



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 15, 2017 – 03:31 am GMT

PDB ID : 4FF1  
Title : N4 mini-vRNAP transcription initiation complex, 1 min after soaking GTP, ATP and Mn  
Authors : Murakami, K.S.; Basu, R.S.  
Deposited on : 2012-05-30  
Resolution : 2.47 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Mogul : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : trunk28620  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : recalc28949

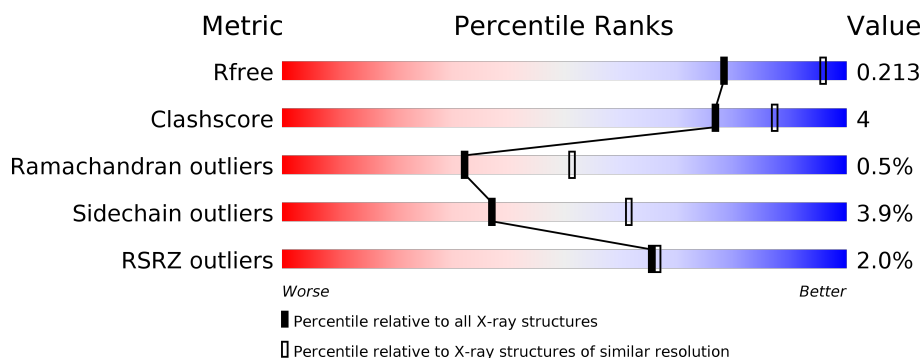
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.47 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))
$R_{free}$	100719	4719 (2.50-2.46)
Clashscore	112137	5483 (2.50-2.46)
Ramachandran outliers	110173	5388 (2.50-2.46)
Sidechain outliers	110143	5390 (2.50-2.46)
RSRZ outliers	101464	4754 (2.50-2.46)

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	1118	<div> <div>2%</div> <div> <div></div> <div>85%</div> <div>11%</div> <div>..</div> </div> </div>
1	B	1118	<div> <div>2%</div> <div> <div></div> <div>87%</div> <div>10%</div> <div>..</div> </div> </div>
2	C	36	<div> <div>47%</div> <div>8%</div> <div>44%</div> </div>
2	D	36	<div> <div>44%</div> <div>11%</div> <div>44%</div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 18400 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 Virion RNA polymerase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1095	Total	C	N	O	S	0	0	0
			8454	5306	1435	1672	41			
1	B	1095	Total	C	N	O	S	0	0	0
			8454	5306	1435	1672	41			

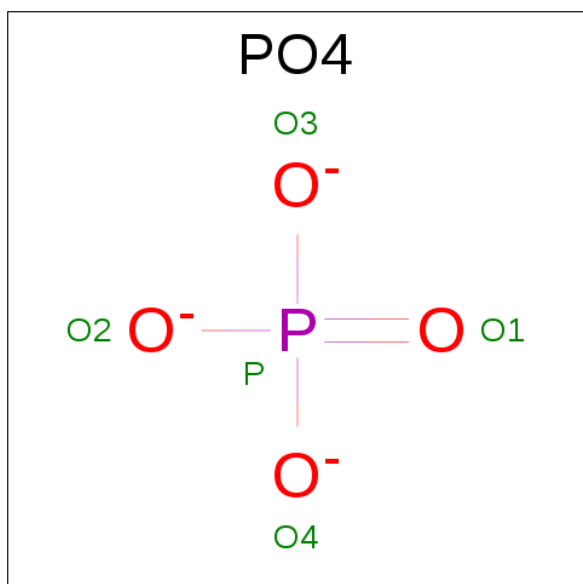
There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-11	MET	-	EXPRESSION TAG	UNP Q859P9
A	-10	GLY	-	EXPRESSION TAG	UNP Q859P9
A	-9	GLY	-	EXPRESSION TAG	UNP Q859P9
A	-8	SER	-	EXPRESSION TAG	UNP Q859P9
A	-7	HIS	-	EXPRESSION TAG	UNP Q859P9
A	-6	HIS	-	EXPRESSION TAG	UNP Q859P9
A	-5	HIS	-	EXPRESSION TAG	UNP Q859P9
A	-4	HIS	-	EXPRESSION TAG	UNP Q859P9
A	-3	HIS	-	EXPRESSION TAG	UNP Q859P9
A	-2	HIS	-	EXPRESSION TAG	UNP Q859P9
A	-1	ARG	-	EXPRESSION TAG	UNP Q859P9
A	0	SER	-	EXPRESSION TAG	UNP Q859P9
B	-11	MET	-	EXPRESSION TAG	UNP Q859P9
B	-10	GLY	-	EXPRESSION TAG	UNP Q859P9
B	-9	GLY	-	EXPRESSION TAG	UNP Q859P9
B	-8	SER	-	EXPRESSION TAG	UNP Q859P9
B	-7	HIS	-	EXPRESSION TAG	UNP Q859P9
B	-6	HIS	-	EXPRESSION TAG	UNP Q859P9
B	-5	HIS	-	EXPRESSION TAG	UNP Q859P9
B	-4	HIS	-	EXPRESSION TAG	UNP Q859P9
B	-3	HIS	-	EXPRESSION TAG	UNP Q859P9
B	-2	HIS	-	EXPRESSION TAG	UNP Q859P9
B	-1	ARG	-	EXPRESSION TAG	UNP Q859P9
B	0	SER	-	EXPRESSION TAG	UNP Q859P9

- Molecule 2 is a DNA chain called Bacteriophag N4 P2 promoter.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	20	Total	C	N	O	P	0	0	0
			413	196	80	117	20			
2	D	20	Total	C	N	O	P	0	0	0
			413	196	80	117	20			

- Molecule 3 is PHOSPHATE ION (three-letter code: PO4) (formula: O<sub>4</sub>P).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	P	0	0
			5	4	1		
3	B	1	Total	O	P	0	0
			5	4	1		

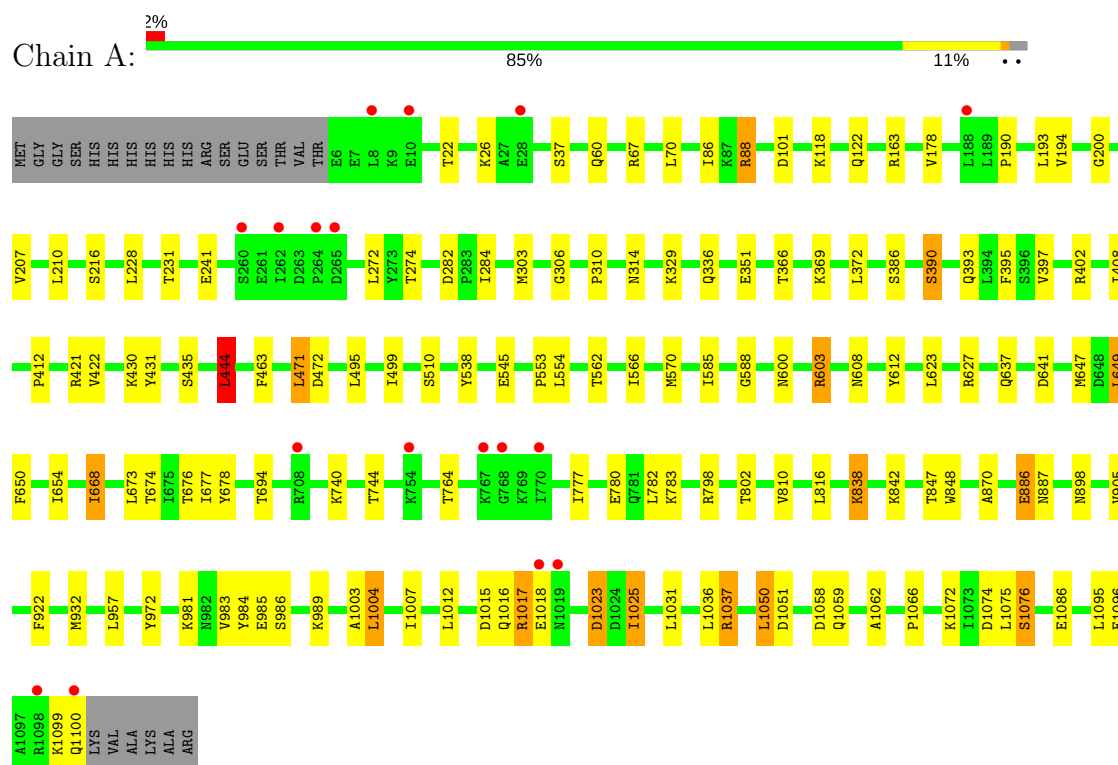
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	283	Total	O	0	0
			283	283		
4	C	15	Total	O	0	0
			15	15		
4	B	332	Total	O	0	0
			332	332		
4	D	26	Total	O	0	0
			26	26		

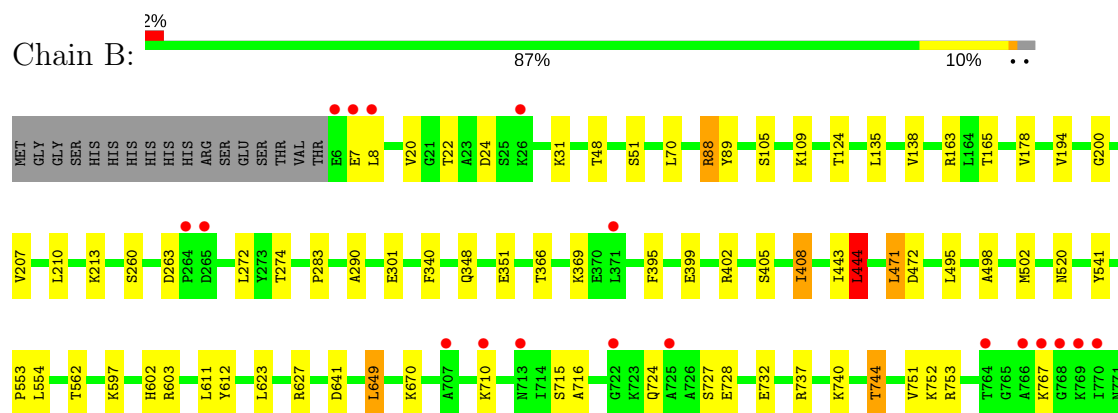
### 3 Residue-property plots [i](#)

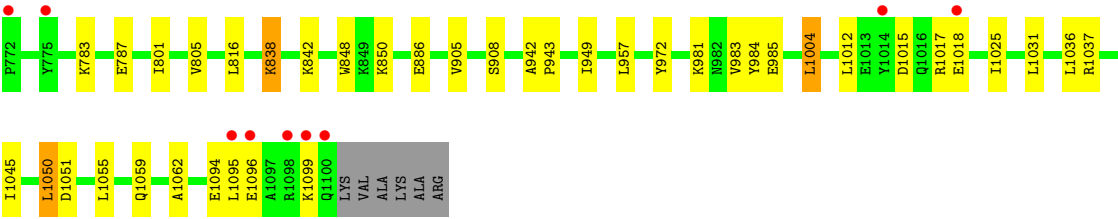
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are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

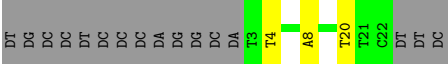


#### • Molecule 1: Virion RNA polymerase





● Molecule 2: Bacteriophag N4 P2 promoter



● Molecule 2: Bacteriophag N4 P2 promoter



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	82.16Å 111.50Å 277.10Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.84 – 2.47 47.84 – 2.46	Depositor EDS
% Data completeness (in resolution range)	95.2 (47.84-2.47) 95.2 (47.84-2.46)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.86 (at 2.45Å)	Xtriage
Refinement program	PHENIX (phenix.refine: dev_1042)	Depositor
R, $R_{free}$	0.153 , 0.211 0.152 , 0.213	Depositor DCC
$R_{free}$ test set	4378 reflections (4.98%)	DCC
Wilson B-factor (Å <sup>2</sup> )	31.7	Xtriage
Anisotropy	0.198	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 44.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	18400	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.93% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: PO4

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	0.38	0/8583	0.53	1/11609 (0.0%)
1	B	0.38	0/8583	0.54	1/11609 (0.0%)
2	C	0.63	0/464	1.25	3/714 (0.4%)
2	D	0.74	1/464 (0.2%)	1.22	1/714 (0.1%)
All	All	0.40	1/18094 (0.0%)	0.60	6/24646 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	15	DG	C3'-O3'	-5.18	1.37	1.44

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	444	LEU	CA-CB-CG	6.07	129.27	115.30
2	C	20	DT	N3-C4-O4	6.05	123.53	119.90
1	B	444	LEU	CA-CB-CG	5.73	128.48	115.30
2	C	4	DT	C5-C4-O4	-5.54	121.02	124.90
2	C	4	DT	N3-C4-O4	5.53	123.22	119.90
2	D	3	DT	O4'-C4'-C3'	-5.22	102.41	104.50

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	8454	0	8479	75	0
1	B	8454	0	8479	47	0
2	C	413	0	225	1	0
2	D	413	0	225	2	0
3	A	5	0	0	0	0
3	B	5	0	0	0	0
4	A	283	0	0	4	0
4	B	332	0	0	1	0
4	C	15	0	0	0	0
4	D	26	0	0	1	0
All	All	18400	0	17408	122	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 4.

All (122) 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:B:1096:GLU:HA	1:B:1099:LYS:HG2	1.68	0.74
1:A:674:THR:O	1:A:677:ILE:HG12	1.94	0.67
1:A:627:ARG:NH2	1:A:641:ASP:OD1	2.27	0.66
1:A:207:VAL:HG11	1:A:905:VAL:HG21	1.78	0.66
1:A:816:LEU:HD13	1:A:983:VAL:HG21	1.80	0.64
1:B:1012:LEU:O	1:B:1017:ARG:NH1	2.30	0.63
1:A:310:PRO:HG2	1:A:336:GLN:HB3	1.82	0.62
1:A:178:VAL:HG21	1:A:194:VAL:HA	1.83	0.61
1:B:597:LYS:HE2	1:B:602:HIS:HB2	1.83	0.61
1:B:627:ARG:NH1	1:B:641:ASP:OD1	2.35	0.59
1:A:422:VAL:HG12	1:A:922:PHE:HA	1.85	0.58
1:B:740:LYS:O	1:B:744:THR:HG22	2.03	0.58
1:A:627:ARG:NH1	1:A:637:GLN:OE1	2.36	0.58
1:B:886:GLU:O	2:D:8:DA:H4'	2.03	0.58
1:A:1096:GLU:HA	1:A:1099:LYS:HG2	1.84	0.58
1:A:981:LYS:O	1:A:985:GLU:HG2	2.03	0.57
1:B:348:GLN:OE1	4:B:1376:HOH:O	2.18	0.57
1:B:981:LYS:O	1:B:985:GLU:HG2	2.05	0.57
1:B:88:ARG:HD2	1:B:283:PRO:HD2	1.86	0.56
1:A:802:THR:HG23	1:A:810:VAL:HG21	1.87	0.56
1:A:88:ARG:HD3	1:A:282:ASP:OD1	2.05	0.56
1:A:677:ILE:HG13	1:A:678:TYR:CD2	2.41	0.55

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:647:MET:HB3	1:A:654:ILE:HG13	1.88	0.55
1:B:816:LEU:HD13	1:B:983:VAL:HG21	1.87	0.55
1:B:984:TYR:CE1	1:B:1037:ARG:HB2	2.42	0.54
1:A:200:GLY:HA2	1:A:274:THR:HG22	1.89	0.54
1:A:886:GLU:O	2:C:8:DA:H4'	2.07	0.54
1:A:306:GLY:HA2	1:A:412:PRO:HG2	1.90	0.54
1:A:351:GLU:HG3	1:A:395:PHE:CE2	2.43	0.54
1:A:444:LEU:HG	1:A:553:PRO:HB2	1.90	0.53
1:B:649:LEU:HD13	1:B:737:ARG:NH2	2.23	0.53
1:A:972:TYR:OH	1:A:1051:ASP:OD1	2.26	0.53
1:B:838:LYS:O	1:B:842:LYS:HG2	2.09	0.53
1:B:395:PHE:O	1:B:399:GLU:HG2	2.10	0.52
1:A:1099:LYS:HG3	1:A:1100:GLN:HG3	1.92	0.51
1:A:1074:ASP:OD1	1:A:1076:SER:OG	2.29	0.51
1:A:570:MET:HG2	1:A:585:ILE:HD11	1.94	0.50
1:B:728:GLU:O	1:B:732:GLU:HB2	2.12	0.50
1:B:178:VAL:HG21	1:B:194:VAL:HA	1.94	0.50
1:A:402:ARG:HA	1:A:408:ILE:HG22	1.92	0.50
1:A:1017:ARG:CG	1:A:1017:ARG:HH11	2.26	0.49
1:A:118:LYS:O	1:A:122:GLN:HG2	2.13	0.49
1:A:26:LYS:HG3	1:A:847:THR:HG21	1.95	0.49
1:A:842:LYS:HB3	1:A:848:TRP:CD2	2.48	0.48
1:A:740:LYS:O	1:A:744:THR:HG22	2.13	0.48
1:A:86:ILE:O	4:A:1499:HOH:O	2.20	0.48
1:A:1003:ALA:O	1:A:1007:ILE:HG13	2.13	0.48
1:A:932:MET:HE3	4:A:1341:HOH:O	2.13	0.48
1:B:972:TYR:OH	1:B:1051:ASP:OD1	2.27	0.48
1:A:1058:ASP:HB2	1:A:1066:PRO:HB3	1.96	0.47
1:A:554:LEU:HD23	1:A:957:LEU:HD11	1.97	0.47
1:A:386:SER:O	1:A:390:SER:HB2	2.15	0.47
1:A:984:TYR:CE1	1:A:1037:ARG:HB2	2.50	0.46
1:A:782:LEU:HD23	1:A:782:LEU:HA	1.81	0.46
1:A:430:LYS:O	1:A:435:SER:OG	2.31	0.46
1:A:603:ARG:NH1	1:A:608:ASN:OD1	2.49	0.46
1:A:463:PHE:HA	1:A:957:LEU:HD13	1.97	0.46
1:A:421:ARG:NH2	4:A:1370:HOH:O	2.33	0.46
1:B:816:LEU:CD1	1:B:983:VAL:HG21	2.46	0.46
1:A:566:ILE:HG13	1:A:588:GLY:HA3	1.97	0.46
1:A:393:GLN:HG2	1:A:431:TYR:HB2	1.97	0.45
1:A:1015:ASP:OD1	1:A:1016:GLN:HG3	2.16	0.45
1:A:887:ASN:ND2	4:A:1510:HOH:O	2.49	0.45

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:20:VAL:HG23	1:B:138:VAL:O	2.16	0.45
1:A:1099:LYS:HE2	1:A:1100:GLN:HG3	1.99	0.45
1:A:37:SER:HB3	1:A:231:THR:HG22	1.99	0.45
1:A:764:THR:HG21	1:A:780:GLU:HB3	1.99	0.45
1:A:562:THR:HG22	1:A:612:TYR:CZ	2.52	0.45
1:B:351:GLU:HG3	1:B:395:PHE:CE2	2.52	0.44
1:A:303:MET:HE1	1:A:397:VAL:HG13	1.99	0.44
1:B:1045:ILE:HG23	1:B:1094:GLU:HG2	2.00	0.44
1:A:1012:LEU:HD11	1:A:1025:ILE:HG22	2.00	0.44
1:A:612:TYR:CE2	1:A:673:LEU:HD23	2.52	0.44
1:A:676:THR:HG23	1:A:798:ARG:NH1	2.33	0.44
1:A:60:GLN:NE2	1:A:67:ARG:HD3	2.33	0.44
1:B:444:LEU:HG	1:B:553:PRO:HB2	1.99	0.44
1:B:783:LYS:O	1:B:787:GLU:HG2	2.17	0.44
1:B:340:PHE:HB3	1:B:443:ILE:HG22	1.98	0.43
1:B:1004:LEU:CD1	1:B:1025:ILE:HG12	2.49	0.43
1:A:314:ASN:HA	1:A:329:LYS:HE2	1.98	0.43
1:A:101:ASP:OD1	1:B:109:LYS:HE3	2.19	0.43
1:B:200:GLY:HA2	1:B:274:THR:HG22	2.00	0.43
1:B:260:SER:HA	1:B:263:ASP:O	2.18	0.43
1:B:612:TYR:CD2	1:B:670:LYS:HA	2.53	0.43
1:A:612:TYR:CD2	1:A:673:LEU:HD23	2.53	0.43
1:B:562:THR:HG22	1:B:612:TYR:CZ	2.54	0.43
1:B:751:VAL:HG21	1:B:753:ARG:NH1	2.33	0.43
1:A:1050:LEU:HA	1:A:1050:LEU:HD12	1.84	0.43
1:A:190:PRO:HG2	1:A:193:LEU:HB2	2.00	0.42
1:B:402:ARG:HA	1:B:408:ILE:HG23	2.01	0.42
1:B:801:ILE:O	1:B:805:VAL:HG22	2.19	0.42
1:A:303:MET:HB3	1:A:303:MET:HE2	1.87	0.42
1:A:654:ILE:HD11	1:A:668:ILE:HG21	2.02	0.42
1:A:471:LEU:HG	1:A:495:LEU:HD11	2.00	0.42
1:B:24:ASP:O	1:B:31:LYS:HG2	2.18	0.42
1:B:498:ALA:O	1:B:502:MET:HG2	2.19	0.42
1:B:207:VAL:HG11	1:B:905:VAL:HG21	2.01	0.42
1:A:1012:LEU:O	1:A:1017:ARG:NH1	2.35	0.42
1:A:668:ILE:HG13	1:A:668:ILE:H	1.67	0.42
1:B:1050:LEU:HD12	1:B:1050:LEU:HA	1.89	0.42
1:A:1004:LEU:HA	1:A:1004:LEU:HD22	1.85	0.42
1:A:1072:LYS:HD3	1:A:1072:LYS:HA	1.80	0.42
1:B:89:TYR:CZ	1:B:290:ALA:HB3	2.54	0.42
1:B:471:LEU:HG	1:B:495:LEU:HD11	2.01	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1023:ASP:OD2	1:A:1023:ASP:N	2.53	0.41
1:A:1075:LEU:HD21	1:A:1086:GLU:HG2	2.03	0.41
1:A:241:GLU:OE2	1:A:898:ASN:HB2	2.21	0.41
1:B:848:TRP:CH2	1:B:850:LYS:HA	2.56	0.41
2:D:22:DC:H1'	4:D:117:HOH:O	2.20	0.41
1:A:694:THR:HG22	1:A:777:ILE:HD12	2.01	0.41
1:A:816:LEU:CD1	1:A:983:VAL:HG21	2.48	0.41
1:B:165:THR:OG1	1:B:301:GLU:OE1	2.35	0.41
1:B:213:LYS:HE2	1:B:301:GLU:OE2	2.21	0.41
1:B:554:LEU:O	1:B:957:LEU:HG	2.20	0.41
1:B:135:LEU:O	1:B:138:VAL:HG22	2.21	0.41
1:A:870:ALA:HB2	1:A:989:LYS:HD3	2.03	0.41
1:B:611:LEU:HA	1:B:611:LEU:HD12	1.95	0.41
1:B:48:THR:O	1:B:51:SER:HB3	2.21	0.41
1:A:499:ILE:HD12	1:A:538:TYR:HD2	1.86	0.40
1:B:942:ALA:HA	1:B:943:PRO:HD3	1.84	0.40
1:A:649:LEU:HB3	1:A:650:PHE:CD1	2.56	0.40
1:A:838:LYS:O	1:A:842:LYS:HG2	2.20	0.40

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	1093/1118 (98%)	1065 (97%)	24 (2%)	4 (0%)	38	57
1	B	1093/1118 (98%)	1058 (97%)	28 (3%)	7 (1%)	28	46
All	All	2186/2236 (98%)	2123 (97%)	52 (2%)	11 (0%)	32	51

All (11) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1018	GLU
1	B	715	SER
1	B	727	SER
1	B	1062	ALA
1	A	1062	ALA
1	B	724	GLN
1	B	716	ALA
1	A	369	LYS
1	A	1017	ARG
1	B	369	LYS
1	B	1018	GLU

### 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	916/935 (98%)	880 (96%)	36 (4%)	37	61
1	B	916/935 (98%)	880 (96%)	36 (4%)	37	61
All	All	1832/1870 (98%)	1760 (96%)	72 (4%)	37	61

All (72) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	22	THR
1	A	70	LEU
1	A	88	ARG
1	A	163	ARG
1	A	210	LEU
1	A	216	SER
1	A	228	LEU
1	A	272	LEU
1	A	284	ILE
1	A	366	THR
1	A	372	LEU
1	A	390	SER
1	A	444	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	471	LEU
1	A	472	ASP
1	A	510	SER
1	A	545	GLU
1	A	600	ASN
1	A	603	ARG
1	A	623	LEU
1	A	649	LEU
1	A	668	ILE
1	A	783	LYS
1	A	838	LYS
1	A	886	GLU
1	A	986	SER
1	A	1004	LEU
1	A	1023	ASP
1	A	1025	ILE
1	A	1031	LEU
1	A	1036	LEU
1	A	1037	ARG
1	A	1050	LEU
1	A	1059	GLN
1	A	1076	SER
1	A	1095	LEU
1	B	7	GLU
1	B	8	LEU
1	B	22	THR
1	B	70	LEU
1	B	88	ARG
1	B	105	SER
1	B	124	THR
1	B	163	ARG
1	B	210	LEU
1	B	272	LEU
1	B	366	THR
1	B	405	SER
1	B	408	ILE
1	B	444	LEU
1	B	471	LEU
1	B	472	ASP
1	B	520	ASN
1	B	541	TYR
1	B	603	ARG

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	623	LEU
1	B	649	LEU
1	B	710	LYS
1	B	744	THR
1	B	752	LYS
1	B	767	LYS
1	B	838	LYS
1	B	908	SER
1	B	949	ILE
1	B	1004	LEU
1	B	1015	ASP
1	B	1031	LEU
1	B	1036	LEU
1	B	1050	LEU
1	B	1055	LEU
1	B	1059	GLN
1	B	1095	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (10) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	375	ASN
1	A	455	GLN
1	A	506	ASN
1	A	1038	ASN
1	B	122	GLN
1	B	455	GLN
1	B	494	ASN
1	B	506	ASN
1	B	671	ASN
1	B	863	GLN

### 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](#)

2 ligands are 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	PO4	A	1201	-	4,4,4	0.68	0	6,6,6	0.54	0
3	PO4	B	1201	-	4,4,4	0.67	0	6,6,6	0.54	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
3	PO4	A	1201	-	-	0/0/0/0	0/0/0/0
3	PO4	B	1201	-	-	0/0/0/0	0/0/0/0

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 ⓘ

There are no chain breaks in this entry.

## 6 Fit of model and data

### 6.1 Protein, DNA and RNA chains

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	1095/1118 (97%)	-0.39	17 (1%) 72 73	15, 32, 74, 110	0
1	B	1095/1118 (97%)	-0.33	27 (2%) 58 59	16, 31, 73, 136	0
2	C	20/36 (55%)	-0.66	0 100 100	32, 41, 65, 79	0
2	D	20/36 (55%)	-0.62	0 100 100	28, 38, 69, 93	0
All	All	2230/2308 (96%)	-0.37	44 (1%) 65 66	15, 32, 74, 136	0

All (44) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	1099	LYS	6.4
1	B	1100	GLN	5.8
1	B	770	ILE	4.6
1	B	769	LYS	4.2
1	B	767	LYS	3.7
1	A	1100	GLN	3.6
1	A	767	LYS	3.4
1	B	766	ALA	3.4
1	A	8	LEU	3.3
1	B	713	ASN	3.2
1	A	1018	GLU	3.2
1	B	768	GLY	3.2
1	B	725	ALA	3.1
1	A	754	LYS	2.9
1	B	710	LYS	2.9
1	B	772	PRO	2.9
1	B	1095	LEU	2.8
1	B	1096	GLU	2.8
1	A	262	ILE	2.7
1	B	264	PRO	2.7
1	A	28	GLU	2.7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	708	ARG	2.7
1	B	1018	GLU	2.6
1	B	775	TYR	2.6
1	B	1014	TYR	2.6
1	A	10	GLU	2.6
1	B	707	ALA	2.5
1	B	1098	ARG	2.5
1	B	722	GLY	2.5
1	A	265	ASP	2.4
1	A	1019	ASN	2.4
1	B	7	GLU	2.4
1	B	764	THR	2.3
1	A	1098	ARG	2.3
1	B	6	GLU	2.3
1	A	264	PRO	2.2
1	A	188	LEU	2.2
1	A	768	GLY	2.2
1	B	8	LEU	2.2
1	B	265	ASP	2.2
1	B	26	LYS	2.1
1	A	770	ILE	2.1
1	A	260	SER	2.0
1	B	371	LEU	2.0

## 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 [i](#)

There are no carbohydrates 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	PO4	A	1201	5/5	0.99	0.13	0.95	31,52,53,61	0
3	PO4	B	1201	5/5	0.98	0.12	0.62	46,50,62,62	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.