



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 27, 2020 – 10:45 pm BST

PDB ID : 2W00  
Title : Crystal structure of the HsdR subunit of the EcoR124I restriction enzyme in complex with ATP  
Authors : Lapkouski, M.; Panjikar, S.; Kuta Smatanova, I.; Ettrich, R.; Csefalvay, E.  
Deposited on : 2008-08-08  
Resolution : 2.60 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

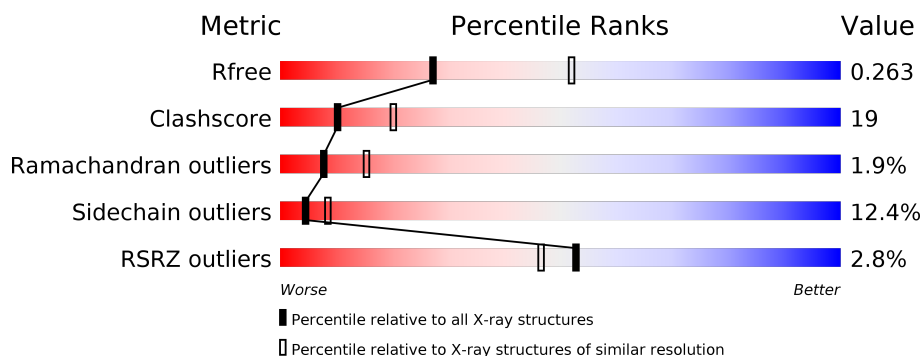
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.60 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	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

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	1038	<div> <div>2%</div> <div> <div></div> <div>54%</div> <div>20%</div> <div>6%</div> <div>19%</div> </div> </div>
1	B	1038	<div> <div>2%</div> <div> <div></div> <div>55%</div> <div>20%</div> <div>6%</div> <div>18%</div> </div> </div>

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 14386 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 HSDR.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	842	Total	C	N	O	S	Se	0	0	0
			6825	4341	1154	1314	3	13			
1	B	852	Total	C	N	O	S	Se	0	0	0
			6915	4393	1172	1334	3	13			

- Molecule 2 is ADENOSINE-5'-TRIPHOSPHATE (three-letter code: ATP) (formula:  $C_{10}H_{16}N_5O_{13}P_3$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
2	B	1	Total	C	N	O	P	0	0
			31	10	5	13	3		

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	286	Total 286	O 286	0	0
4	B	296	Total 296	O 296	0	0



Chain B:

Sequence logo for Chain B, showing amino acid conservation across 100 positions. The y-axis represents information content in bits (0.00 to 2.00). The x-axis shows positions 1 to 100. The color scale indicates conservation levels: 29% (red), 55% (green), 20% (yellow), 6% (orange), and 18% (grey).

Key features of the sequence logo:

- High Conservation (Red/Green):** Positions 1-10, 15-20, 25-30, 35-40, 45-50, 55-60, 65-70, 75-80, 85-90, 95-100.
- Low Conservation (Yellow/Orange):** Positions 11-14, 21-24, 31-34, 41-44, 51-54, 61-64, 71-74, 81-84, 91-94.
- Medium Conservation (Grey):** Positions 16-19, 26-29, 36-39, 46-49, 56-59, 66-69, 76-79, 86-89, 96-99.

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	123.91Å 129.92Å 161.01Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.97 – 2.60 19.97 – 2.60	Depositor EDS
% Data completeness (in resolution range)	100.0 (19.97-2.60) 100.0 (19.97-2.60)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.82 (at 2.59Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.221 , 0.265 0.220 , 0.263	Depositor DCC
$R_{free}$ test set	977 reflections (1.22%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	43.4	Xtriage
Anisotropy	0.071	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.27 , 53.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.000 for k,h,-l	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	14386	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 44.47 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.5201e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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: MG, ATP

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	0.62	3/6942 (0.0%)	0.72	2/9341 (0.0%)
1	B	0.65	8/7034 (0.1%)	0.70	5/9464 (0.1%)
All	All	0.64	11/13976 (0.1%)	0.71	7/18805 (0.0%)

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
1	A	1	40
1	B	2	37
All	All	3	77

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	827	GLU	CD-OE2	13.17	1.40	1.25
1	A	827	GLU	CD-OE2	8.36	1.34	1.25
1	B	892	TYR	CG-CD2	6.75	1.48	1.39
1	B	825	ASP	CG-OD1	6.38	1.40	1.25
1	B	892	TYR	CE1-CZ	6.15	1.46	1.38

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	811	ASP	C-N-CD	-18.35	80.23	120.60
1	A	533	PRO	N-CA-CB	6.11	110.63	103.30
1	B	511	MSE	N-CA-C	6.09	127.44	111.00
1	B	511	MSE	CB-CG-SE	-5.99	94.74	112.70

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	422	LEU	CA-CB-CG	5.89	128.85	115.30

All (3) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	207	THR	CB
1	B	207	THR	CB
1	B	511	MSE	CA

5 of 77 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	140	PHE	Peptide
1	A	170	GLY	Peptide
1	A	190	SER	Peptide
1	A	26	THR	Peptide
1	A	28	ASP	Peptide

## 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	6825	0	6665	257	0
1	B	6915	0	6741	258	0
2	A	31	0	12	1	0
2	B	31	0	12	2	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	286	0	0	11	0
4	B	296	0	0	14	0
All	All	14386	0	13430	512	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 19.

The worst 5 of 512 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:A:810:SER:CA	1:A:812:PRO:CD	1.90	1.46
1:B:769:ILE:HG22	1:B:770:GLU:HA	1.20	1.17
1:A:533:PRO:HA	1:A:535:SER:H	1.12	1.14
1:A:769:ILE:HD13	1:A:769:ILE:H	1.06	1.12
1:A:769:ILE:CD1	1:A:769:ILE:H	1.60	1.12

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	832/1038 (80%)	754 (91%)	60 (7%)	18 (2%)	6	12
1	B	842/1038 (81%)	768 (91%)	61 (7%)	13 (2%)	10	21
All	All	1674/2076 (81%)	1522 (91%)	121 (7%)	31 (2%)	8	15

5 of 31 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	28	ASP
1	A	411	HIS
1	A	532	PHE
1	A	568	SER
1	A	763	PHE

### 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	
1	A	740/912 (81%)	648 (88%)	92 (12%)	4	8
1	B	750/912 (82%)	657 (88%)	93 (12%)	4	8
All	All	1490/1824 (82%)	1305 (88%)	185 (12%)	4	8

5 of 185 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	834	THR
1	B	89	LEU
1	B	769	ILE
1	A	842	LYS
1	B	26	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 63 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	647	ASN
1	B	47	GLN
1	B	561	GLN
1	A	676	ASN
1	A	805	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 ⓘ

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	$\# Z  > 2$
2	ATP	B	1893	3	26,33,33	0.87	0	31,52,52	1.43	5 (16%)
2	ATP	A	1886	3	26,33,33	0.82	1 (3%)	31,52,52	1.51	5 (16%)

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
2	ATP	B	1893	3	-	0/18/38/38	0/3/3/3
2	ATP	A	1886	3	-	0/18/38/38	0/3/3/3

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1886	ATP	C5-C4	2.12	1.46	1.40

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1886	ATP	N3-C2-N1	-3.72	122.86	128.68
2	B	1893	ATP	PB-O3B-PG	-3.63	120.37	132.83
2	A	1886	ATP	PB-O3B-PG	-2.84	123.09	132.83
2	B	1893	ATP	C4-C5-N7	-2.78	106.51	109.40
2	A	1886	ATP	C2-N1-C6	2.60	123.20	118.75

There are no chirality outliers.

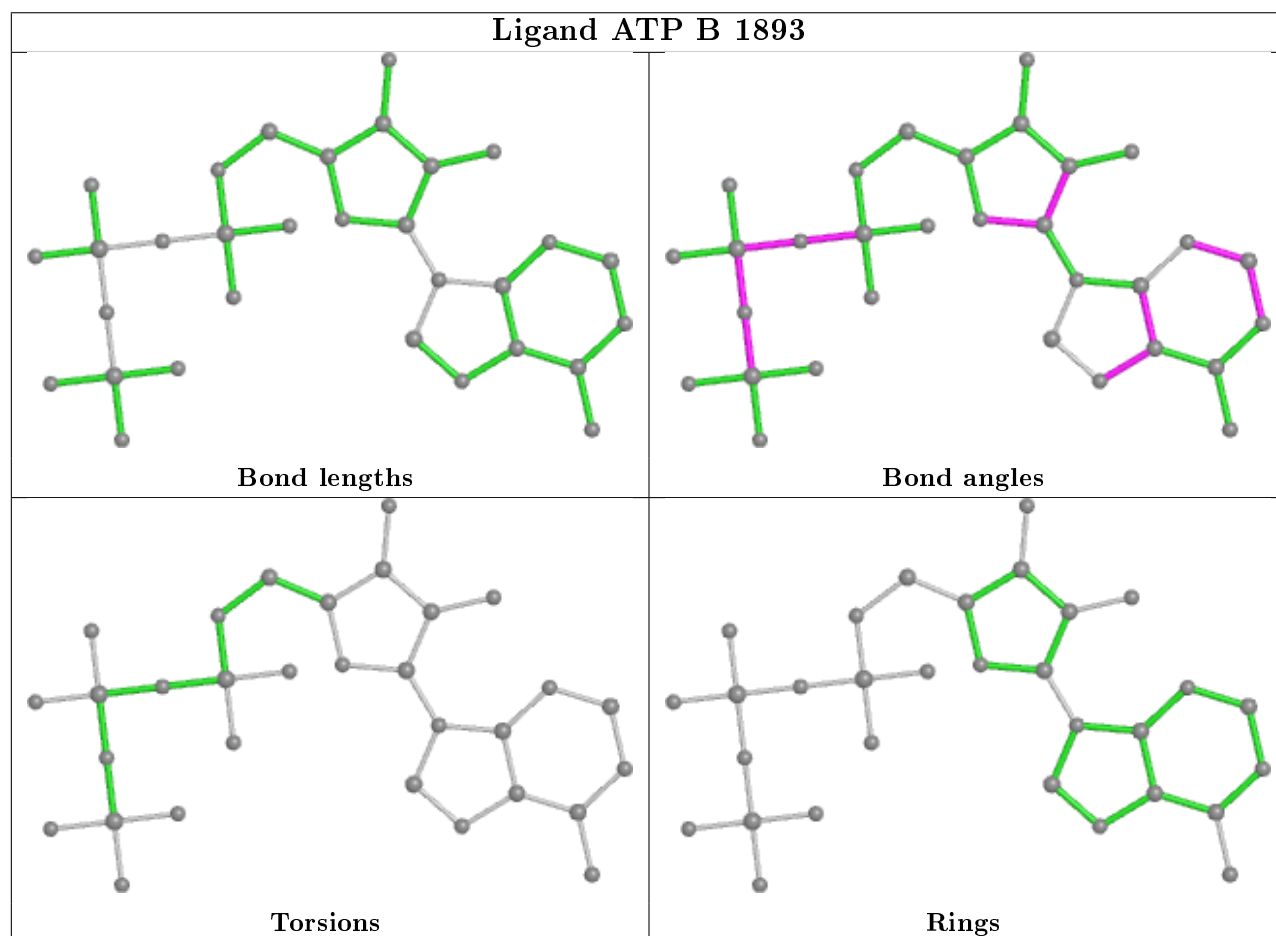
There are no torsion outliers.

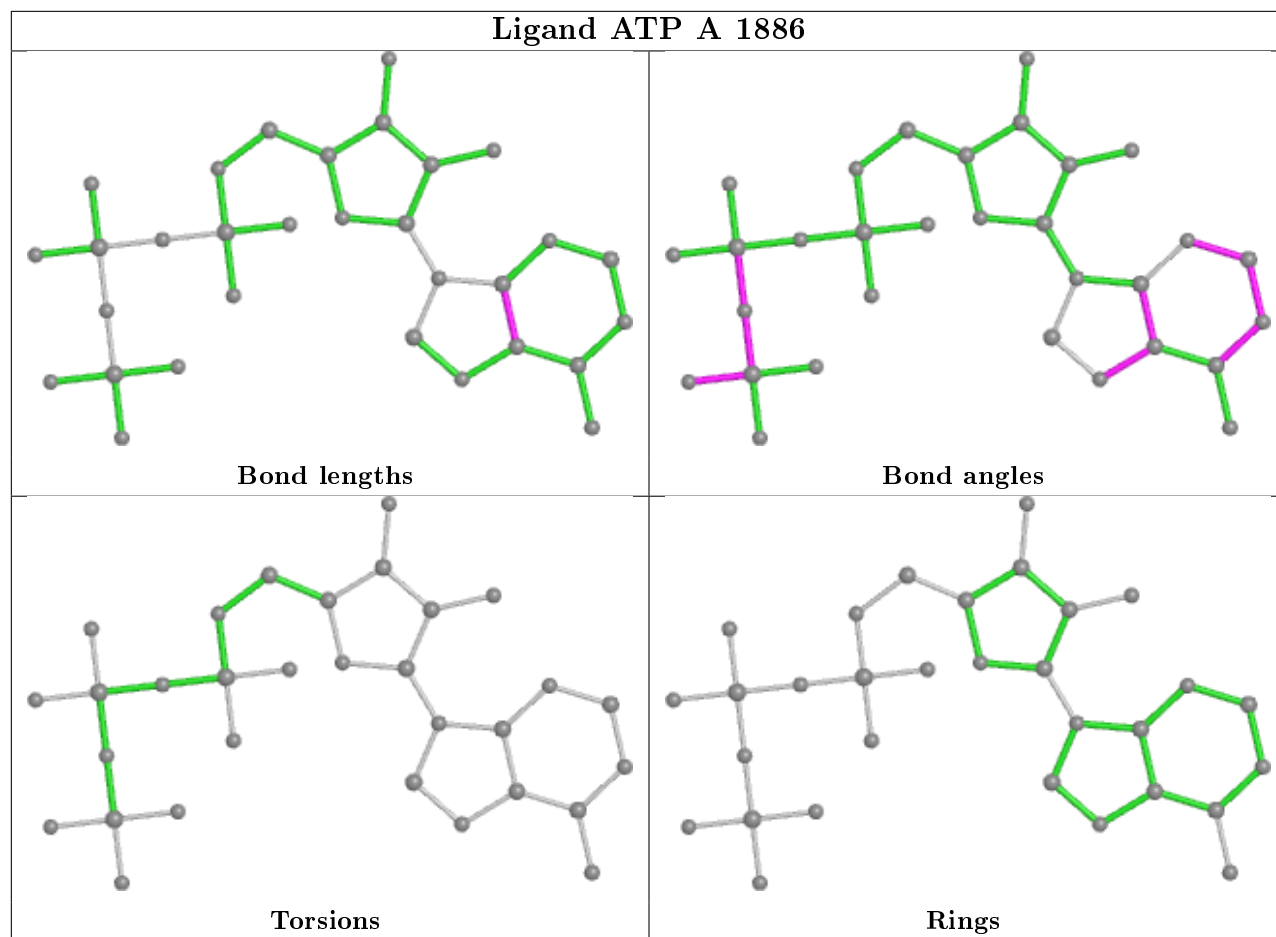
There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	1893	ATP	2	0
2	A	1886	ATP	1	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	829/1038 (79%)	-0.12	22 (2%) 54 48	24, 40, 53, 66	0
1	B	839/1038 (80%)	-0.08	24 (2%) 51 45	25, 40, 53, 66	0
All	All	1668/2076 (80%)	-0.10	46 (2%) 53 46	24, 40, 53, 66	0

The worst 5 of 46 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	445	LEU	4.7
1	B	375	ASP	4.5
1	B	26	THR	4.4
1	B	446	GLY	4.1
1	B	141	GLU	4.1

### 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](#)

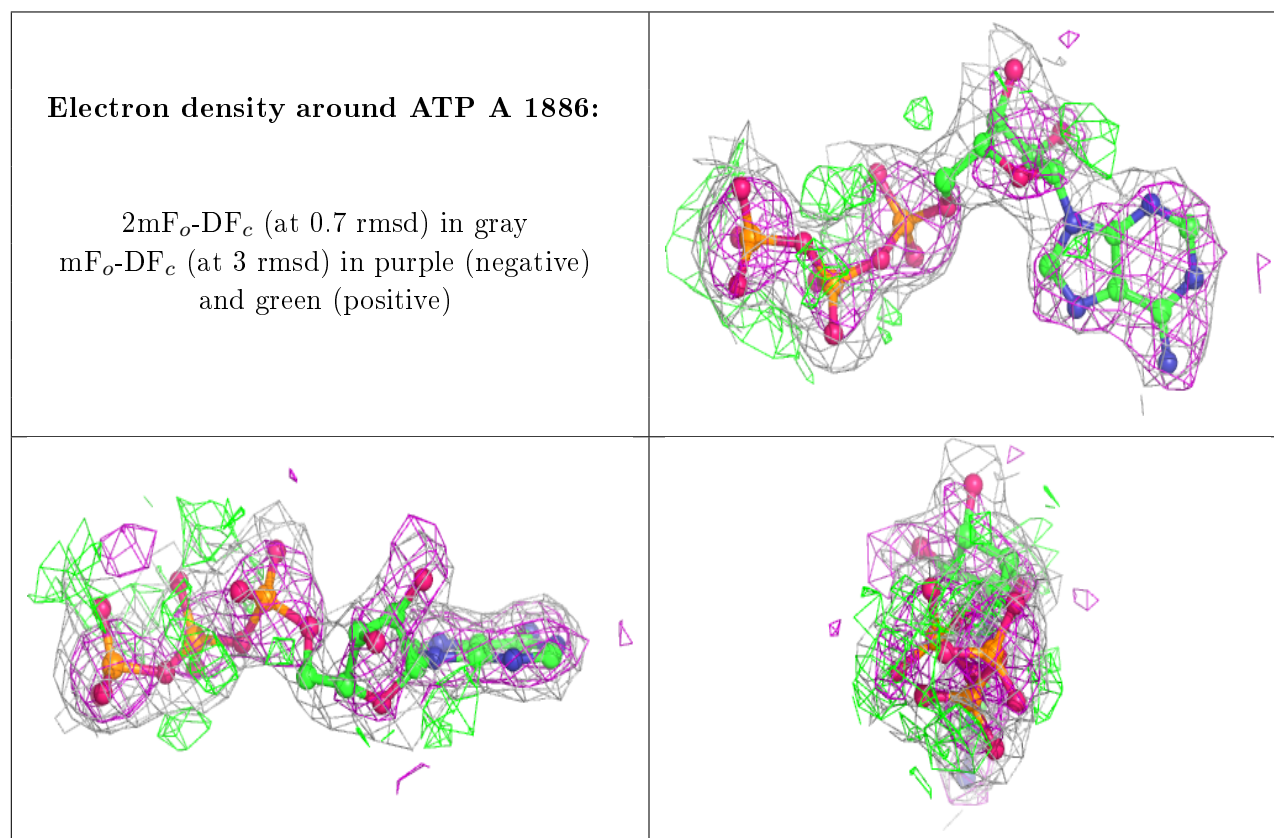
There are no carbohydrates in this entry.

### 6.4 Ligands [i](#)

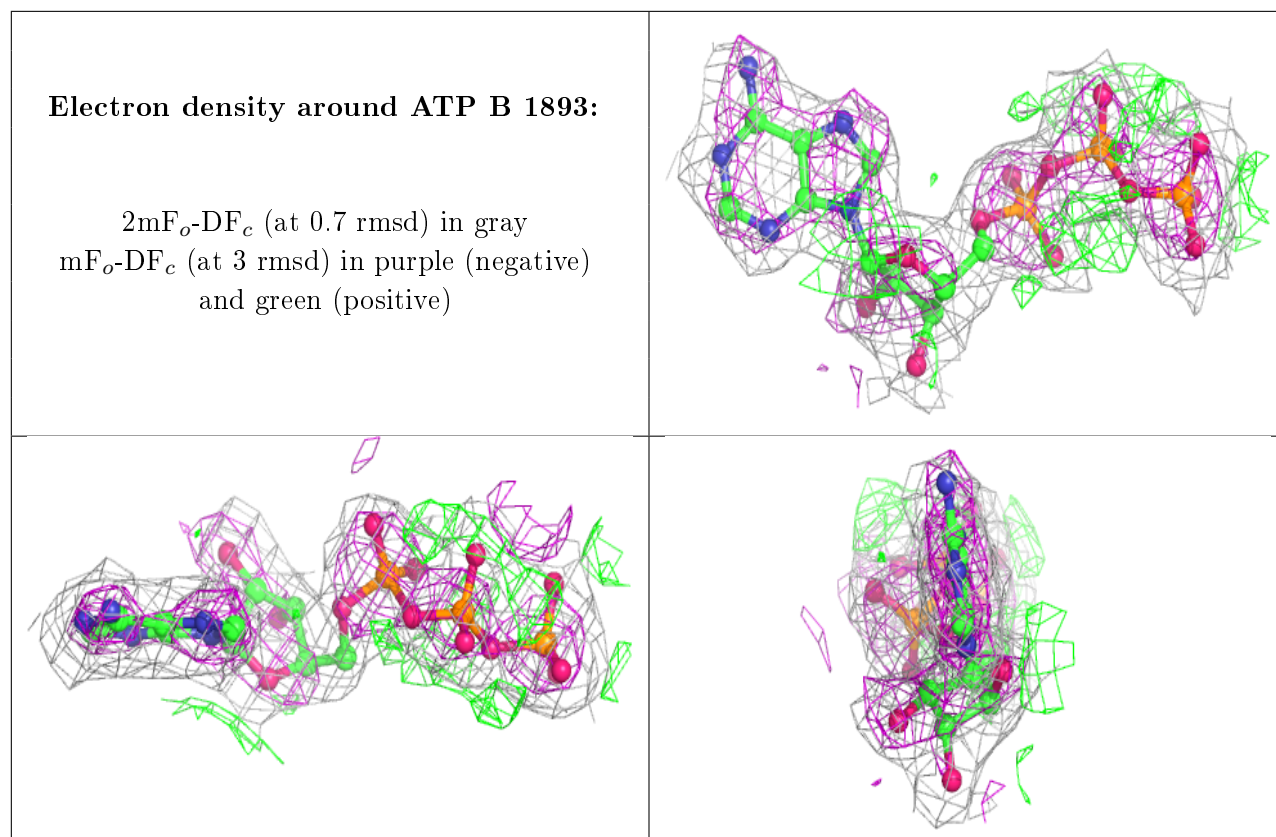
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 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	MG	A	1887	1/1	0.97	0.04	11,11,11,11	0
3	MG	B	1894	1/1	0.98	0.09	13,13,13,13	0
2	ATP	A	1886	31/31	0.98	0.12	17,26,31,33	0
2	ATP	B	1893	31/31	0.99	0.10	20,26,30,33	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.