



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 10, 2020 – 05:55 AM BST

PDB ID : 1RWQ  
Title : Human Dipeptidyl peptidase IV in complex with 5-aminomethyl-6-(2,4-dichloro-phenyl)-2-(3,5-dimethoxy-phenyl)-pyrimidin-4-ylamine  
Authors : Hennig, M.; Thoma, R.; Stihle, M.  
Deposited on : 2003-12-17  
Resolution : 2.20 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)  
Xtrriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.13.1

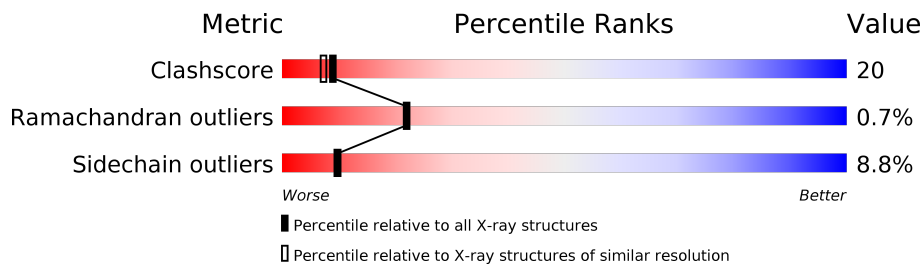
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.20 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))
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)

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  $\leq 5\%$ .

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	728	
1	B	728	

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 12190 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 Dipeptidyl peptidase IV.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	728	Total 5963	C 3827	N 982	O 1128	S 26	0	0	0
1	B	728	Total 5963	C 3827	N 982	O 1128	S 26	0	0	0

- Molecule 2 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



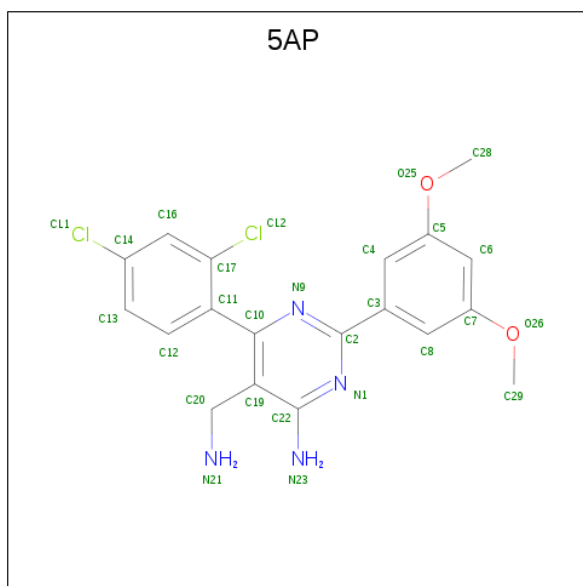
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
2	A	1	Total 14	C 8	N 1	O 5	0	0
2	A	1	Total 14	C 8	N 1	O 5	0	0
2	A	1	Total 14	C 8	N 1	O 5	0	0
2	A	1	Total 14	C 8	N 1	O 5	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	B	1	Total	C	N	O	0	0
			14	8	1	5		
2	B	1	Total	C	N	O	0	0
			14	8	1	5		
2	B	1	Total	C	N	O	0	0
			14	8	1	5		
2	B	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 3 is 5-(AMINOMETHYL)-6-(2,4-DICHLOROPHENYL)-2-(3,5-DIMETHOXYPHENYL)PYRIMIDIN-4-AMINE (three-letter code: 5AP) (formula: C<sub>19</sub>H<sub>18</sub>Cl<sub>2</sub>N<sub>4</sub>O<sub>2</sub>).

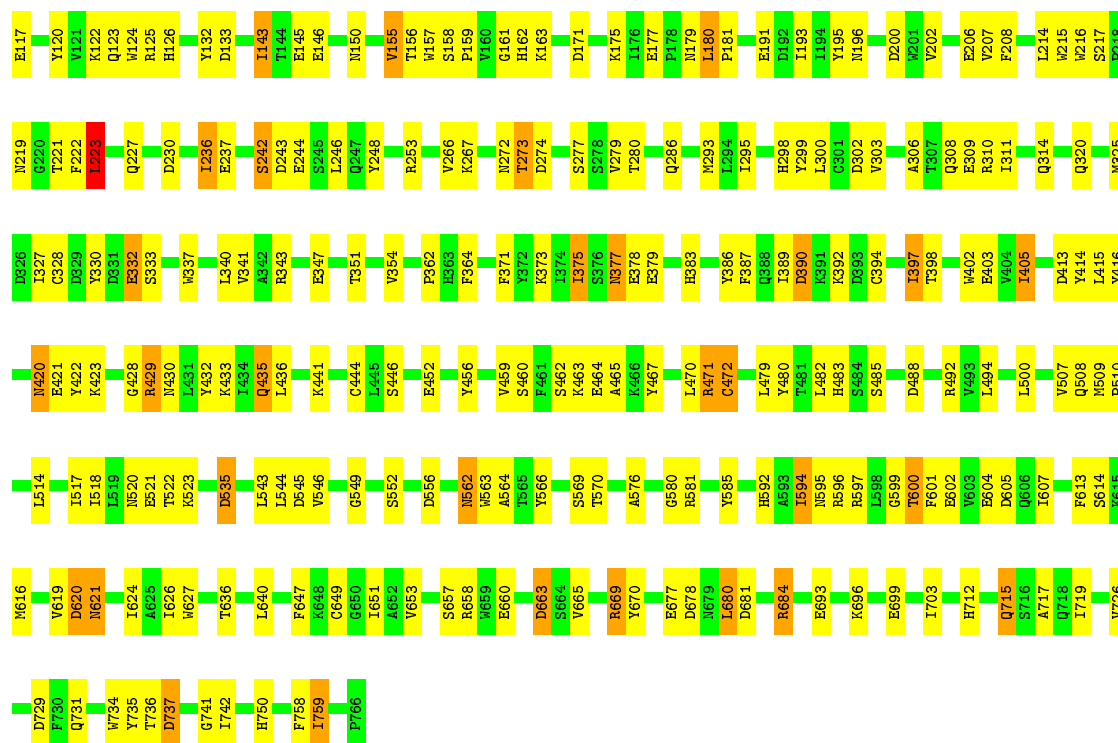


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
3	A	1	Total	C	Cl	N	O	0	0
			27	19	2	4	2		
3	B	1	Total	C	Cl	N	O	0	0
			27	19	2	4	2		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	47	Total	O	0	0
			47	47		
4	B	51	Total	O	0	0
			51	51		





## 4 Data and refinement statistics

Xtrriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	65.68Å 69.14Å 422.56Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	15.00 – 2.20	Depositor
% Data completeness (in resolution range)	84.3 (15.00-2.20)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.10	Depositor
Refinement program	REFMAC 5.0	Depositor
R, $R_{free}$	0.239 , 0.290	Depositor
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	12190	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	43.0	wwPDB-VP

## 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: 5AP, NAG

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	0.99	2/6135 (0.0%)	1.14	30/8344 (0.4%)
1	B	1.03	4/6135 (0.1%)	1.16	32/8344 (0.4%)
All	All	1.01	6/12270 (0.0%)	1.15	62/16688 (0.4%)

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	124	TRP	CG-CD1	5.87	1.45	1.36
1	B	429	ARG	NE-CZ	5.84	1.40	1.33
1	A	745	SER	CB-OG	-5.64	1.34	1.42
1	A	248	TYR	CD1-CE1	5.59	1.47	1.39
1	B	243	ASP	CB-CG	5.44	1.63	1.51

The worst 5 of 62 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	413	ASP	CB-CG-OD2	9.51	126.86	118.30
1	A	301	CYS	CA-CB-SG	-9.32	97.23	114.00
1	A	501	ASP	CB-CG-OD2	8.29	125.76	118.30
1	A	729	ASP	CB-CG-OD2	8.19	125.67	118.30
1	B	274	ASP	CB-CG-OD2	8.10	125.59	118.30

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	5963	0	5681	241	0
1	B	5963	0	5683	232	0
2	A	56	0	52	2	0
2	B	56	0	52	8	0
3	A	27	0	18	5	0
3	B	27	0	18	0	0
4	A	47	0	0	4	0
4	B	51	0	0	0	0
All	All	12190	0	11504	473	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 20.

The worst 5 of 473 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:150:ASN:HD21	2:B:793:NAG:C1	1.45	1.28
1:B:600:THR:CG2	1:B:601:PHE:H	1.59	1.15
1:B:267:LYS:HD2	1:B:286:GLN:HE22	1.06	1.12
1:B:92:ASN:HD22	2:B:797:NAG:C1	1.65	1.08
1:B:600:THR:HG22	1:B:601:PHE:H	1.23	1.02

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	726/728 (100%)	665 (92%)	54 (7%)	7 (1%)	15 14
1	B	726/728 (100%)	665 (92%)	58 (8%)	3 (0%)	34 37

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	1452/1456 (100%)	1330 (92%)	112 (8%)	10 (1%)	22	22

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	40	ARG
1	A	488	ASP
1	B	111	GLY
1	B	242	SER
1	A	615	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	
1	A	653/653 (100%)	596 (91%)	57 (9%)	10	10
1	B	653/653 (100%)	595 (91%)	58 (9%)	9	9
All	All	1306/1306 (100%)	1191 (91%)	115 (9%)	10	10

5 of 115 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	677	GLU
1	B	87	SER
1	B	621	ASN
1	A	684	ARG
1	B	42	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 35 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	750	HIS
1	B	263	ASN
1	B	679	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	66	HIS
1	B	126	HIS

### 5.3.3 RNA [i](#)

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 monosaccharides in this entry.

### 5.6 Ligand geometry [i](#)

10 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
2	NAG	B	796	1	14,14,15	0.87	1 (7%)	17,19,21	1.39	3 (17%)
2	NAG	B	793	1	14,14,15	0.70	1 (7%)	17,19,21	1.26	2 (11%)
3	5AP	B	1900	-	28,29,29	1.50	5 (17%)	36,41,41	1.94	10 (27%)
2	NAG	A	793	1	14,14,15	0.72	1 (7%)	17,19,21	2.26	6 (35%)
2	NAG	A	794	1	14,14,15	1.01	1 (7%)	17,19,21	2.14	7 (41%)
2	NAG	B	794	1	14,14,15	0.73	1 (7%)	17,19,21	1.51	2 (11%)
2	NAG	A	796	1	14,14,15	0.74	0	17,19,21	1.93	5 (29%)
2	NAG	B	797	1	14,14,15	0.63	0	17,19,21	1.89	4 (23%)
3	5AP	A	900	-	28,29,29	2.01	5 (17%)	36,41,41	2.29	12 (33%)
2	NAG	A	795	1	14,14,15	0.72	0	17,19,21	2.24	4 (23%)

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	NAG	B	796	1	-	1/6/23/26	0/1/1/1
2	NAG	B	793	1	-	5/6/23/26	0/1/1/1
3	5AP	B	1900	-	-	5/12/14/14	0/3/3/3
2	NAG	A	793	1	-	5/6/23/26	0/1/1/1
2	NAG	A	794	1	-	0/6/23/26	0/1/1/1
2	NAG	B	794	1	-	3/6/23/26	0/1/1/1
2	NAG	A	796	1	-	0/6/23/26	0/1/1/1
2	NAG	B	797	1	-	2/6/23/26	0/1/1/1
3	5AP	A	900	-	-	3/12/14/14	0/3/3/3
2	NAG	A	795	1	-	4/6/23/26	0/1/1/1

The worst 5 of 15 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	900	5AP	C11-C10	-8.15	1.41	1.49
3	B	1900	5AP	C10-N9	-3.90	1.28	1.34
3	A	900	5AP	C10-C19	3.28	1.47	1.41
3	B	1900	5AP	C12-C11	-3.04	1.35	1.40
3	B	1900	5AP	C10-C19	2.75	1.46	1.41

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	793	NAG	C4-C3-C2	-5.88	102.40	111.02
3	A	900	5AP	C16-C17-C11	5.67	125.94	121.91
3	A	900	5AP	C22-N1-C2	5.17	124.22	117.50
2	A	795	NAG	C4-C3-C2	-5.01	103.67	111.02
3	A	900	5AP	C17-C16-C14	-4.79	113.36	118.71

There are no chirality outliers.

5 of 28 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	793	NAG	C8-C7-N2-C2
2	B	793	NAG	O7-C7-N2-C2
2	A	793	NAG	C3-C2-N2-C7

*Continued on next page...*

*Continued from previous page...*

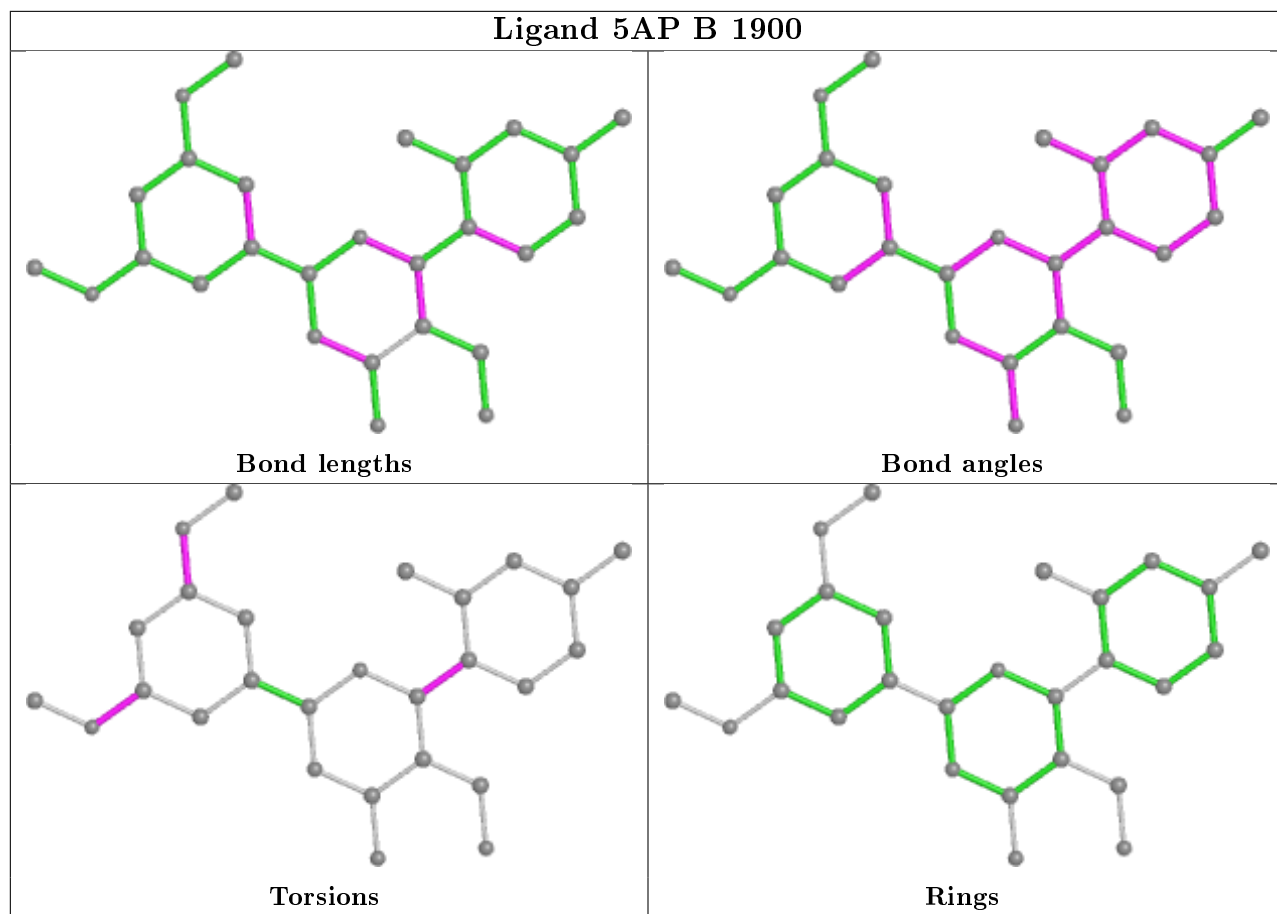
Mol	Chain	Res	Type	Atoms
2	A	793	NAG	C8-C7-N2-C2
2	A	793	NAG	O7-C7-N2-C2

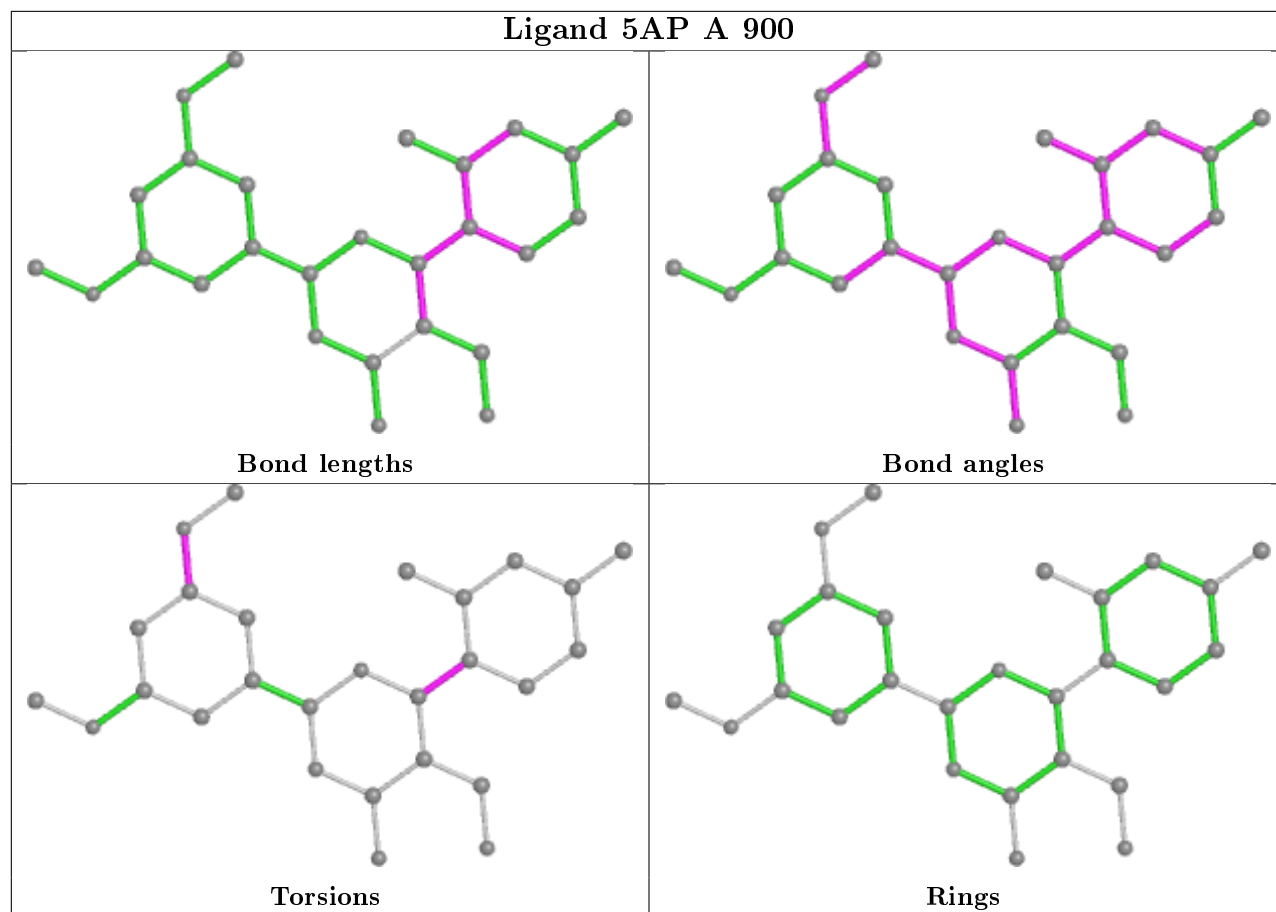
There are no ring outliers.

6 monomers are involved in 15 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	793	NAG	5	0
2	A	793	NAG	1	0
2	B	794	NAG	1	0
2	B	797	NAG	2	0
3	A	900	5AP	5	0
2	A	795	NAG	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](#)

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates [i](#)

EDS was not executed - this section is therefore empty.

### 6.4 Ligands [i](#)

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers [i](#)

EDS was not executed - this section is therefore empty.