5	F	4	Total	O	0	0
			4	4		

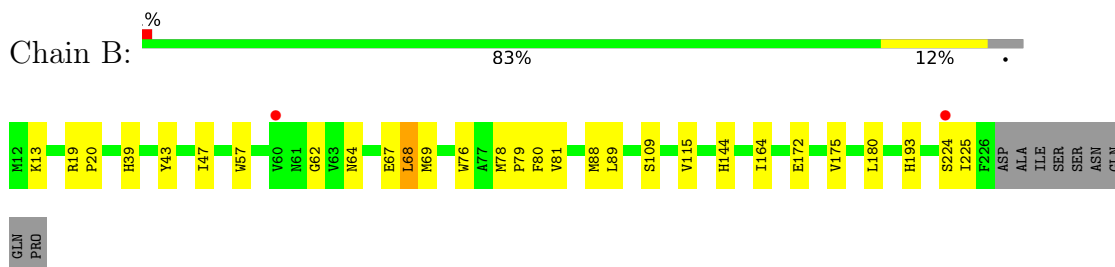
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

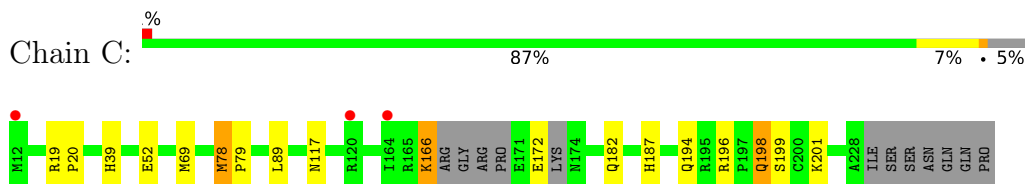
- Molecule 1: Deoxynucleoside kinase



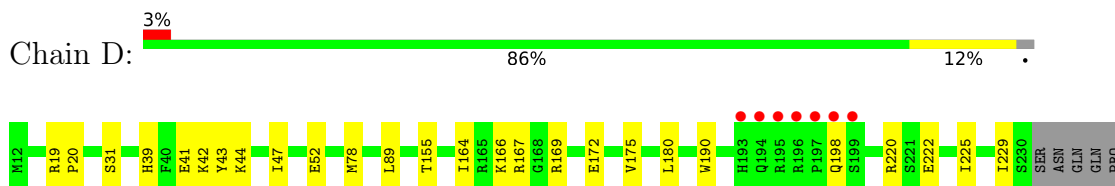
- Molecule 1: Deoxynucleoside kinase



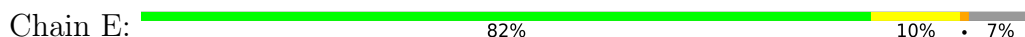
- Molecule 1: Deoxynucleoside kinase



- Molecule 1: Deoxynucleoside kinase



- Molecule 1: Deoxynucleoside kinase

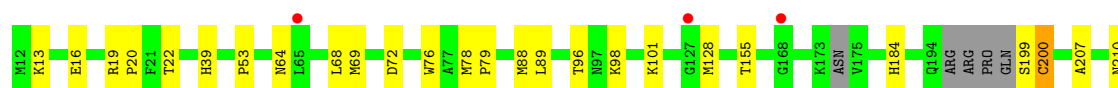




ASP  
ALA  
ILE  
SER  
SER  
ASN  
GLN  
GLN  
PRO

● Molecule 1: Deoxynucleoside kinase

Chain F: %



M213  
T216  
A228  
ILE  
SER  
SER  
ASN  
GLN  
GLN  
PRO

## 4 Data and refinement statistics

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	190.93Å 190.93Å 115.46Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	42.66 – 2.40 42.62 – 2.40	Depositor EDS
% Data completeness (in resolution range)	100.0 (42.66-2.40) 100.0 (42.62-2.40)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.85 (at 2.39Å)	Xtrriage
Refinement program	REFMAC 5.8.0258	Depositor
R, $R_{free}$	0.206 , 0.263 0.211 , 0.266	Depositor DCC
$R_{free}$ test set	2939 reflections (4.79%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	38.4	Xtrriage
Anisotropy	0.444	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	(Not available) , (Not available)	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	0.016 for $-2/3^*h-1/3^*k-4/3^*l,-1/3^*h-2/3^*k+4/3^*l,-1/3^*h+1/3^*k+1/3^*l$ 0.011 for $-h,1/3^*h-1/3^*k-4/3^*l,-1/3^*h-2/3^*k+1/3^*l$ 0.012 for $-1/3^*h+1/3^*k+4/3^*l,-k,2/3^*h+1/3^*k+1/3^*l$ 0.012 for $-h,2/3^*h+1/3^*k+4/3^*l,1/3^*h+2/3^*k-1/3^*l$ 0.018 for $-1/3^*h-2/3^*k+4/3^*l,-2/3^*h-1/3^*k-4/3^*l,1/3^*h-1/3^*k-1/3^*l$ 0.011 for $1/3^*h+2/3^*k-4/3^*l,-k,-2/3^*h-1/3^*k-1/3^*l$ 0.245 for h,-h-k,-l	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	10856	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	44.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.50% 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: 8LK, SO4, GOL

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	0.75	0/1831	0.91	0/2482
1	B	0.74	0/1820	0.87	0/2466
1	C	0.73	1/1775 (0.1%)	0.88	0/2402
1	D	0.78	2/1863 (0.1%)	0.87	0/2521
1	E	0.77	0/1725	0.90	0/2342
1	F	0.78	0/1707	0.88	0/2318
All	All	0.76	3/10721 (0.0%)	0.89	0/14531

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	52	GLU	CD-OE1	6.46	1.32	1.25
1	D	222	GLU	CD-OE1	5.30	1.31	1.25
1	C	52	GLU	CD-OE1	5.16	1.31	1.25

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	1788	0	1731	21	1
1	B	1777	0	1726	20	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	1736	0	1684	11	0
1	D	1820	0	1787	13	0
1	E	1684	0	1621	12	0
1	F	1670	0	1563	19	0
2	A	19	0	0	0	0
2	B	19	0	0	0	0
2	C	19	0	0	0	0
2	D	19	0	0	1	0
2	E	19	0	0	2	0
2	F	19	0	0	0	0
3	A	30	0	0	0	0
3	B	20	0	0	0	0
3	C	15	0	0	0	0
3	D	15	0	0	0	0
3	E	15	0	0	0	0
3	F	10	0	0	0	0
4	A	12	0	16	2	0
4	B	6	0	8	4	0
4	C	6	0	8	0	0
4	D	18	0	24	3	0
4	E	18	0	24	0	0
4	F	6	0	8	5	0
5	A	23	0	0	0	0
5	B	20	0	0	0	0
5	C	23	0	0	2	0
5	D	15	0	0	0	0
5	E	11	0	0	0	0
5	F	4	0	0	0	0
All	All	10856	0	10200	95	1

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.

The worst 5 of 95 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:F:96:THR:OG1	4:F:304:GOL:O2	1.92	0.87
1:F:96:THR:HG1	4:F:304:GOL:HO2	1.28	0.73
1:E:172:GLU:OE2	2:E:301:8LK:O3'	2.07	0.71
1:A:64:ASN:O	1:A:68:LEU:HD13	1.90	0.71
1:A:64:ASN:O	1:A:68:LEU:CD1	2.45	0.65

All (1) 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)
1:A:193:HIS:O	1:B:193:HIS:O[9_544]	1.54	0.66

## 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	215/224 (96%)	207 (96%)	8 (4%)	0	100	100
1	B	213/224 (95%)	201 (94%)	11 (5%)	1 (0%)	29	41
1	C	206/224 (92%)	194 (94%)	12 (6%)	0	100	100
1	D	217/224 (97%)	209 (96%)	8 (4%)	0	100	100
1	E	207/224 (92%)	196 (95%)	10 (5%)	1 (0%)	29	41
1	F	206/224 (92%)	194 (94%)	10 (5%)	2 (1%)	15	23
All	All	1264/1344 (94%)	1201 (95%)	59 (5%)	4 (0%)	41	55

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	F	72	ASP
1	F	200	CYS
1	B	68	LEU
1	E	197	PRO

### 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	192/208 (92%)	186 (97%)	6 (3%)	40	60
1	B	192/208 (92%)	187 (97%)	5 (3%)	46	66
1	C	186/208 (89%)	180 (97%)	6 (3%)	39	59
1	D	199/208 (96%)	193 (97%)	6 (3%)	41	61
1	E	175/208 (84%)	169 (97%)	6 (3%)	37	56
1	F	171/208 (82%)	168 (98%)	3 (2%)	59	76
All	All	1115/1248 (89%)	1083 (97%)	32 (3%)	42	62

5 of 32 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	199	SER
1	F	39	HIS
1	C	69	MET
1	C	39	HIS
1	F	69	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	F	184	HIS
1	F	181	GLN
1	C	198	GLN
1	C	187	HIS
1	E	38	ASN

### 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 monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

38 ligands are modelled in this entry.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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	217/224 (96%)	-0.42	0 <b>100</b> <b>100</b>	18, 34, 63, 92	0
1	B	215/224 (95%)	-0.33	2 (0%) <b>84</b> <b>82</b>	20, 34, 66, 95	0
1	C	212/224 (94%)	-0.28	3 (1%) <b>75</b> <b>73</b>	23, 38, 69, 91	0
1	D	219/224 (97%)	-0.22	7 (3%) <b>47</b> <b>46</b>	23, 38, 62, 113	0
1	E	209/224 (93%)	-0.20	1 (0%) <b>91</b> <b>89</b>	29, 47, 73, 133	0
1	F	212/224 (94%)	0.01	3 (1%) <b>75</b> <b>73</b>	31, 55, 88, 146	0
All	All	1284/1344 (95%)	-0.24	16 (1%) <b>79</b> <b>77</b>	18, 41, 74, 146	0

The worst 5 of 16 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	168	GLY	7.8
1	D	195	ARG	5.3
1	D	199	SER	5.1
1	D	198	GLN	3.4
1	D	197	PRO	3.2

### 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 [i](#)

There are no monosaccharides in this entry.

## 6.4 Ligands [i](#)

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
2	8LK	E	301	19/19	0.68	0.25	69,92,109,109	0
4	GOL	A	308	6/6	0.79	0.15	39,53,54,57	0
4	GOL	F	304	6/6	0.80	0.17	41,54,59,60	0
4	GOL	C	305	6/6	0.83	0.20	46,65,68,69	0
4	GOL	D	305	6/6	0.85	0.17	38,43,49,52	0
2	8LK	C	301	19/19	0.85	0.22	43,58,71,73	0
2	8LK	F	301	19/19	0.86	0.20	61,86,92,97	0
4	GOL	D	307	6/6	0.87	0.23	62,63,65,68	0
4	GOL	E	306	6/6	0.87	0.20	52,56,60,66	0
3	SO4	E	304	5/5	0.87	0.23	70,74,78,81	0
4	GOL	A	309	6/6	0.88	0.15	42,49,51,51	0
3	SO4	B	303	5/5	0.89	0.39	88,95,100,104	0
3	SO4	F	302	5/5	0.90	0.14	56,61,63,69	0
3	SO4	A	304	5/5	0.90	0.44	77,90,96,111	0
3	SO4	D	304	5/5	0.90	0.22	84,85,89,96	0
4	GOL	E	307	6/6	0.90	0.12	50,61,64,68	0
3	SO4	B	302	5/5	0.90	0.21	64,67,79,85	0
3	SO4	F	303	5/5	0.91	0.16	64,74,82,91	0
3	SO4	A	307	5/5	0.91	0.28	77,79,91,112	0
4	GOL	E	305	6/6	0.91	0.16	58,64,69,70	0
3	SO4	D	303	5/5	0.93	0.15	70,73,87,89	0
3	SO4	B	304	5/5	0.93	0.09	48,59,61,73	0
3	SO4	C	304	5/5	0.94	0.12	61,67,80,88	0
4	GOL	D	306	6/6	0.94	0.22	35,43,50,52	0
3	SO4	A	302	5/5	0.94	0.16	55,66,76,77	0
2	8LK	B	301	19/19	0.94	0.19	51,58,66,76	0
3	SO4	E	302	5/5	0.94	0.14	57,57,60,65	0
4	GOL	B	306	6/6	0.94	0.13	45,52,56,62	0
2	8LK	A	301	19/19	0.94	0.19	55,64,78,82	0
3	SO4	D	302	5/5	0.95	0.14	62,64,67,77	0
2	8LK	D	301	19/19	0.95	0.13	39,42,46,48	0
3	SO4	A	303	5/5	0.95	0.19	70,73,79,87	0
3	SO4	C	303	5/5	0.97	0.13	70,71,72,75	0
3	SO4	A	306	5/5	0.98	0.13	48,59,60,61	0
3	SO4	C	302	5/5	0.98	0.12	42,46,51,52	0
3	SO4	E	303	5/5	0.99	0.13	40,43,46,48	0
3	SO4	B	305	5/5	0.99	0.11	27,29,31,37	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	SO4	A	305	5/5	0.99	0.10	29,32,32,34	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.