



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2026 – 05:55 AM UTC

PDB ID : 6P4F / pdb_00006p4f
Title : Crystal structure of the XPB-Bax1-forked DNA ternary complex
Authors : He, F.; Hilario, E.; Fan, L.
Deposited on : 2019-05-27
Resolution : 3.55 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

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	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

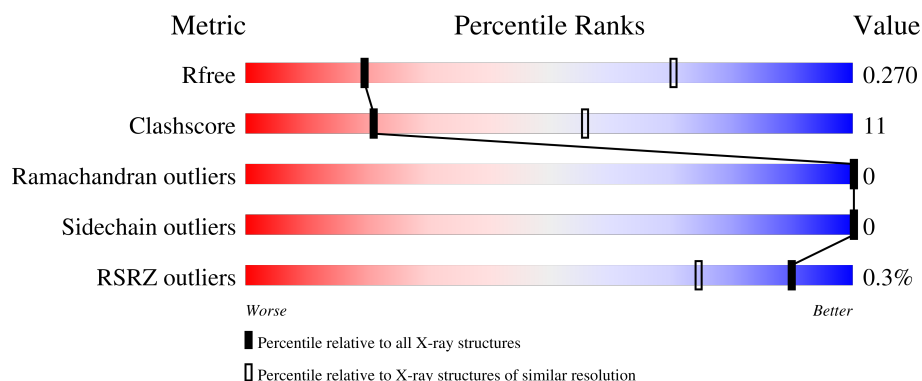
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

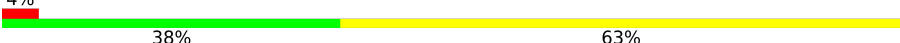
The reported resolution of this entry is 3.55 Å.

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}	180053	1410 (3.62-3.50)
Clashscore	190562	1480 (3.62-3.50)
Ramachandran outliers	187476	1440 (3.62-3.50)
Sidechain outliers	187428	1441 (3.62-3.50)
RSRZ outliers	180081	1409 (3.62-3.50)

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 of the lower bar indicate the fraction of residues that contain outliers for ≥ 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	440	 75% 24% .
1	C	440	 75% 24% .
2	B	373	 % 73% 27%
2	D	373	 78% 22%
3	E	24	 4% 38% 63%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	G	24	50% 50%
4	F	23	30% 65% .
4	H	23	43% 48% 9%

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 14208 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 DNA-dependent ATPase XPBII.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	435	Total	C	N	O	S	0	0	0
			3295	2137	537	618	3			
1	C	436	Total	C	N	O	S	0	0	0
			3266	2107	535	620	4			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	GLY	-	expression tag	UNP Q970I2
A	1	SER	MET	conflict	UNP Q970I2
C	0	GLY	-	expression tag	UNP Q970I2
C	1	SER	MET	conflict	UNP Q970I2

- Molecule 2 is a protein called Endonuclease Bax1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	373	Total	C	N	O	S	0	0	0
			2877	1898	464	512	3			
2	D	373	Total	C	N	O	S	0	0	0
			2871	1895	460	514	2			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	1	GLY	-	insertion	UNP Q970I1
D	1	GLY	-	insertion	UNP Q970I1

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	24	Total	C	N	O	P	0	0	0
			481	233	82	143	23			
3	G	24	Total	C	N	O	P	0	0	0
			481	233	82	143	23			

- Molecule 4 is a DNA chain called DNA (5'-D(P*TP*GP*TP*AP*GP*GP*TP*TP*TP*CP*CP*AP*TP*GP*TP*TP*GP*AP*GP*TP*CP*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	F	22	Total	C	N	O	P	0	0	0
			453	217	77	137	22			
4	H	21	Total	C	N	O	P	0	0	0
			433	207	75	130	21			

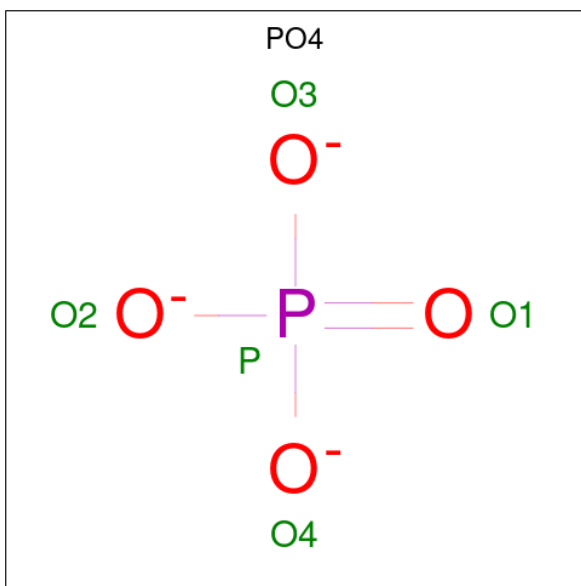
- Molecule 5 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	6	Total	Mg	0	0
			6	6		
5	B	6	Total	Mg	0	0
			6	6		
5	C	5	Total	Mg	0	0
			5	5		
5	D	4	Total	Mg	0	0
			4	4		
5	E	2	Total	Mg	0	0
			2	2		
5	F	1	Total	Mg	0	0
			1	1		
5	H	1	Total	Mg	0	0
			1	1		

- Molecule 6 is CHLORIDE ION (CCD ID: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	4	Total	Cl	0	0
			4	4		
6	B	5	Total	Cl	0	0
			5	5		
6	D	1	Total	Cl	0	0
			1	1		
6	G	1	Total	Cl	0	0
			1	1		

- Molecule 7 is PHOSPHATE ION (CCD ID: PO4) (formula: O_4P).

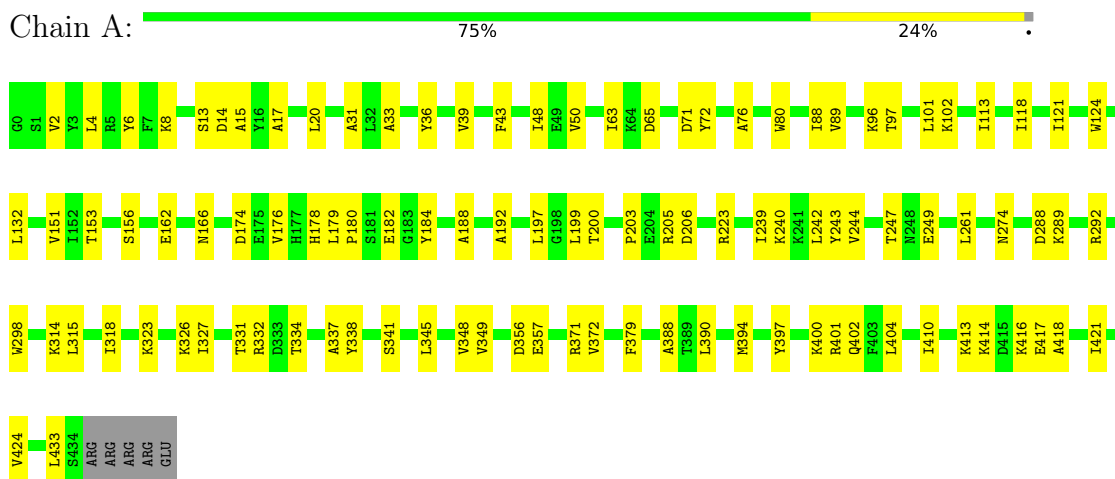


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	B	1	Total	O	P	0	0
			5	4	1		
7	D	1	Total	O	P	0	0
			5	4	1		
7	F	1	Total	O	P	0	0
			5	4	1		

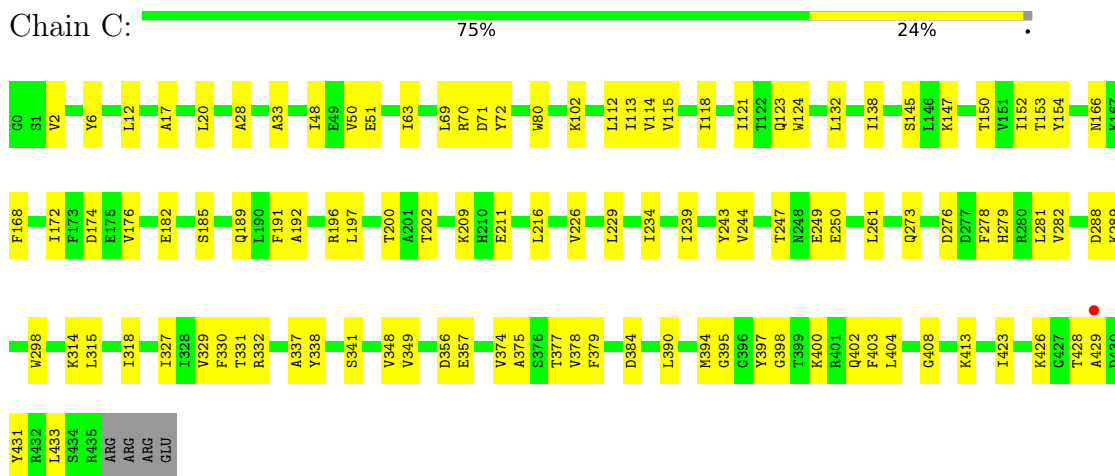
3 Residue-property plots

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.

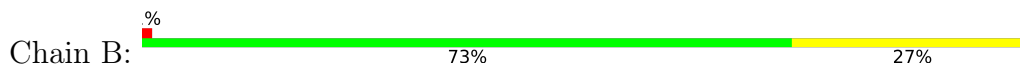
• Molecule 1: DNA-dependent ATPase XPBII

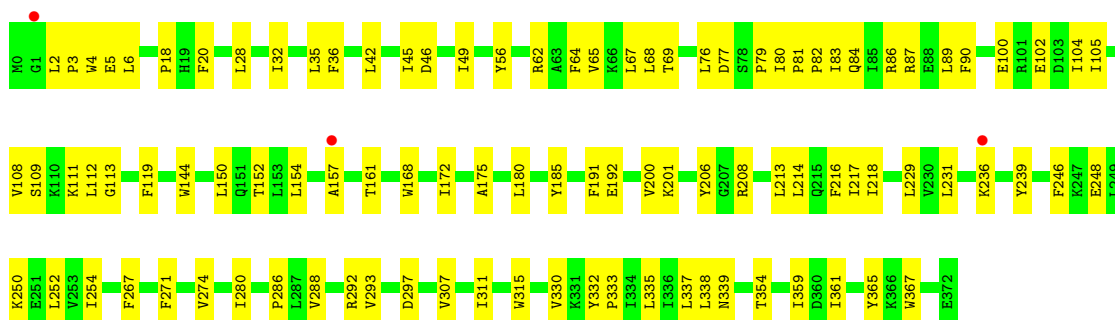


• Molecule 1: DNA-dependent ATPase XPBII



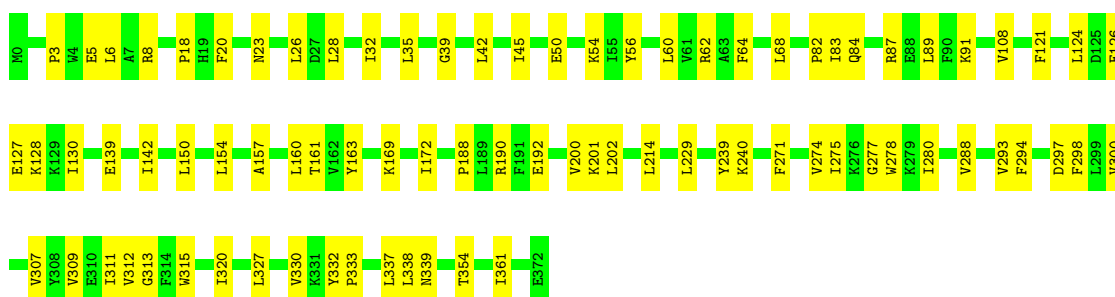
• Molecule 2: Endonuclease Bax1





• Molecule 2: Endonuclease Bax1

Chain D: 78% 22%



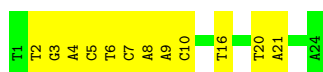
• Molecule 3: DNA (5'-D(*TP*TP*GP*AP*CP*TP*CP*AP*AP*CP*AP*TP*CP*CP*TP*TP*TP*GP*CP*TP*AP*CP*AP*A)-3')

Chain E: 4% 38% 63%



• Molecule 3: DNA (5'-D(*TP*TP*GP*AP*CP*TP*CP*AP*AP*CP*AP*TP*CP*CP*TP*TP*TP*GP*CP*TP*AP*CP*AP*A)-3')

Chain G: 50% 50%



• Molecule 4: DNA (5'-D(P*TP*GP*TP*AP*GP*GP*TP*TP*TP*CP*CP*AP*TP*GP*TP*T P*GP*AP*GP*TP*CP*A)-3')

Chain F: 30% 65%



• Molecule 4: DNA (5'-D(P*TP*GP*TP*AP*GP*GP*TP*TP*TP*CP*CP*AP*TP*GP*TP*T P*GP*AP*GP*TP*CP*A)-3')



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	214.69Å 92.15Å 172.13Å 90.00° 132.24° 90.00°	Depositor
Resolution (Å)	39.73 – 3.55 39.73 – 3.55	Depositor EDS
% Data completeness (in resolution range)	93.5 (39.73-3.55) 93.4 (39.73-3.55)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.22	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.51 (at 3.57Å)	Xtriage
Refinement program	PHENIX 1.15.2-3472	Depositor
R, R_{free}	0.250 , 0.272 0.250 , 0.270	Depositor DCC
R_{free} test set	1402 reflections (4.62%)	wwPDB-VP
Wilson B-factor (Å ²)	104.0	Xtriage
Anisotropy	0.128	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.25 , 79.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.43$, $\langle L^2 \rangle = 0.26$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	14208	wwPDB-VP
Average B, all atoms (Å ²)	108.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.09% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: MG, PO4, 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	0.17	0/3357	0.44	0/4567
1	C	0.17	0/3326	0.43	0/4526
2	B	0.17	0/2940	0.42	0/3999
2	D	0.16	0/2936	0.42	0/3998
3	E	0.24	0/537	0.47	0/825
3	G	0.23	0/537	0.43	0/825
4	F	0.25	0/506	0.48	0/780
4	H	0.24	0/484	0.44	0/746
All	All	0.18	0/14623	0.43	0/20266

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	3295	0	3201	67	0
1	C	3266	0	3127	72	0
2	B	2877	0	2763	69	0
2	D	2871	0	2737	55	0
3	E	481	0	274	13	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	G	481	0	274	10	0
4	F	453	0	252	16	0
4	H	433	0	240	14	0
5	A	6	0	0	0	0
5	B	6	0	0	0	0
5	C	5	0	0	0	0
5	D	4	0	0	0	0
5	E	2	0	0	0	0
5	F	1	0	0	0	0
5	H	1	0	0	0	0
6	A	4	0	0	0	0
6	B	5	0	0	0	0
6	D	1	0	0	0	0
6	G	1	0	0	0	0
7	B	5	0	0	0	0
7	D	5	0	0	0	0
7	F	5	0	0	1	0
All	All	14208	0	12868	296	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 11.

The worst 5 of 296 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:C:123:GLN:HE21	1:C:384:ASP:HB2	1.42	0.84
1:C:123:GLN:NE2	1:C:384:ASP:HB2	1.94	0.82
2:B:62:ARG:NH2	4:F:14:DA:OP2	2.13	0.81
3:E:4:DA:H61	4:F:22:DT:H3	1.25	0.80
1:C:341:SER:HB2	1:C:348:VAL:HG13	1.65	0.78

There are no symmetry-related clashes.

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	433/440 (98%)	404 (93%)	29 (7%)	0	100	100
1	C	434/440 (99%)	401 (92%)	33 (8%)	0	100	100
2	B	371/373 (100%)	334 (90%)	37 (10%)	0	100	100
2	D	371/373 (100%)	338 (91%)	33 (9%)	0	100	100
All	All	1609/1626 (99%)	1477 (92%)	132 (8%)	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	324/384 (84%)	324 (100%)	0	100	100
1	C	317/384 (83%)	317 (100%)	0	100	100
2	B	276/344 (80%)	276 (100%)	0	100	100
2	D	275/344 (80%)	275 (100%)	0	100	100
All	All	1192/1456 (82%)	1192 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	73	GLN
1	C	123	GLN

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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 39 ligands modelled in this entry, 36 are monoatomic - leaving 3 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	$\# Z > 2$
7	PO4	D	401	-	4,4,4	0.99	0	6,6,6	0.48	0
7	PO4	B	401	-	4,4,4	1.00	0	6,6,6	0.44	0
7	PO4	F	101	-	4,4,4	1.08	0	6,6,6	0.40	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.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	F	101	PO4	1	0

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 [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, 95th 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 ²)	Q<0.9
1	A	435/440 (98%)	-0.32	0 100 100	44, 98, 152, 179	0
1	C	436/440 (99%)	-0.31	1 (0%) 91 80	54, 104, 161, 190	0
2	B	373/373 (100%)	-0.34	3 (0%) 82 58	64, 109, 150, 187	0
2	D	373/373 (100%)	-0.39	0 100 100	57, 107, 153, 190	0
3	E	24/24 (100%)	0.31	1 (4%) 40 21	65, 143, 195, 198	0
3	G	24/24 (100%)	0.13	0 100 100	70, 156, 205, 232	0
4	F	22/23 (95%)	-0.03	0 100 100	61, 104, 183, 215	0
4	H	21/23 (91%)	0.06	0 100 100	70, 104, 196, 280	0
All	All	1708/1720 (99%)	-0.31	5 (0%) 90 74	44, 104, 159, 280	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	1	GLY	2.5
3	E	24	DA	2.4
2	B	157	ALA	2.4
2	B	236	LYS	2.1
1	C	429	ALA	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 oligosaccharides 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, 95th 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 ²)	Q<0.9
5	MG	E	101	1/1	0.46	0.13	86,86,86,86	0
6	CL	A	510	1/1	0.69	0.08	104,104,104,104	0
5	MG	A	503	1/1	0.74	0.10	67,67,67,67	0
5	MG	F	102	1/1	0.75	0.07	64,64,64,64	0
5	MG	A	501	1/1	0.79	0.23	57,57,57,57	0
6	CL	B	409	1/1	0.80	0.07	88,88,88,88	0
5	MG	C	503	1/1	0.81	0.10	83,83,83,83	0
5	MG	B	407	1/1	0.82	0.08	69,69,69,69	0
5	MG	D	405	1/1	0.83	0.09	70,70,70,70	0
5	MG	C	501	1/1	0.84	0.12	91,91,91,91	0
7	PO4	B	401	5/5	0.84	0.09	104,119,122,128	0
6	CL	B	410	1/1	0.85	0.07	113,113,113,113	0
6	CL	B	408	1/1	0.86	0.05	100,100,100,100	0
5	MG	A	505	1/1	0.86	0.13	58,58,58,58	0
6	CL	G	101	1/1	0.87	0.06	33,33,33,33	0
5	MG	C	502	1/1	0.88	0.04	77,77,77,77	0
6	CL	B	412	1/1	0.88	0.04	101,101,101,101	0
7	PO4	F	101	5/5	0.88	0.09	61,71,73,81	0
5	MG	D	402	1/1	0.89	0.03	66,66,66,66	0
5	MG	B	404	1/1	0.89	0.10	76,76,76,76	0
5	MG	C	505	1/1	0.89	0.11	69,69,69,69	0
5	MG	B	403	1/1	0.90	0.03	65,65,65,65	0
5	MG	B	402	1/1	0.90	0.03	49,49,49,49	0
5	MG	E	102	1/1	0.90	0.18	31,31,31,31	0
7	PO4	D	401	5/5	0.90	0.07	64,77,81,84	0
5	MG	D	404	1/1	0.90	0.06	69,69,69,69	0
5	MG	A	506	1/1	0.91	0.09	80,80,80,80	0
5	MG	H	101	1/1	0.91	0.13	62,62,62,62	0
6	CL	A	509	1/1	0.91	0.07	97,97,97,97	0
5	MG	A	502	1/1	0.91	0.04	35,35,35,35	0
5	MG	B	406	1/1	0.92	0.03	75,75,75,75	0
5	MG	D	403	1/1	0.92	0.14	31,31,31,31	0
6	CL	B	411	1/1	0.93	0.11	77,77,77,77	0
6	CL	A	508	1/1	0.93	0.03	75,75,75,75	0
5	MG	A	504	1/1	0.93	0.14	60,60,60,60	0
6	CL	A	507	1/1	0.94	0.08	108,108,108,108	0
6	CL	D	406	1/1	0.94	0.11	69,69,69,69	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($\text{\AA}^2$)	Q<0.9
5	MG	C	504	1/1	0.96	0.09	40,40,40,40	0
5	MG	B	405	1/1	0.97	0.04	54,54,54,54	0

6.5 Other polymers [i](#)

There are no such residues in this entry.