



# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 10, 2023 – 02:23 AM EDT

PDB ID : 4IQJ  
Title : Structure of PolIIIalpha-Tauc-DNA complex suggests an atomic model of the replisome  
Authors : Liu, B.; Lin, J.; Steitz, T.  
Deposited on : 2013-01-11  
Resolution : 3.20 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.35.1  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.35.1

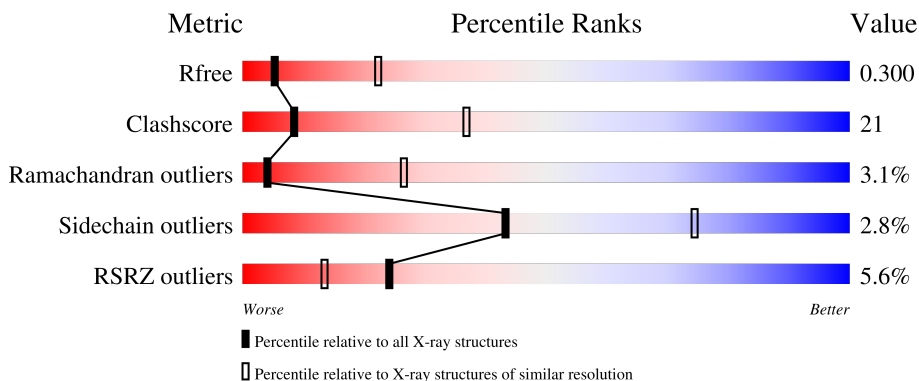
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.20 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	1133 (3.20-3.20)
Clashscore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)
RSRZ outliers	127900	1095 (3.20-3.20)

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	E	20	 5% 90% 10%
1	G	20	 15% 75% 20% 5%
1	K	20	 15% 90% 10%
2	F	28	 18% 50% 14% 7% 29%
2	H	28	 25% 82% 11% . .

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	J	28	
2	L	28	
3	I	21	
4	A	1220	
4	B	1220	
4	C	1220	
4	D	1220	
5	M	177	
5	N	177	
5	O	177	
5	P	177	

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
3	DOC	I	21	-	-	-	X

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 45265 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 DNA chain called DNA (5'-D(P\*CP\*GP\*AP\*AP\*AP\*CP\*GP\*AP\*CP\*GP\*GP\*CP\*CP\*AP\*GP\*TP\*GP\*CP\*CP\*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	E	20	Total	C	N	O	P	0	0	0
			411	193	83	115	20			
1	G	20	Total	C	N	O	P	0	0	0
			411	193	83	115	20			
1	K	20	Total	C	N	O	P	0	0	0
			411	193	83	115	20			

- Molecule 2 is a DNA chain called DNA (5'-D(\*TP\*TP\*TP\*TP\*TP\*TP\*TP\*GP\*TP\*GP\*GP\*CP\*AP\*CP\*TP\*GP\*GP\*CP\*CP\*GP\*TP\*CP\*GP\*TP\*TP\*TP\*CP\*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	F	20	Total	C	N	O	P	0	0	0
			409	194	70	125	20			
2	H	27	Total	C	N	O	P	0	0	0
			551	264	87	173	27			
2	J	23	Total	C	N	O	P	0	0	0
			471	224	79	145	23			
2	L	24	Total	C	N	O	P	0	0	0
			491	234	81	152	24			

- Molecule 3 is a DNA chain called DNA (5'-D(P\*CP\*GP\*AP\*AP\*AP\*CP\*GP\*AP\*CP\*GP\*GP\*CP\*CP\*AP\*GP\*TP\*GP\*CP\*CP\*AP\*(DOC))-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	I	21	Total	C	N	O	P	0	0	0
			429	202	86	120	21			

- Molecule 4 is a protein called DNA polymerase III subunit alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	A	1164	Total	C	N	O	S	0	0	0
			9280	5918	1620	1714	28			
4	B	1167	Total	C	N	O	S	0	0	0
			9295	5926	1623	1716	30			
4	C	1166	Total	C	N	O	S	0	0	0
			9293	5930	1621	1714	28			
4	D	1185	Total	C	N	O	S	0	0	0
			9445	6026	1651	1738	30			

- Molecule 5 is a protein called DNA polymerase III subunit gamma/tau.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	M	138	Total	C	N	O	S	0	0	0
			1103	702	203	197	1			
5	N	135	Total	C	N	O	S	0	0	0
			1078	687	200	190	1			
5	O	135	Total	C	N	O	S	0	0	0
			1078	687	200	190	1			
5	P	138	Total	C	N	O	S	0	0	0
			1093	695	203	194	1			

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
M	367	HIS	-	expression tag	UNP A0A0M9ACL9
M	368	HIS	-	expression tag	UNP A0A0M9ACL9
M	369	HIS	-	expression tag	UNP A0A0M9ACL9
M	370	HIS	-	expression tag	UNP A0A0M9ACL9
M	371	HIS	-	expression tag	UNP A0A0M9ACL9
M	372	HIS	-	expression tag	UNP A0A0M9ACL9
M	373	LYS	GLU	conflict	UNP A0A0M9ACL9
M	541	MET	-	expression tag	UNP A0A0M9ACL9
M	542	PRO	-	expression tag	UNP A0A0M9ACL9
M	543	PRO	-	expression tag	UNP A0A0M9ACL9
N	367	HIS	-	expression tag	UNP A0A0M9ACL9
N	368	HIS	-	expression tag	UNP A0A0M9ACL9
N	369	HIS	-	expression tag	UNP A0A0M9ACL9
N	370	HIS	-	expression tag	UNP A0A0M9ACL9
N	371	HIS	-	expression tag	UNP A0A0M9ACL9
N	372	HIS	-	expression tag	UNP A0A0M9ACL9
N	373	LYS	GLU	conflict	UNP A0A0M9ACL9
N	541	MET	-	expression tag	UNP A0A0M9ACL9
N	542	PRO	-	expression tag	UNP A0A0M9ACL9

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
N	543	PRO	-	expression tag	UNP A0A0M9ACL9
O	367	HIS	-	expression tag	UNP A0A0M9ACL9
O	368	HIS	-	expression tag	UNP A0A0M9ACL9
O	369	HIS	-	expression tag	UNP A0A0M9ACL9
O	370	HIS	-	expression tag	UNP A0A0M9ACL9
O	371	HIS	-	expression tag	UNP A0A0M9ACL9
O	372	HIS	-	expression tag	UNP A0A0M9ACL9
O	373	LYS	GLU	conflict	UNP A0A0M9ACL9
O	541	MET	-	expression tag	UNP A0A0M9ACL9
O	542	PRO	-	expression tag	UNP A0A0M9ACL9
O	543	PRO	-	expression tag	UNP A0A0M9ACL9
P	367	HIS	-	expression tag	UNP A0A0M9ACL9
P	368	HIS	-	expression tag	UNP A0A0M9ACL9
P	369	HIS	-	expression tag	UNP A0A0M9ACL9
P	370	HIS	-	expression tag	UNP A0A0M9ACL9
P	371	HIS	-	expression tag	UNP A0A0M9ACL9
P	372	HIS	-	expression tag	UNP A0A0M9ACL9
P	373	LYS	GLU	conflict	UNP A0A0M9ACL9
P	541	MET	-	expression tag	UNP A0A0M9ACL9
P	542	PRO	-	expression tag	UNP A0A0M9ACL9
P	543	PRO	-	expression tag	UNP A0A0M9ACL9

- Molecule 6 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

- Molecule 7 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	1	Total Mg 1 1	0	0
7	B	1	Total Mg 1 1	0	0

Continued on next page...

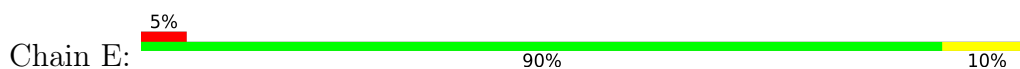
*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
7	C	1	Total 1	Mg 1	0	0
7	D	1	Total 1	Mg 1	0	0

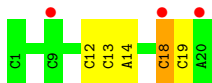
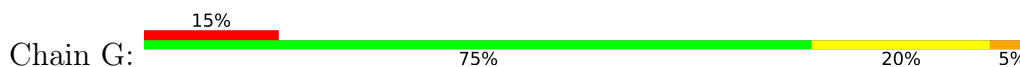
### 3 Residue-property plots [i](#)

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.

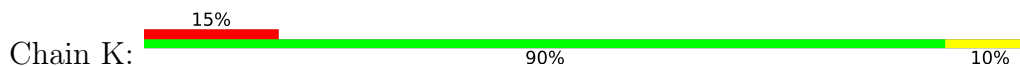
- Molecule 1: DNA (5'-D(P\*CP\*GP\*AP\*AP\*AP\*CP\*GP\*AP\*CP\*GP\*GP\*CP\*CP\*AP\*GP\*T P\*GP\*CP\*CP\*A)-3')



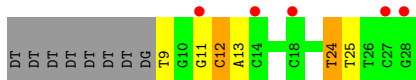
- Molecule 1: DNA (5'-D(P\*CP\*GP\*AP\*AP\*AP\*CP\*GP\*AP\*CP\*GP\*GP\*CP\*CP\*AP\*GP\*T P\*GP\*CP\*CP\*A)-3')



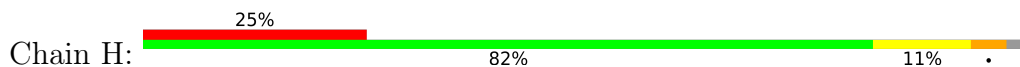
- Molecule 1: DNA (5'-D(P\*CP\*GP\*AP\*AP\*AP\*CP\*GP\*AP\*CP\*GP\*GP\*CP\*CP\*AP\*GP\*T P\*GP\*CP\*CP\*A)-3')



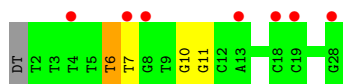
- Molecule 2: DNA (5'-D(\*TP\*TP\*TP\*TP\*TP\*TP\*TP\*GP\*TP\*GP\*GP\*CP\*AP\*CP\*TP\*GP \*GP\*CP\*CP\*GP\*TP\*CP\*GP\*TP\*TP\*TP\*CP\*G)-3')



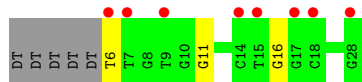
- Molecule 2: DNA (5'-D(\*TP\*TP\*TP\*TP\*TP\*TP\*TP\*GP\*TP\*GP\*GP\*CP\*AP\*CP\*TP\*GP \*GP\*CP\*CP\*GP\*TP\*CP\*GP\*TP\*TP\*TP\*CP\*G)-3')



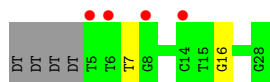
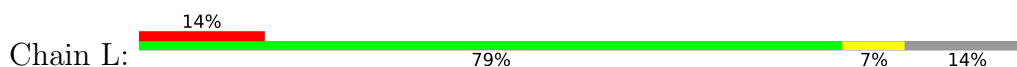




- Molecule 2: DNA (5'-D(\*TP\*TP\*TP\*TP\*TP\*TP\*TP\*TP\*GP\*TP\*GP\*GP\*CP\*AP\*CP\*TP\*GP\*GP\*CP\*CP\*GP\*TP\*CP\*GP\*TP\*TP\*TP\*CP\*G)-3')



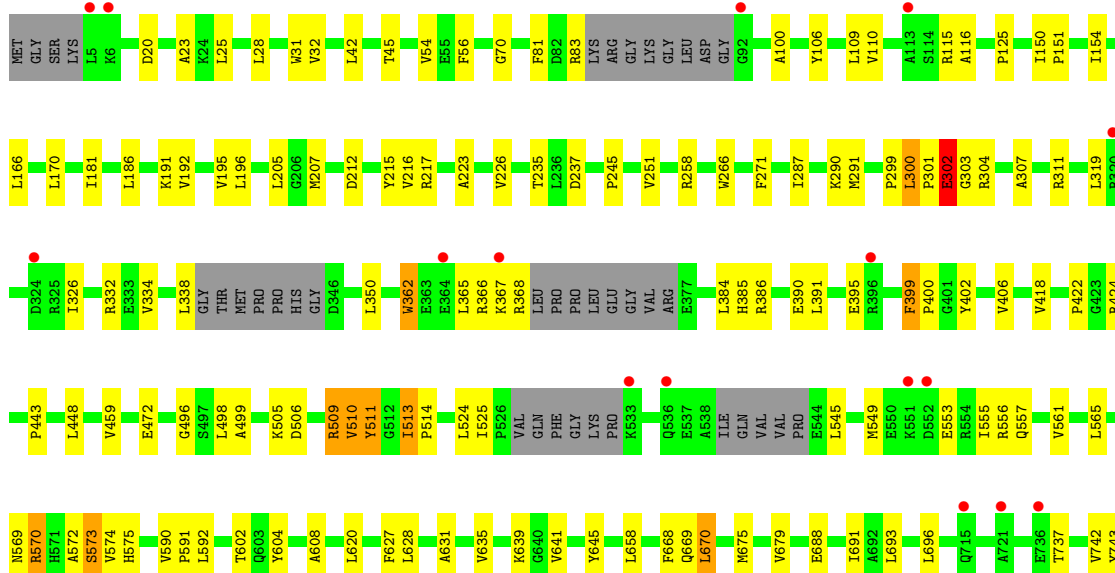
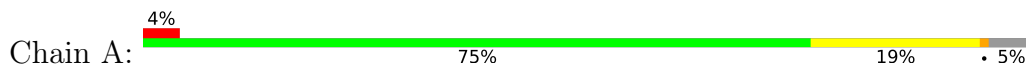
- Molecule 2: DNA (5'-D(\*TP\*TP\*TP\*TP\*TP\*TP\*TP\*TP\*GP\*TP\*GP\*GP\*CP\*AP\*CP\*TP\*GP\*GP\*CP\*CP\*GP\*TP\*CP\*GP\*TP\*TP\*TP\*CP\*G)-3')



- Molecule 3: DNA (5'-D(P\*CP\*GP\*AP\*AP\*AP\*CP\*GP\*AP\*CP\*GP\*GP\*CP\*CP\*AP\*GP\*T\*P\*GP\*CP\*CP\*AP\*(DOC))-3')

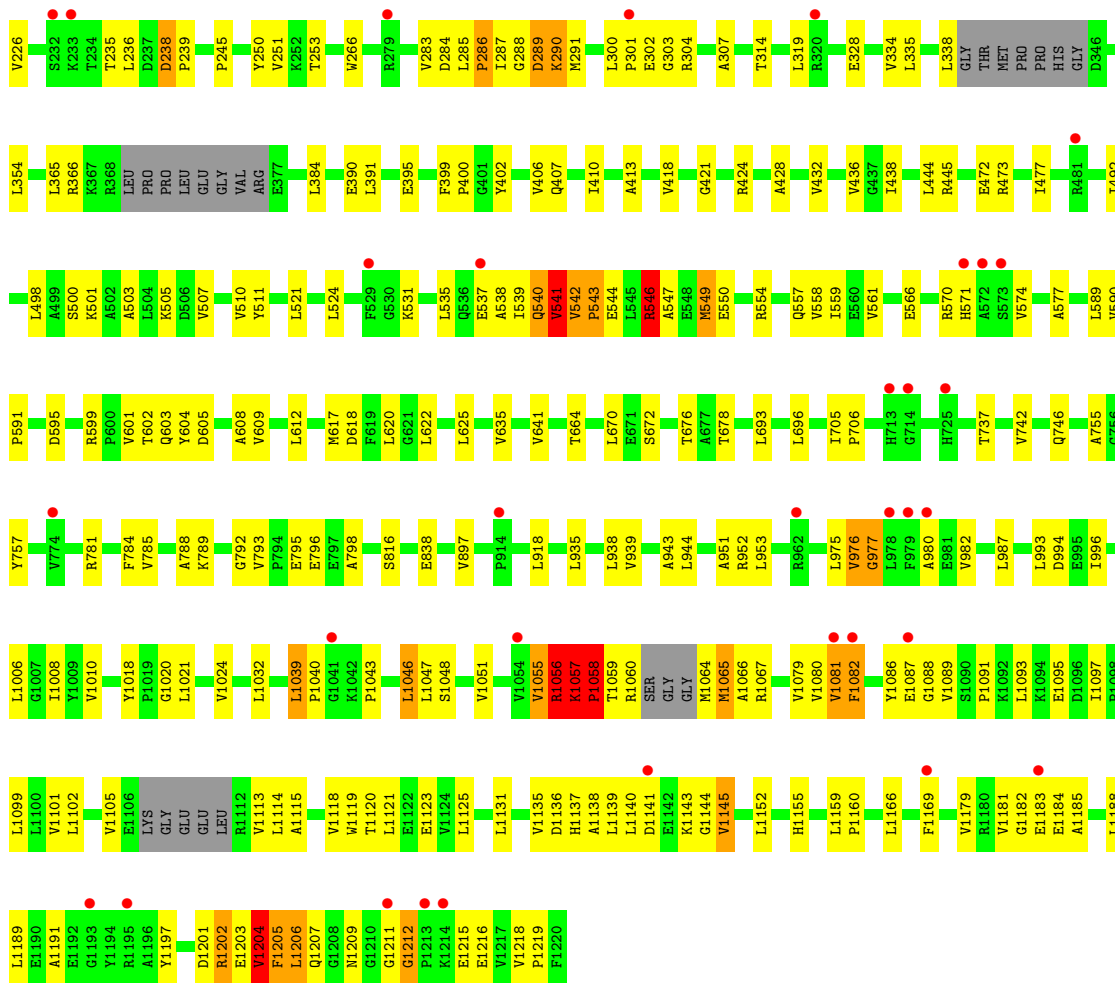


- Molecule 4: DNA polymerase III subunit alpha

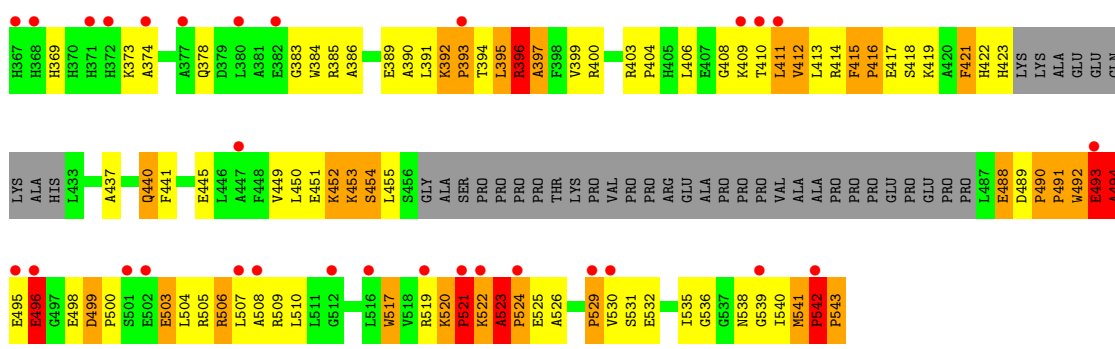
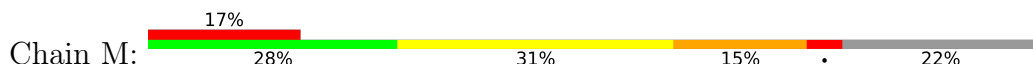




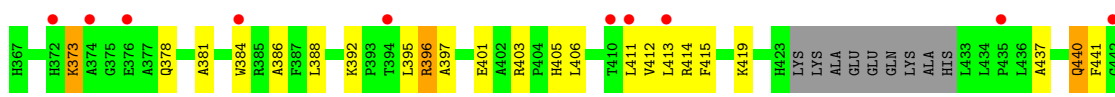




• Molecule 5: DNA polymerase III subunit gamma/tau



• Molecule 5: DNA polymerase III subunit gamma/tau





## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	188.53Å 94.97Å 204.08Å 90.00° 89.97° 90.00°	Depositor
Resolution (Å)	20.00 – 3.20 20.00 – 3.20	Depositor EDS
% Data completeness (in resolution range)	96.0 (20.00-3.20) 96.0 (20.00-3.20)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.04	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	15.90 (at 3.22Å)	Xtrriage
Refinement program	REFMAC 5.5.0109	Depositor
R, $R_{free}$	0.264 , 0.305 0.258 , 0.300	Depositor DCC
$R_{free}$ test set	5710 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	99.9	Xtrriage
Anisotropy	0.095	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.26 , 77.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.43$ , $\langle L^2 \rangle = 0.25$	Xtrriage
Estimated twinning fraction	0.018 for h,-k,-l	Xtrriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	45265	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	151.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.18% 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: DOC, ZN, MG

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	E	0.31	0/462	0.78	0/710
1	G	0.34	0/462	0.81	1/710 (0.1%)
1	K	0.30	0/462	0.80	1/710 (0.1%)
2	F	0.31	0/456	0.88	2/702 (0.3%)
2	H	0.41	0/613	0.84	1/945 (0.1%)
2	J	0.34	0/525	0.85	1/809 (0.1%)
2	L	0.38	0/547	0.80	0/843
3	I	0.33	0/462	0.80	1/710 (0.1%)
4	A	0.50	2/9466 (0.0%)	0.60	5/12781 (0.0%)
4	B	0.50	1/9480 (0.0%)	0.61	2/12800 (0.0%)
4	C	0.51	3/9481 (0.0%)	0.61	3/12805 (0.0%)
4	D	0.52	1/9638 (0.0%)	0.62	10/13019 (0.1%)
5	M	0.85	7/1133 (0.6%)	1.06	11/1528 (0.7%)
5	N	0.63	0/1108	0.80	3/1493 (0.2%)
5	O	0.69	0/1108	0.87	5/1493 (0.3%)
5	P	0.82	2/1123 (0.2%)	1.05	7/1513 (0.5%)
All	All	0.53	16/46526 (0.0%)	0.67	53/63571 (0.1%)

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

Mol	Chain	#Chirality outliers	#Planarity outliers
4	A	0	1
4	B	0	1
4	C	0	3
4	D	0	4
5	M	0	7
5	N	0	1
5	O	0	2

*Continued on next page...*

Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
5	P	0	7
All	All	0	26

The worst 5 of 16 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	P	543	PRO	N-CD	5.93	1.56	1.47
5	M	524	PRO	N-CD	5.75	1.55	1.47
5	M	416	PRO	N-CD	5.70	1.55	1.47
4	A	31	TRP	CD2-CE2	5.58	1.48	1.41
4	D	1058	PRO	N-CD	5.57	1.55	1.47

The worst 5 of 53 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	212	ASP	CB-CG-OD1	9.95	127.25	118.30
2	F	24	DT	P-O3'-C3'	9.83	131.50	119.70
5	P	541	MET	C-N-CD	8.76	146.79	128.40
5	P	415	PHE	C-N-CD	8.45	146.15	128.40
5	M	541	MET	C-N-CD	8.39	146.02	128.40

There are no chirality outliers.

5 of 26 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	A	570	ARG	Peptide
4	B	1084	ARG	Sidechain
4	C	1080	VAL	Peptide
4	C	972	ARG	Mainchain
4	C	974	GLY	Peptide

## 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	E	411	0	222	2	0

Continued on next page...



Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	G	411	0	222	5	0
1	K	411	0	222	1	0
2	F	409	0	227	5	0
2	H	551	0	310	4	0
2	J	471	0	262	2	0
2	L	491	0	274	2	0
3	I	429	0	233	5	0
4	A	9280	0	9298	227	2
4	B	9295	0	9323	298	3
4	C	9293	0	9318	274	0
4	D	9445	0	9488	387	2
5	M	1103	0	1085	176	0
5	N	1078	0	1065	134	1
5	O	1078	0	1065	143	0
5	P	1093	0	1078	208	2
6	A	3	0	0	0	0
6	B	3	0	0	0	0
6	C	3	0	0	0	0
6	D	3	0	0	0	0
7	A	1	0	0	0	0
7	B	1	0	0	0	0
7	C	1	0	0	0	0
7	D	1	0	0	0	0
All	All	45265	0	43692	1851	7

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 21.

The worst 5 of 1851 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)
5:P:528:GLU:CG	5:P:529:PRO:HD3	1.29	1.53
5:O:387:PHE:CE2	5:O:441:PHE:CE2	2.03	1.47
4:D:1080:VAL:HG11	4:D:1082:PHE:CE2	1.50	1.45
5:P:523:ALA:CB	5:P:524:PRO:HA	1.38	1.43
4:D:1080:VAL:HA	4:D:1081:VAL:CG2	1.47	1.41

The worst 5 of 7 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)
4:D:175:ASP:OD1	5:P:414:ARG:NH1[2_655]	1.84	0.36
4:A:115:ARG:NE	4:A:976:VAL:CG1[2_555]	1.90	0.30
4:A:115:ARG:NH2	4:A:976:VAL:CG2[2_555]	1.95	0.25
4:B:363:GLU:OE2	4:B:969:GLU:OE2[1_545]	2.01	0.19
4:D:160:ASP:N	5:N:454:SER:OG[1_545]	2.01	0.19

## 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	
4	A	1146/1220 (94%)	1036 (90%)	87 (8%)	23 (2%)	7	38
4	B	1147/1220 (94%)	1051 (92%)	76 (7%)	20 (2%)	9	42
4	C	1148/1220 (94%)	1056 (92%)	79 (7%)	13 (1%)	14	51
4	D	1173/1220 (96%)	1062 (90%)	94 (8%)	17 (1%)	11	46
5	M	132/177 (75%)	84 (64%)	31 (24%)	17 (13%)	0	1
5	N	129/177 (73%)	76 (59%)	28 (22%)	25 (19%)	0	0
5	O	129/177 (73%)	77 (60%)	28 (22%)	24 (19%)	0	0
5	P	132/177 (75%)	80 (61%)	33 (25%)	19 (14%)	0	1
All	All	5136/5588 (92%)	4522 (88%)	456 (9%)	158 (3%)	4	26

5 of 158 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	A	506	ASP
4	A	1081	VAL
4	A	1201	ASP
4	A	1204	VAL
4	B	301	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
4	A	965/1009 (96%)	958 (99%)	7 (1%)	84 94
4	B	968/1009 (96%)	946 (98%)	22 (2%)	50 78
4	C	968/1009 (96%)	949 (98%)	19 (2%)	55 80
4	D	984/1009 (98%)	961 (98%)	23 (2%)	50 78
5	M	112/144 (78%)	102 (91%)	10 (9%)	9 35
5	N	109/144 (76%)	104 (95%)	5 (5%)	27 63
5	O	109/144 (76%)	97 (89%)	12 (11%)	6 26
5	P	110/144 (76%)	89 (81%)	21 (19%)	1 8
All	All	4325/4612 (94%)	4206 (97%)	119 (3%)	43 74

5 of 119 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	D	976	VAL
5	P	522	LYS
5	M	395	LEU
5	P	520	LYS
5	P	543	PRO

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 26 such sidechains are listed below:

Mol	Chain	Res	Type
4	D	411	ASN
4	D	799	ASN
5	O	369	HIS
4	D	540	GLN
4	D	896	ASN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	DOC	I	21	3,2	16,19,20	1.02	1 (6%)	20,26,29	1.37	4 (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	DOC	I	21	3,2	-	1/7/18/19	0/2/2/2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	I	21	DOC	C6-C5	2.16	1.40	1.35

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	I	21	DOC	C3'-C2'-C1'	2.96	106.20	102.78
3	I	21	DOC	O4'-C1'-N1	2.44	112.22	107.86
3	I	21	DOC	O4'-C4'-C3'	2.35	108.71	104.80
3	I	21	DOC	O2-C2-N3	-2.28	118.62	122.33

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	I	21	DOC	C3'-C4'-C5'-O5'

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	I	21	DOC	1	0

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 16 ligands modelled in this entry, 16 are monoatomic - leaving 0 for Mogul analysis.

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	E	20/20 (100%)	0.86	1 (5%) 28 16	159, 177, 203, 209	0
1	G	20/20 (100%)	0.82	3 (15%) 2 1	160, 175, 204, 216	0
1	K	20/20 (100%)	0.94	3 (15%) 2 1	157, 168, 206, 213	0
2	F	20/28 (71%)	1.38	5 (25%) 0 0	165, 187, 208, 217	0
2	H	27/28 (96%)	1.25	7 (25%) 0 0	145, 168, 184, 186	0
2	J	23/28 (82%)	1.60	8 (34%) 0 0	148, 171, 192, 201	0
2	L	24/28 (85%)	1.41	4 (16%) 1 1	150, 176, 197, 204	0
3	I	20/21 (95%)	1.42	5 (25%) 0 0	198, 224, 252, 255	0
4	A	1164/1220 (95%)	0.19	50 (4%) 35 22	109, 140, 177, 212	0
4	B	1167/1220 (95%)	0.21	48 (4%) 37 24	107, 140, 174, 218	0
4	C	1166/1220 (95%)	0.22	46 (3%) 39 25	108, 148, 176, 203	0
4	D	1185/1220 (97%)	0.16	39 (3%) 46 30	111, 145, 175, 230	0
5	M	138/177 (77%)	1.09	30 (21%) 0 0	159, 176, 196, 209	0
5	N	135/177 (76%)	0.87	14 (10%) 6 4	143, 169, 187, 203	0
5	O	135/177 (76%)	1.03	19 (14%) 2 2	155, 181, 206, 227	0
5	P	138/177 (77%)	0.95	22 (15%) 1 1	139, 177, 202, 226	0
All	All	5402/5781 (93%)	0.31	304 (5%) 24 13	107, 147, 186, 255	0

The worst 5 of 304 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	P	446	LEU	9.6
4	D	714	GLY	6.0
4	D	1087	GLU	5.8
4	D	978	LEU	5.4
4	B	1085	ALA	5.3

## 6.2 Non-standard residues in protein, DNA, RNA chains [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
3	DOC	I	21	18/19	0.77	0.44	173,185,229,232	0

## 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
6	ZN	C	1301	1/1	0.76	0.14	197,197,197,197	0
7	MG	C	1304	1/1	0.79	0.07	177,177,177,177	0
6	ZN	D	1302	1/1	0.83	0.16	159,159,159,159	0
6	ZN	C	1303	1/1	0.83	0.20	179,179,179,179	0
7	MG	B	1304	1/1	0.84	0.27	132,132,132,132	0
6	ZN	A	1302	1/1	0.85	0.20	168,168,168,168	0
7	MG	A	1304	1/1	0.85	0.17	156,156,156,156	0
6	ZN	B	1302	1/1	0.88	0.15	141,141,141,141	0
6	ZN	C	1302	1/1	0.89	0.11	158,158,158,158	0
6	ZN	B	1303	1/1	0.90	0.21	166,166,166,166	0
6	ZN	D	1303	1/1	0.91	0.14	176,176,176,176	0
6	ZN	A	1301	1/1	0.91	0.17	187,187,187,187	0
6	ZN	D	1301	1/1	0.93	0.11	233,233,233,233	0
6	ZN	A	1303	1/1	0.94	0.10	141,141,141,141	0
6	ZN	B	1301	1/1	0.95	0.14	163,163,163,163	0
7	MG	D	1304	1/1	0.95	0.28	139,139,139,139	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.