



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 11:28 PM GMT

PDB ID : 1X9N
Title : Crystal Structure of Human DNA Ligase I bound to 5'-adenylated, nicked DNA
Authors : Pascal, J.M.; O'Brien, P.J.; Tomkinson, A.E.; Ellenberger, T.
Deposited on : 2004-08-23
Resolution : 3.00 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
<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 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
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) : trunk26865

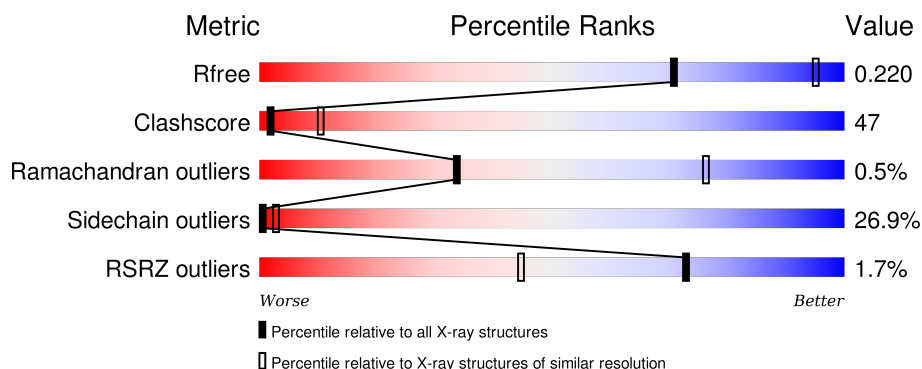
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.00 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}	91344	1578 (3.00-3.00)
Clashscore	102246	1912 (3.00-3.00)
Ramachandran outliers	100387	1853 (3.00-3.00)
Sidechain outliers	100360	1856 (3.00-3.00)
RSRZ outliers	91569	1592 (3.00-3.00)

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 ≥ 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	B	13	<div> <div>8%</div> <div>46%</div> <div>38%</div> <div>15%</div> </div>
2	C	15	<div> <div>7%</div> <div>7%</div> <div>13%</div> <div>40%</div> <div>40%</div> </div>
3	D	28	<div> <div>7%</div> <div>43%</div> <div>29%</div> <div>29%</div> </div>
4	A	688	<div> <div>%</div> <div>31%</div> <div>47%</div> <div>14%</div> <div>8%</div> </div>

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 5730 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 dideoxy terminated DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	11	Total	C	N	O	P	0	0	0
			222	107	40	65	10			

- Molecule 2 is a DNA chain called 5'-phosphorylated DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	9	Total	C	N	O	P	0	0	0
			187	88	35	55	9			

- Molecule 3 is a DNA chain called template DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	20	Total	C	N	O	P	0	0	0
			404	192	78	115	19			

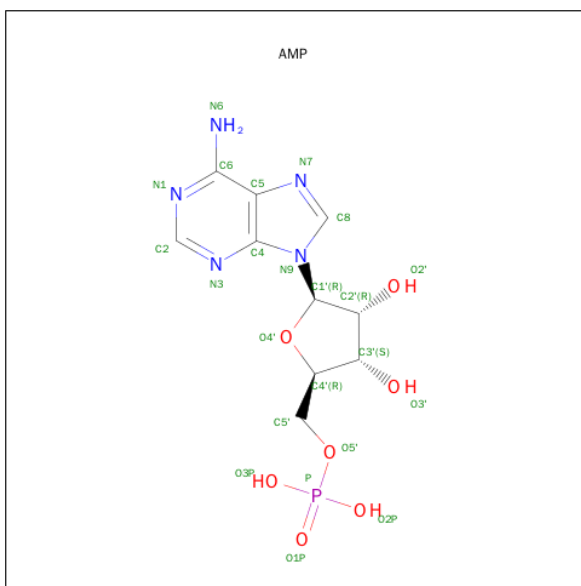
- Molecule 4 is a protein called DNA ligase I.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
4	A	632	Total	C	N	O	S	Se	0	0	0
			4894	3109	849	920	10	6			

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	232	MET	-	INITIATING METHIONINE	UNP P18858
A	308	MSE	MET	MODIFIED RESIDUE	UNP P18858
A	393	MSE	MET	MODIFIED RESIDUE	UNP P18858
A	480	MSE	MET	MODIFIED RESIDUE	UNP P18858
A	501	MSE	MET	MODIFIED RESIDUE	UNP P18858
A	543	MSE	MET	MODIFIED RESIDUE	UNP P18858
A	723	MSE	MET	MODIFIED RESIDUE	UNP P18858

- Molecule 5 is ADENOSINE MONOPHOSPHATE (three-letter code: AMP) (formula: $C_{10}H_{14}N_5O_7P$).

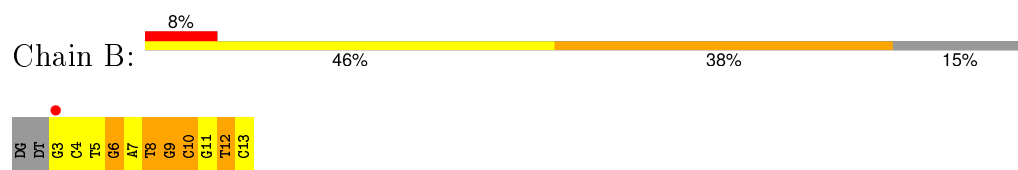


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	C	1	Total	C	N	O	P	0	0
			23	10	5	7	1		

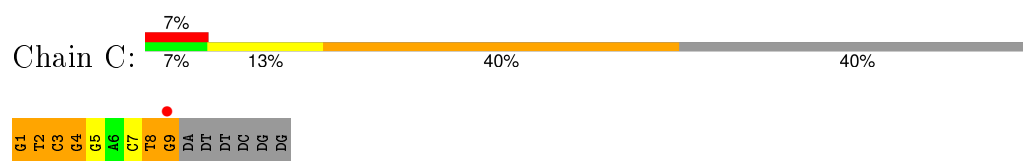
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

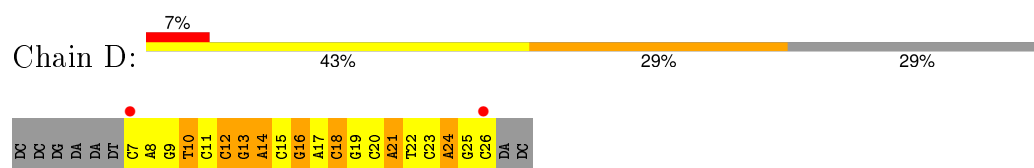
- Molecule 1: dideoxy terminated DNA



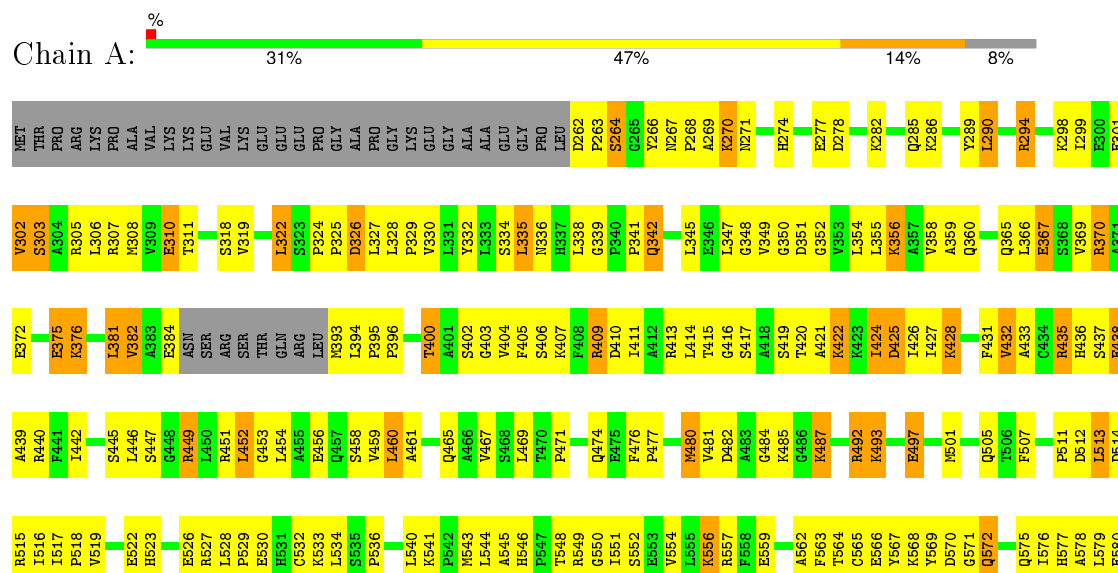
- Molecule 2: 5'-phosphorylated DNA



- Molecule 3: template DNA



- Molecule 4: DNA ligase I



GLY	V844	S782	Q713	D647	V584
GLU	K845	Y783	S714	A648	K585
ASP	C846	D784	V715	S649	I586
SER	A847	E785	K716	E850	
GLY	D848	D786	D717	I651	
SER	L849	S787	S718	Q652	R589
ASP	S850	E788	C719	V653	I590
PRO	L851	E789	E720	Q684	Q591
GLU	S852	L790	G721	V655	S592
ASP	P853	Q791	L722	C656	D593
THR	I854	A792	M723	L657	
TYR	Y855	I793	Y724	Y658	
	P856	C794	K725		K597
	A857	K795	T726		Y598
		L796	L727	D661	P599
	V862	G797		I662	D600
	D863	T798	T732	I663	I601
	S864	T799	Y733	Y664	I602
	D865	F800	Y734	L665	S603
	K866	S801	E735	N666	S604
	I868	D802	I736		I605
	S869		K737	L670	P606
	L870	L805	R738	V671	K607
			S739	E673	I608
	R871		R740	P674	K609
	F872	H808	I741	L675	I610
	P873	H809	M742	S676	P611
	R874	Q810	L743	R677	S612
	F875	S811	L744	R678	I613
	I876	L812	K745	R679	T614
	R877	K813	L746	Q680	S615
	V878	A814	K747	I681	I616
		L815	D748	L682	I617
		V816	Y749	R683	I618
	D881	P818	L750	E884	D619
	K882	S819	D751	N685	E621
	Q883	P820	G752	F686	A622
	P884	R821	V753	V687	I623
	E885	P822	G754	E688	A624
	Q886	Y823	D755	I689	I625
	A887	V824	T756		D626
	T888	R825	L757	F693	
	T889	I826	D758	V694	K629
	S890	D827	L759	F695	K630
	A891	G828	V760	A696	Q631
	V893	A829	I762	S698	I633
		V830		L699	P634
		I831	Y765	D700	F635
	C895	P832		T701	I636
	L896	D833		K702	V637
	Y897	H834	R768	D703	I638
	R898	W835	G769	I704	T639
	K899	L836	K770	E705	T640
	Q900	D837	R771	Q706	I641
	S901	P838		I707	I642
	GLN	S839	R774		
	ILE	A840			
	GLN	W841	L779	F710	R643
	ASN	I842	L780	L711	E644
	GLN	E843	A781	E712	E645
	GLN				V646

4 Data and refinement statistics

Property	Value	Source
Space group	P 63	Depositor
Cell constants a, b, c, α , β , γ	161.89 Å 161.89 Å 88.45 Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	20.00 – 3.00 19.99 – 3.00	Depositor EDS
% Data completeness (in resolution range)	50.0 (20.00-3.00) 99.0 (19.99-3.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.08	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.60 (at 2.98 Å)	Xtriage
Refinement program	REFMAC 5.1.80	Depositor
R, R_{free}	0.235 , 0.268 0.215 , 0.220	Depositor DCC
R_{free} test set	1287 reflections (5.16%)	DCC
Wilson B-factor (Å ²)	78.1	Xtriage
Anisotropy	0.568	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 79.5	EDS
Estimated twinning fraction	0.026 for h,-h-k,-l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	0 of 26244 reflections	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	5730	wwPDB-VP
Average B, all atoms (Å ²)	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.83% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: DOC, AMP

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	B	1.01	0/228	2.16	12/351 (3.4%)
2	C	1.40	3/209 (1.4%)	2.04	9/321 (2.8%)
3	D	1.15	0/453	1.95	19/696 (2.7%)
4	A	0.58	0/4987	0.82	15/6753 (0.2%)
All	All	0.71	3/5877 (0.1%)	1.12	55/8121 (0.7%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	1	DG	C3'-O3'	-7.01	1.34	1.44
2	C	3	DC	C3'-O3'	-5.51	1.36	1.44
2	C	4	DG	C3'-O3'	-5.22	1.37	1.44

The worst 5 of 55 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	8	DT	O4'-C1'-N1	18.87	121.21	108.00
3	D	16	DG	O4'-C4'-C3'	-9.17	100.50	106.00
3	D	15	DC	O4'-C4'-C3'	-8.89	100.67	106.00
3	D	11	DC	O4'-C1'-N1	-7.93	102.45	108.00
3	D	21	DA	O4'-C1'-N9	-7.78	102.56	108.00

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	B	222	0	126	18	0
2	C	187	0	102	15	0
3	D	404	0	224	30	0
4	A	4894	0	4925	468	0
5	C	23	0	12	3	0
All	All	5730	0	5389	520	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 47.

The worst 5 of 520 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)
4:A:480:MSE:CE	4:A:480:MSE:SE	2.15	1.44
4:A:548:THR:CG2	4:A:745:LEU:HD13	1.72	1.19
2:C:7:DC:H2"	2:C:8:DT:H5"	1.24	1.15
4:A:400:THR:HG22	4:A:403:GLY:H	1.08	1.14
4:A:551:ILE:CD1	4:A:707:ILE:HG22	1.79	1.13

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
4	A	628/688 (91%)	559 (89%)	66 (10%)	3 (0%)	34 76

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	A	748	ASP
4	A	753	VAL
4	A	644	LYS

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	525/576 (91%)	384 (73%)	141 (27%)	0 3

5 of 141 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	A	597	LYS
4	A	652	GLN
4	A	862	VAL
4	A	607	LYS
4	A	633	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
4	A	474	GLN
4	A	594	ASN
4	A	713	GLN
4	A	360	GLN
4	A	706	GLN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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	$\# Z > 2$
1	DOC	B	13	1,3	11,19,20	0.94	1 (9%)	14,26,29	1.70	3 (21%)

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
1	DOC	B	13	1,3	-	0/3/18/19	0/2/2/2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	13	DOC	C6-C5	-2.42	1.32	1.38

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	13	DOC	C2'-C1'-N1	-3.68	105.04	112.49
1	B	13	DOC	C3'-C2'-C1'	-3.08	99.26	102.71
1	B	13	DOC	C2-N3-C4	2.35	118.92	115.61

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	B	13	DOC	2	0

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

1 ligand 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	$\# Z > 2$
5	AMP	C	100	2	20,25,25	1.16	2 (10%)	22,38,38	2.40	1 (4%)

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
5	AMP	C	100	2	-	0/6/26/26	0/3/3/3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	C	100	AMP	C2-N1	2.28	1.38	1.33
5	C	100	AMP	C2-N3	3.45	1.38	1.32

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	C	100	AMP	N3-C2-N1	-10.48	120.87	128.89

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	C	100	AMP	3	0

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

6 Fit of model and data [i](#)

6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 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	B	10/13 (76%)	0.32	1 (10%) 9 4	44, 50, 65, 66	0
2	C	9/15 (60%)	-0.19	1 (11%) 7 3	40, 46, 55, 61	0
3	D	20/28 (71%)	-0.02	2 (10%) 9 4	38, 46, 58, 70	0
4	A	626/688 (90%)	-0.29	7 (1%) 82 58	26, 41, 50, 61	0
All	All	665/744 (89%)	-0.27	11 (1%) 73 45	26, 41, 52, 70	0

The worst 5 of 11 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	D	26	DC	3.5
1	B	3	DG	3.2
4	A	866	LYS	3.0
4	A	739	SER	2.8
4	A	781	ALA	2.7

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. 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, 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	LLDF	B-factors(Å ²)	Q<0.9
1	DOC	B	13	18/19	0.97	0.16	-	35,37,42,42	0

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, 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	LLDF	B-factors(\AA^2)	Q<0.9
5	AMP	C	100	23/23	0.98	0.14	-1.31	36,38,45,48	0

6.5 Other polymers [i](#)

There are no such residues in this entry.