



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 25, 2020 – 07:35 pm BST

PDB ID : 1V9P  
Title : Crystal Structure Of Nad<sup>+</sup>-Dependent DNA Ligase  
Authors : Lee, J.Y.; Chang, C.; Song, H.K.; Moon, J.; Yang, J.K.; Kim, H.K.; Kwon, S.K.; Suh, S.W.  
Deposited on : 2004-01-27  
Resolution : 2.90 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	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
buster-report	:	1.1.7 (2018)
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)	:	2.11

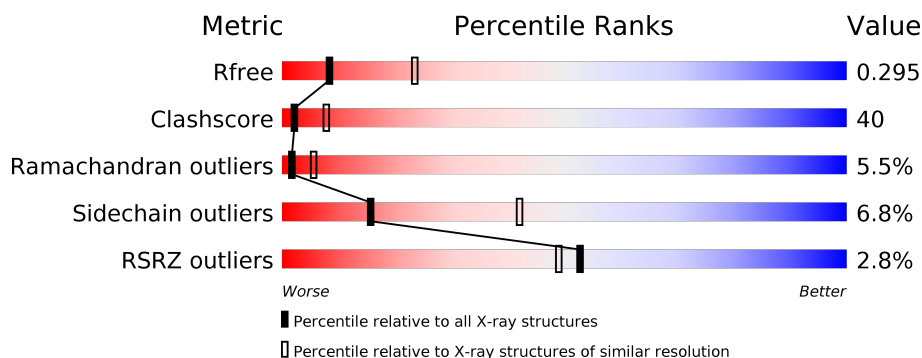
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.90 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	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)

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 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	584	<div> <div>3%</div> <div> <div></div> <div>44%</div> <div>48%</div> <div>8%</div> </div> </div>
1	B	584	<div> <div>3%</div> <div> <div></div> <div>43%</div> <div>50%</div> <div>7%</div> </div> </div>

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 9768 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.

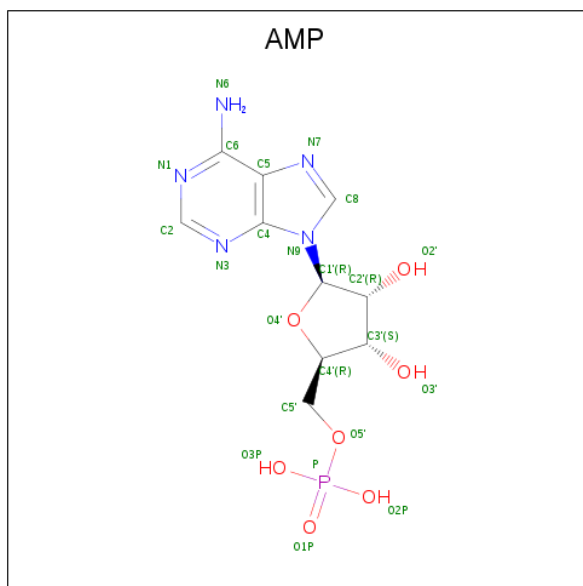
- Molecule 1 is a protein called DNA ligase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	584	Total	C	N	O	S	0	0	0
			4742	3000	860	871	11			
1	B	584	Total	C	N	O	S	0	0	0
			4742	3000	860	871	11			

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Zn	0	0
			1	1		
2	A	1	Total	Zn	0	0
			1	1		

- Molecule 3 is ADENOSINE MONOPHOSPHATE (three-letter code: AMP) (formula: C<sub>10</sub>H<sub>14</sub>N<sub>5</sub>O<sub>7</sub>P).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			22	10	5	6	1		
3	B	1	Total	C	N	O	P	0	0
			22	10	5	6	1		

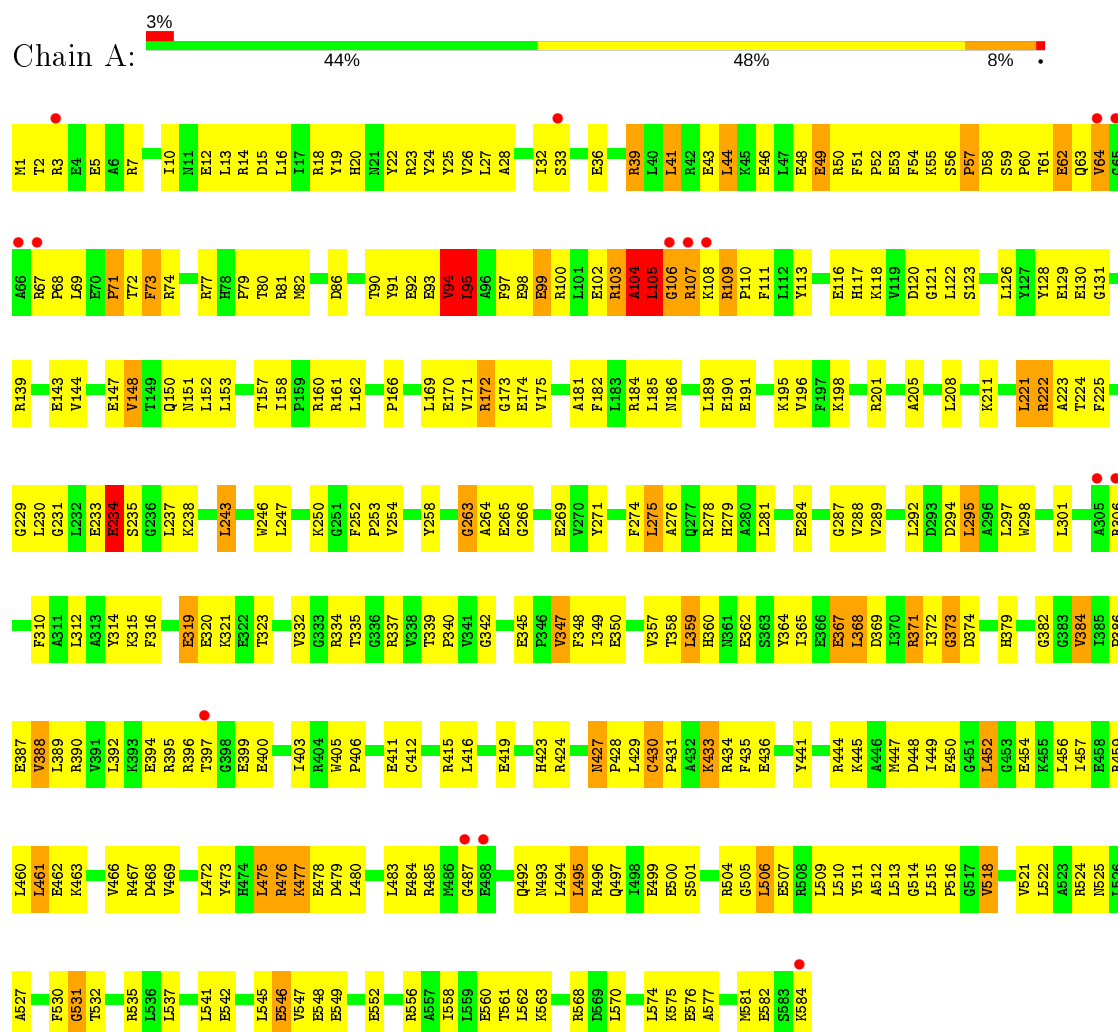
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	120	Total	O	0	0
			120	120		
4	B	118	Total	O	0	0
			118	118		

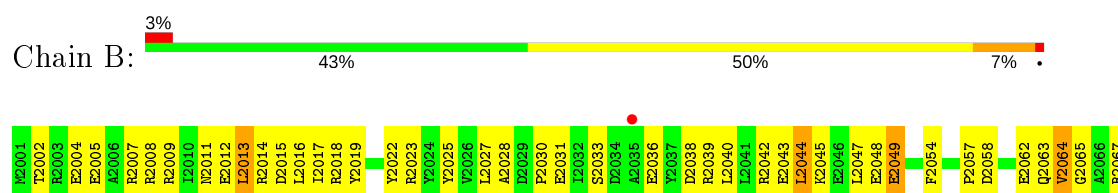
### 3 Residue-property plots

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



#### • Molecule 1: DNA ligase



L2510	R2439	L2377	G2362	L2228	T2138	P2068
L2511	H2440	V2378	T2303	G2229	R2139	P2071
A2512	Y2441	H2379	T2304	L2230	G2140	F2072
L2513	A2442	R2380	A2305	G2231	D2141	F2073
G2514	S2443	G2383	R2306	L2232	G2142	R2074
L2515	R2444	G2384	A2307	E2233	E2143	
	K2445	V2384	R2308	E2234	V2144	P2075
V2518	A2446	I2385	R2309	S2235	G2145	V2076
G2519	M2447	F2386		G2236	E2146	V2077
E2520		E2387	Y2314	L2237	E2147	H2078
V2521	G2451	V2388	K2315	K2238	E2148	P2079
L2522	L2452	L2389	F2316	S2239	V2149	T2080
		L2390	P2317	Q2240	Q2150	
A2527	L2456	V2391	E2320	Y2241	N2151	N2087
	I2457	L2392		E2242	L2152	A2088
F2530	L2460	K2393	T2323	L2243	L2153	F2089
G2531		E2394		L2244		T2090
T2532		R2395	D2327	L2245	T2157	V2091
V2533	K2463	R2396	V2328	K2246		E2092
	G2464	T2397	V2329	L2247		E2093
A2539	L2465	G2398	F2330		R2160	V2094
S2540	V2466	E2399	Q2331	K2250	R2161	L2095
L2541	R2467	E2400	V2332	G2251	L2162	A2096
E2542	D2468	R2401	G2333	F2252		F2097
E2543	V2469	P2402	R2334	P2253	P2166	E2098
	A2470	L2403	T2335	V2254		E2099
E2546		R2404	G2336	E2255	L2169	R2100
V2547	Y2473	W2405	G2337	H2256	E2170	L2101
E2548	H2474	P2406	V2338	G2257	V2171	E2102
E2549	L2475	E2407	V2339	Y2258	G2173	R2103
	R2476	T2408	P2340	E2259	E2174	A2104
	G2477	C2409	V2341	K2260		L2105
	E2478	P2410	G2342	A2261	I2179	G2106
	D2479	E2411		L2262	L2183	R2107
		C2412		G2263		K2108
	E2484	G2413	E2345	A2264	R2184	R2109
	R2485	H2414	P2346	E2265		P2110
	V2486	R2415		G2266	E2188	F2111
G2487	G2487	L2416	T2349	V2267	L2189	L2112
E2488	E2488	V2417	E2350	E2268	E2190	Y2113
K2489	K2489	K2418	G2351	E2269		T2114
					G2193	V2115
Q2492	Q2492	K2421	V2357	R2273	E2194	E2116
N2493	N2493	V2422	T2358		H2117	H2117
L2494	L2494	R2423	L2359	Q2277	K2195	K2118
L2495	L2495	R2424	E2360	R2278	V2196	
R2496	R2496	C2425	N2361	H2279		V2119
Q2497	Q2497	P2426	E2362	A2280	R2201	G2121
I2498	I2498	N2427	S2363	L2281	L2208	D2120
E2499	E2499	P2428	Y2364	P2282	R2209	G2121
E2500	E2500	L2429		F2283		S2123
S2501	S2501	C2430	L2368	E2284	R2214	V2124
K2502	K2502	P2431	D2369	A2285		N2125
H2503	H2503	A2432	T2370	D2286	L2221	L2126
R2504	R2504	K2433	R2371		R2222	Y2127
		E2434	T2372	L2292	A2223	Y2128
G2505	G2505	F2435	G2373		E2224	E2129
L2506	L2506	E2436	D2374	W2298	T2224	E2130
E2507	E2507	A2437	V2375		F2225	G2131
R2508	R2508	I2438	V2376	L2301	Y2226	V2132
L2509	L2509				A2227	

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	89.21Å 117.33Å 97.48Å 90.00° 115.09° 90.00°	Depositor
Resolution (Å)	19.93 – 2.90 24.75 – 2.88	Depositor EDS
% Data completeness (in resolution range)	89.4 (19.93-2.90) 94.4 (24.75-2.88)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.48 (at 2.89Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.232 , 0.293 0.238 , 0.295	Depositor DCC
$R_{free}$ test set	3869 reflections (9.94%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	53.4	Xtriage
Anisotropy	0.238	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 61.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.022 for h,-k,-h-l	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	9768	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	52.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.58% 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: AMP, ZN

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	0.47	2/4832 (0.0%)	0.75	7/6514 (0.1%)
1	B	0.42	2/4832 (0.0%)	0.70	5/6514 (0.1%)
All	All	0.45	4/9664 (0.0%)	0.72	12/13028 (0.1%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	359	LEU	N-CA	-6.58	1.33	1.46
1	B	2120	ASP	C-N	-6.48	1.21	1.33
1	A	487	GLY	N-CA	-5.20	1.38	1.46
1	B	2360	HIS	C-N	-5.09	1.22	1.34

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	359	LEU	CB-CA-C	11.56	132.16	110.20
1	B	2360	HIS	CA-C-N	9.32	137.70	117.20
1	A	359	LEU	N-CA-C	-8.49	88.08	111.00
1	A	94	VAL	CB-CA-C	-7.80	96.57	111.40
1	B	2360	HIS	N-CA-C	7.72	131.85	111.00

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	4742	0	4797	388	0
1	B	4742	0	4794	369	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	22	0	12	2	0
3	B	22	0	12	1	0
4	A	120	0	0	15	0
4	B	118	0	0	6	0
All	All	9768	0	9615	757	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 40.

The worst 5 of 757 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:2360:HIS:O	1:B:2361:ASN:HB3	1.37	1.11
1:A:10:ILE:HG23	1:A:14:ARG:NH1	1.68	1.09
1:A:166:PRO:CG	1:A:234:GLU:HB3	1.88	1.03
1:A:537:LEU:HD22	1:A:568:ARG:HE	1.24	1.02
1:B:2360:HIS:O	1:B:2361:ASN:CB	2.00	1.00

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	582/584 (100%)	478 (82%)	71 (12%)	33 (6%)	1	5
1	B	582/584 (100%)	463 (80%)	88 (15%)	31 (5%)	2	6
All	All	1164/1168 (100%)	941 (81%)	159 (14%)	64 (6%)	2	5

5 of 64 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	62	GLU
1	A	64	VAL
1	A	94	VAL
1	A	104	ALA
1	A	109	ARG

### 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	499/499 (100%)	469 (94%)	30 (6%)	19	49
1	B	499/499 (100%)	461 (92%)	38 (8%)	13	36
All	All	998/998 (100%)	930 (93%)	68 (7%)	16	42

5 of 68 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	2013	LEU
1	B	2153	LEU
1	B	2499	GLU
1	B	2044	LEU
1	B	2105	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	2011	ASN
1	B	2087	ASN
1	B	2277	GLN
1	A	427	ASN
1	A	525	ASN

### 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 ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 4 ligands modelled in this entry, 2 are monoatomic - leaving 2 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	$\# Z  > 2$
3	AMP	B	2700	-	18,24,25	1.26	3 (16%)	18,35,38	1.95	4 (22%)
3	AMP	A	700	-	18,24,25	1.20	3 (16%)	18,35,38	1.89	4 (22%)

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	AMP	B	2700	-	-	1/3/25/26	0/3/3/3
3	AMP	A	700	-	-	1/3/25/26	0/3/3/3

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	2700	AMP	O4'-C1'	2.66	1.44	1.41

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	700	AMP	C5-N7	-2.49	1.30	1.39
3	B	2700	AMP	C5-N7	-2.46	1.30	1.39
3	A	700	AMP	C2-N3	2.43	1.36	1.32
3	B	2700	AMP	C2-N3	2.19	1.35	1.32

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	2700	AMP	N3-C2-N1	-6.46	118.58	128.68
3	A	700	AMP	N3-C2-N1	-6.20	118.99	128.68
3	A	700	AMP	C3'-C2'-C1'	2.90	105.34	100.98
3	B	2700	AMP	C3'-C2'-C1'	2.78	105.17	100.98
3	B	2700	AMP	C2'-C3'-C4'	2.31	107.14	102.64

There are no chirality outliers.

All (2) torsion outliers are listed below:

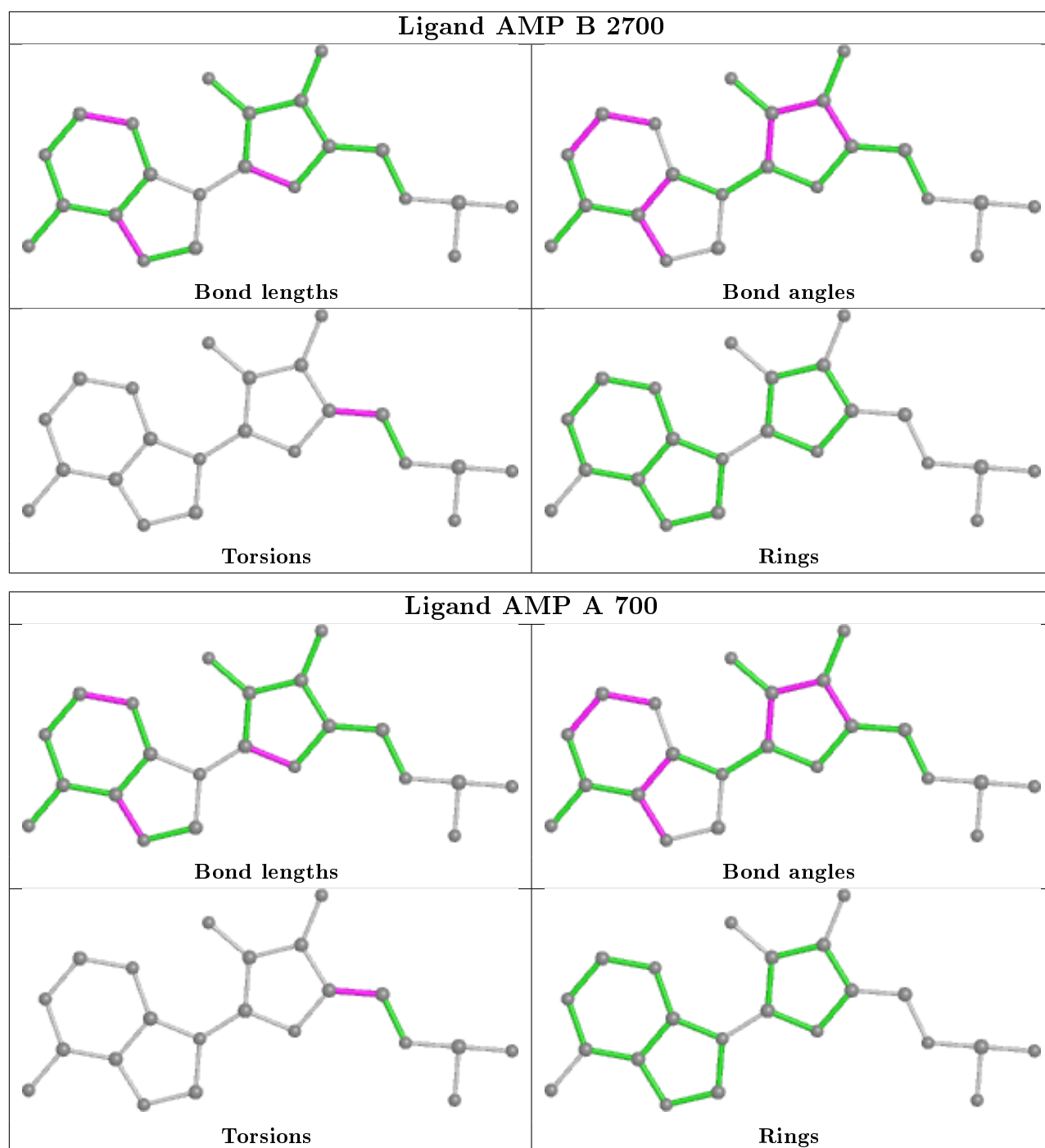
Mol	Chain	Res	Type	Atoms
3	A	700	AMP	O4'-C4'-C5'-O5'
3	B	2700	AMP	O4'-C4'-C5'-O5'

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	2700	AMP	1	0
3	A	700	AMP	2	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.



## 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	584/584 (100%)	-0.12	15 (2%) 56 52	12, 46, 92, 142	0
1	B	584/584 (100%)	-0.05	18 (3%) 49 44	12, 50, 98, 151	0
All	All	1168/1168 (100%)	-0.08	33 (2%) 53 49	12, 48, 95, 151	0

The worst 5 of 33 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	106	GLY	8.4
1	A	107	ARG	7.3
1	B	2107	ARG	5.6
1	B	2351	GLY	5.5
1	B	2305	ALA	4.8

### 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](#)

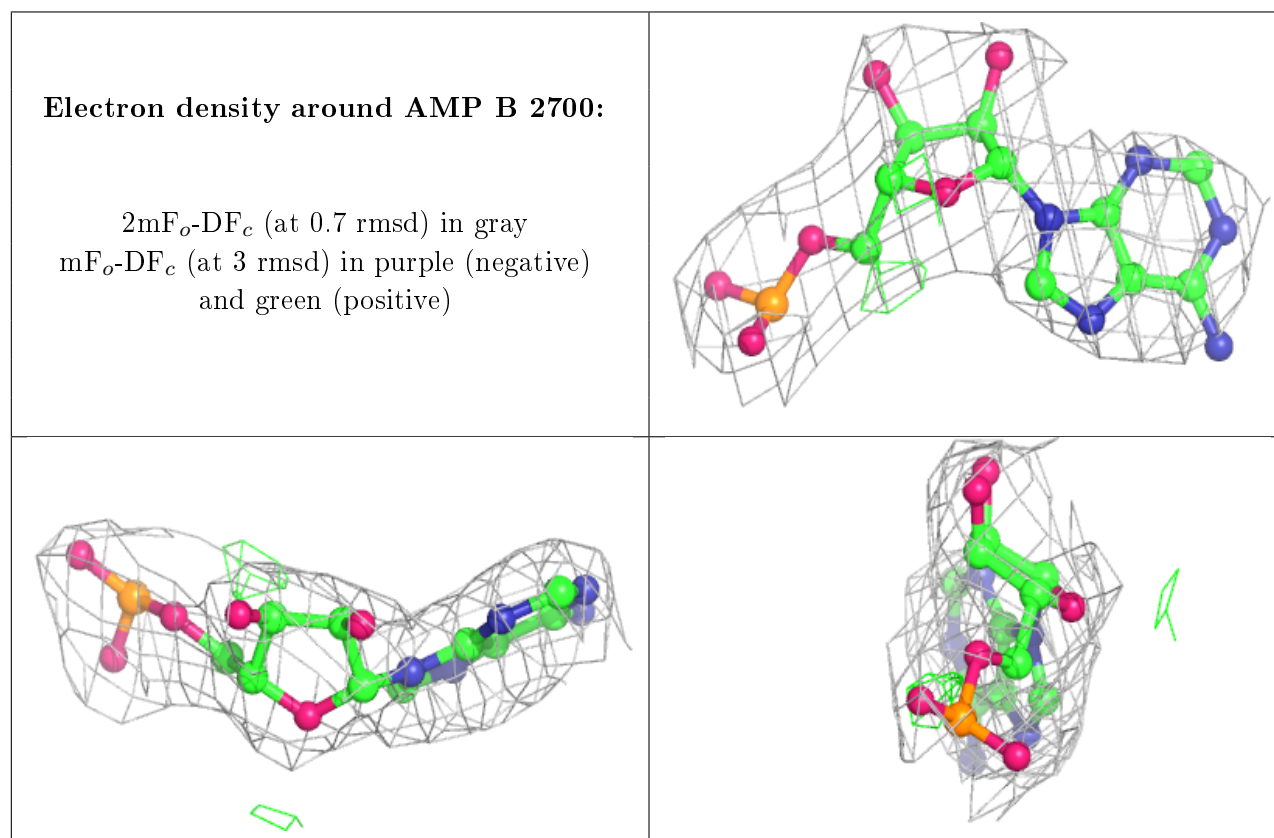
There are no carbohydrates in this entry.

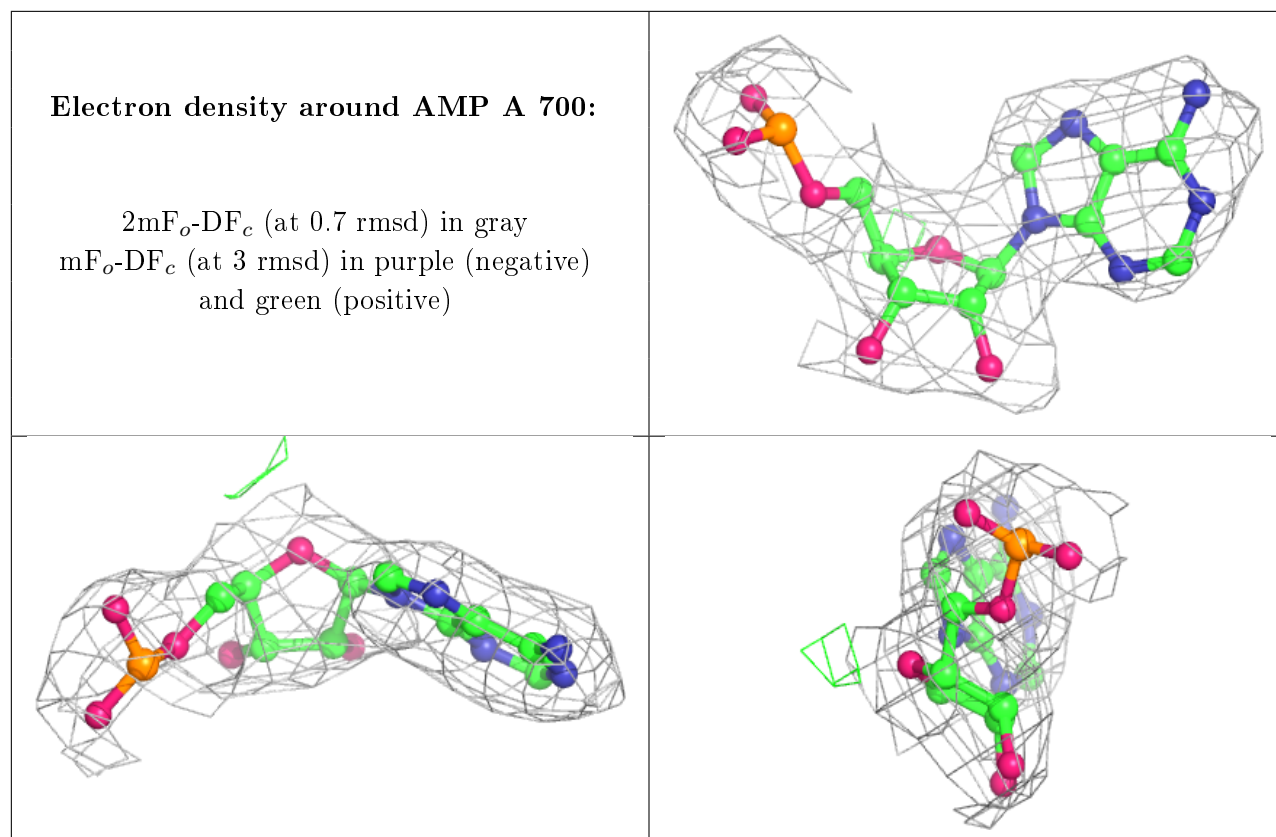
### 6.4 Ligands [i](#)

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. 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( $\text{\AA}^2$ )	Q<0.9
3	AMP	B	2700	22/23	0.91	0.21	81,90,93,93	0
3	AMP	A	700	22/23	0.94	0.16	49,53,56,57	0
2	ZN	B	2701	1/1	0.99	0.10	29,29,29,29	0
2	ZN	A	701	1/1	1.00	0.10	19,19,19,19	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.





## 6.5 Other polymers [i](#)

There are no such residues in this entry.