

wwPDB X-ray Structure Validation Summary Report (i)

Nov 7, 2023 – 05:17 PM JST

ant

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 (i) 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 (1)) were used in the production of this report:

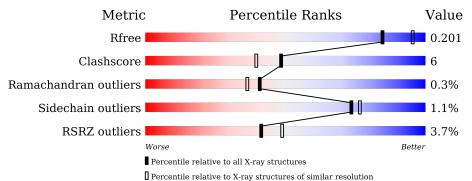
MolProbity Mogul Xtriage (Phenix) EDS	:	4.02b-467 1.8.5 (274361), CSD as541be (2020) 1.13 2.36
buster-report Percentile statistics Refmac	: : :	1.1.7 (2018) 20191225.v01 (using entries in the PDB archive December 25th 2019) 5.8.0158 7.0.044 (Gargrove)
Ideal geometry (DNA, RNA) Validation Pipeline (wwPDB-VP)		Parkinson et al. (1996) 2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure: $X\text{-}RAY \, DIFFRACTION$

The reported resolution of this entry is 2.14 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	$\begin{array}{l} \textbf{Whole archive} \\ (\#\textbf{Entries}) \end{array}$	${f Similar\ resolution}\ (\#{ m Entries,\ resolution\ range}({ m \AA}))$
R _{free}	130704	2523 (2.16-2.12)
Clashscore	141614	2653(2.16-2.12)
Ramachandran outliers	138981	2618 (2.16-2.12)
Sidechain outliers	138945	2617 (2.16-2.12)
RSRZ outliers	127900	2485 (2.16-2.12)

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 <=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	А	524	89%	10%	
1	В	524	% 8 9%	10%	
1	С	524	87%	11%	•
1	D	524	<u>6%</u> 83%	15%	••



6M2T

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 17180 atoms, of which 28 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.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	Λ	518	Total	С	Ν	Ο	\mathbf{S}	0	3	0
	А	510	3914	2508	676	721	9	0		0
1	В	518	Total	С	Ν	0	S	0	3	0
	D	510	3914	2508	676	721	9	0	5	0
1	С	518	Total	С	Ν	0	S	0	2	0
	U	510	3909	2503	676	721	9	0	Ζ	0
1	л	518	Total	С	Ν	0	S	0	2	0
	D	510	3914	2508	676	721	9	0	3	U

• Molecule 1 is a protein called Benzoate-coenzyme A ligase.

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
А	1	MET	-	initiating methionine	UNP Q93TK0
А	83	ALA	THR	engineered mutation	UNP Q93TK0
А	333	ALA	HIS	engineered mutation	UNP Q93TK0
А	334	ALA	ILE	engineered mutation	UNP Q93TK0
А	341	ASP	GLY	engineered mutation	UNP Q93TK0
A	524	GLY	-	expression tag	UNP Q93TK0
В	1	MET	-	initiating methionine	UNP Q93TK0
В	83	ALA	THR	engineered mutation	UNP Q93TK0
В	333	ALA	HIS	engineered mutation	UNP Q93TK0
В	334	ALA	ILE	engineered mutation	UNP Q93TK0
В	341	ASP	GLY	engineered mutation	UNP Q93TK0
В	524	GLY	-	expression tag	UNP Q93TK0
С	1	MET	-	initiating methionine	UNP Q93TK0
С	83	ALA	THR	engineered mutation	UNP Q93TK0
С	333	ALA	HIS	engineered mutation	UNP Q93TK0
С	334	ALA	ILE	engineered mutation	UNP Q93TK0
С	341	ASP	GLY	engineered mutation	UNP Q93TK0
С	524	GLY	-	expression tag	UNP Q93TK0
D	1	MET	-	initiating methionine	UNP Q93TK0
D	83	ALA	THR	engineered mutation	UNP Q93TK0
D	333	ALA	HIS	engineered mutation	UNP Q93TK0

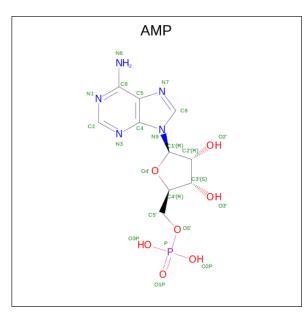
Continued on next page...



	<i>J</i> 1	ere ae page			1
Chain	Residue	Modelled	Actual	Comment	Reference
D	334	ALA	ILE	engineered mutation	UNP Q93TK0
D	341	ASP	GLY	engineered mutation	UNP Q93TK0
D	524	GLY	-	expression tag	UNP Q93TK0

Continued from previous page...

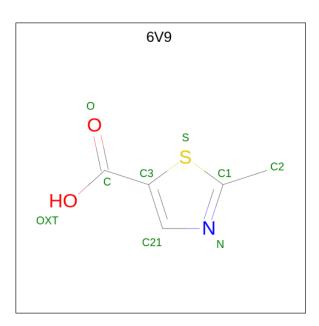
• Molecule 2 is ADENOSINE MONOPHOSPHATE (three-letter code: AMP) (formula: $C_{10}H_{14}N_5O_7P$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
2	А	1	Total	С	Η	Ν	Ο	Р	0	0
	Л	1	29	10	6	5	7	1	0	0
2	В	1	Total	С	Η	Ν	Ο	Р	0	0
	D	1	29	10	6	5	7	1	0	0
2	C	1	Total	С	Η	Ν	Ο	Р	0	0
	U	1	29	10	6	5	7	1	0	0
2	П	1	Total	С	Η	Ν	0	Р	0	0
		1	29	10	6	5	7	1		0

• Molecule 3 is 2-methyl-1,3-thiazole-5-carboxylic acid (three-letter code: 6V9) (formula: $C_5H_5NO_2S$) (labeled as "Ligand of Interest" by depositor).





Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	Δ	1	Total C H N O S	0	0
5	Л	1	9 5 1 1 1 1	0	0
3	В	1	Total C H N O S	0	0
5	D	1	9 5 1 1 1 1	0	0
3	С	1	Total C H N O S	0	0
5	U	1	9 5 1 1 1 1	0	0
2	Л	1	Total C H N O S	0	0
0		1	9 5 1 1 1 1		0

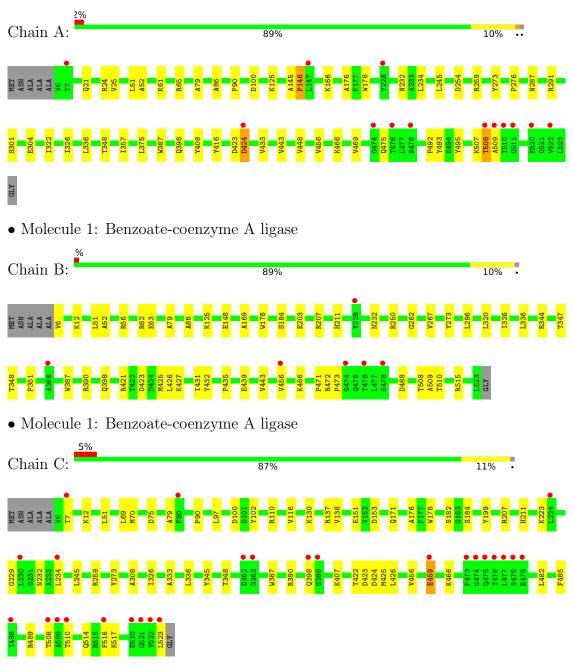
• Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	А	350	Total O 350 350	0	0
4	В	357	Total O 357 357	0	0
4	С	341	Total O 341 341	0	0
4	D	329	Total O 329 329	0	0

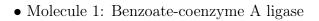


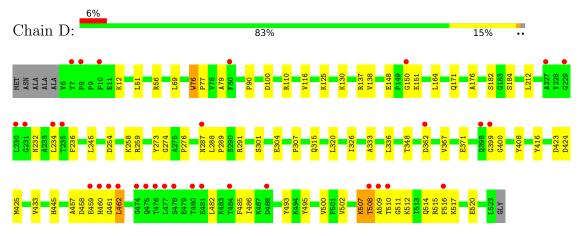
3 Residue-property plots (i)

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: Benzoate-coenzyme A ligase







4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants	98.05Å 95.14Å 119.52Å	Depositor
a, b, c, α , β , γ	90.00° 110.58° 90.00°	Depositor
Resolution (Å)	27.78 - 2.14	Depositor
Resolution (A)	27.78 - 2.14	EDS
% Data completeness	99.4 (27.78-2.14)	Depositor
(in resolution range)	99.5(27.78-2.14)	EDS
R _{merge}	0.20	Depositor
R _{sym}	(Not available)	Depositor
$< I/\sigma(I) > 1$	$6.01 (at 2.14 \text{\AA})$	Xtriage
Refinement program	PHENIX (1.11.1_2575: ???)	Depositor
D D.	0.160 , 0.198	Depositor
R, R_{free}	0.165 , 0.201	DCC
R_{free} test set	2003 reflections $(1.77%)$	wwPDB-VP
Wilson B-factor $(Å^2)$	23.8	Xtriage
Anisotropy	0.055	Xtriage
Bulk solvent $k_{sol}(e/Å^3), B_{sol}(Å^2)$	0.35 , 45.6	EDS
L-test for twinning ²	$ \langle L \rangle = 0.49, \langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	17180	wwPDB-VP
Average B, all atoms $(Å^2)$	24.0	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 77.70 % 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 8.0603e-07. The detected translational NCS is most likely also responsible for the elevated intensity ratio.

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.



¹Intensities estimated from amplitudes.

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: $6\mathrm{V9},$ 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		
	Unam	RMSZ	# Z > 5	RMSZ	# Z > 5	
1	А	0.40	0/4026	0.56	0/5494	
1	В	0.40	0/4026	0.56	0/5494	
1	С	0.39	0/4018	0.56	0/5483	
1	D	0.38	0/4026	0.55	0/5494	
All	All	0.39	0/16096	0.56	0/21965	

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	А	3914	0	3835	36	2
1	В	3914	0	3835	39	0
1	С	3909	0	3824	39	1
1	D	3914	0	3835	75	2
2	А	23	6	12	0	0
2	В	23	6	12	0	0
2	С	23	6	12	0	0
2	D	23	6	12	0	0
3	А	8	1	0	0	0

Continued on next page...



Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	В	8	1	0	0	0
3	С	8	1	0	1	0
3	D	8	1	0	1	0
4	А	350	0	0	4	0
4	В	357	0	0	7	0
4	С	341	0	0	5	1
4	D	329	0	0	7	1
All	All	17152	28	15377	187	4

Continued from previous page...

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 6.

The worst 5 of 187 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:D:254:ASP:OD1	1:D:287:ASN:ND2	1.88	1.07
1:D:508:THR:HG23	1:D:512:LYS:O	1.51	1.06
1:D:460:HIS:HB2	1:D:462:LEU:CD2	1.83	1.05
1:D:100:ASP:OD1	1:D:130:LYS:NZ	1.94	1.00
1:D:460:HIS:HB2	1:D:462:LEU:HD22	1.45	0.94

All (4) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:61:LYS:NZ	$1:A:508:THR:O[2_655]$	1.77	0.43
1:C:137:ARG:NH2	$1:D:520:GLU:O[2_556]$	2.08	0.12
1:A:100:ASP:OD2	$1:D:100:ASP:CB[1_655]$	2.10	0.10
4:C:998:HOH:O	4:D:982:HOH:O[2_556]	2.14	0.06

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	n Analysed Favoured Allowed		Outliers	Perce	entiles	
1	А	519/524~(99%)	507~(98%)	8 (2%)	4 (1%)	19	11
1	В	519/524~(99%)	507~(98%)	11 (2%)	1 (0%)	47	45
1	С	518/524~(99%)	505~(98%)	13 (2%)	0	100	100
1	D	519/524~(99%)	504 (97%)	13 (2%)	2~(0%)	34	29
All	All	2075/2096~(99%)	2023 (98%)	45 (2%)	7~(0%)	41	36

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	А	508	THR
1	А	146	PRO
1	В	509	ALA
1	D	508	THR
1	А	509	ALA

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	totameric Outliers		Percentiles		
1	А	395/412~(96%)	391~(99%)	4 (1%)	76	79		
1	В	395/412~(96%)	392~(99%)	3 (1%)	81	85		
1	С	394/412~(96%)	390~(99%)	4 (1%)	76	79		
1	D	395/412~(96%)	389~(98%)	6(2%)	65	68		
All	All	1579/1648~(96%)	1562 (99%)	17 (1%)	73	76		

5 of 17 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	462	LEU
1	D	510	THR
1	С	182	SER
1	С	390	ARG
1	С	426	LEU



Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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)

8 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).

Mal	Mol Type		e Chain Res Lin			Bond lengths			Bond angles		
WIOI	Type	Ullaili	nes	LIIIK	Counts	RMSZ	# Z >2	Counts	RMSZ	# Z >2	
3	6V9	С	602	2	6,8,9	5.92	3 (50%)	$3,\!10,\!12$	3.27	2 (66%)	
2	AMP	С	601	3	22,25,25	2.46	8 (36%)	25,38,38	<mark>3.52</mark>	7 (28%)	
3	6V9	В	602	2	6,8,9	5.13	1 (16%)	3,10,12	3.47	1 (33%)	
2	AMP	В	601	3	22,25,25	2.17	7 (31%)	25,38,38	1.86	4 (16%)	
3	6V9	D	602	2	6,8,9	5.39	2 (33%)	3,10,12	3.82	2 (66%)	
2	AMP	D	601	3	22,25,25	1.98	7 (31%)	25,38,38	1.78	3 (12%)	
2	AMP	А	601	3	22,25,25	2.25	6 (27%)	25,38,38	1.79	3 (12%)	
3	6V9	А	602	2	6,8,9	4.74	1 (16%)	3,10,12	3.40	1 (33%)	

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.



6M2T

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	6V9	С	602	2	-	0/0/2/4	0/1/1/1
2	AMP	С	601	3	-	0/6/26/26	0/3/3/3
3	6V9	В	602	2	-	0/0/2/4	0/1/1/1
2	AMP	В	601	3	-	0/6/26/26	0/3/3/3
3	6V9	D	602	2	-	0/0/2/4	0/1/1/1
2	AMP	D	601	3	-	1/6/26/26	0/3/3/3
2	AMP	А	601	3	-	0/6/26/26	0/3/3/3
3	6V9	А	602	2	-	0/0/2/4	0/1/1/1

'-' means no outliers of that kind were identified.

The worst 5 of 35 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	С	602	6V9	С3-С	-14.09	1.34	1.48
3	D	602	6V9	С3-С	-12.83	1.35	1.48
3	В	602	6V9	С3-С	-12.33	1.35	1.48
3	А	602	6V9	С3-С	-11.30	1.36	1.48
2	С	601	AMP	O4'-C1'	6.08	1.49	1.41

The worst 5 of 23 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
2	С	601	AMP	O3P-P-O5'	-10.32	79.26	106.73
2	С	601	AMP	O2P-P-O5'	-7.71	86.21	106.73
2	С	601	AMP	O5'-P-O1P	-7.24	86.16	106.47
2	В	601	AMP	N3-C2-N1	-6.88	117.93	128.68
2	D	601	AMP	N3-C2-N1	-6.50	118.52	128.68

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	D	601	AMP	C3'-C4'-C5'-O5'

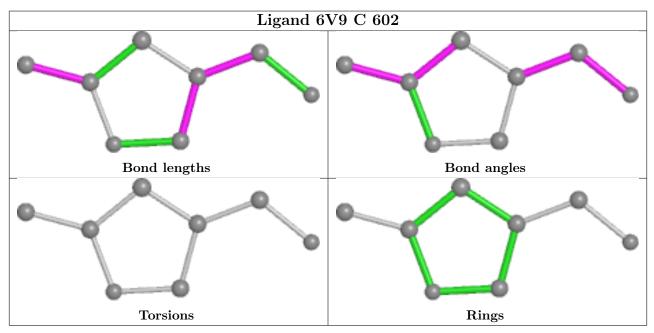
There are no ring outliers.

2 monomers are involved in 2 short contacts:

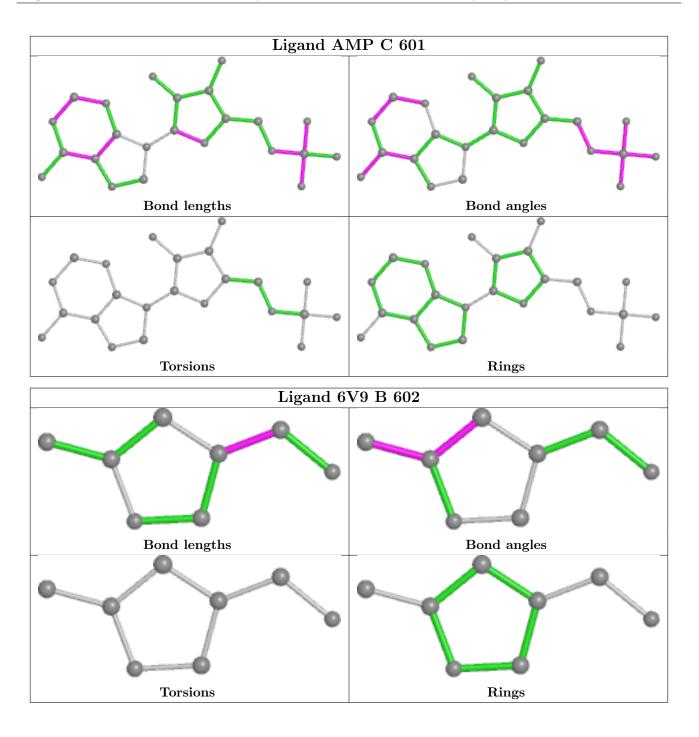
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	С	602	6V9	1	0
3	D	602	6V9	1	0



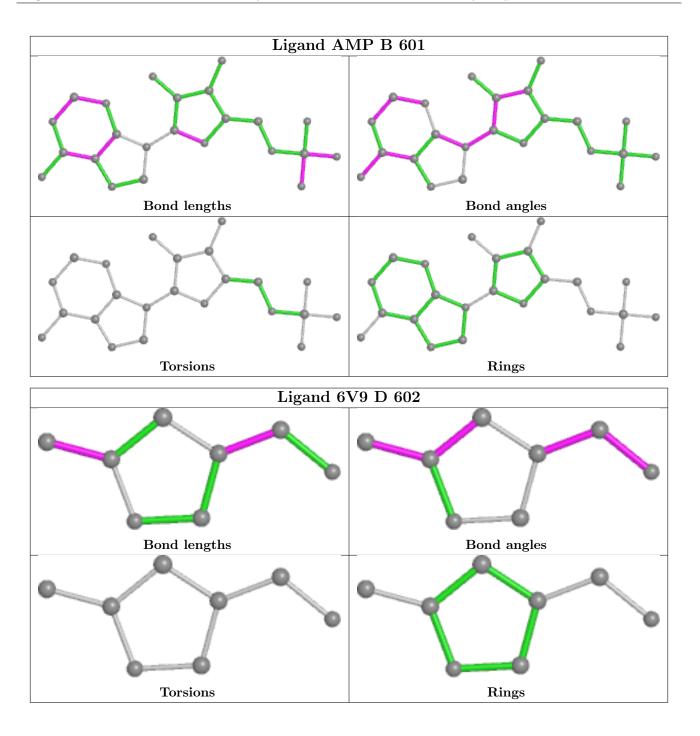
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 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 sufficient the outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.



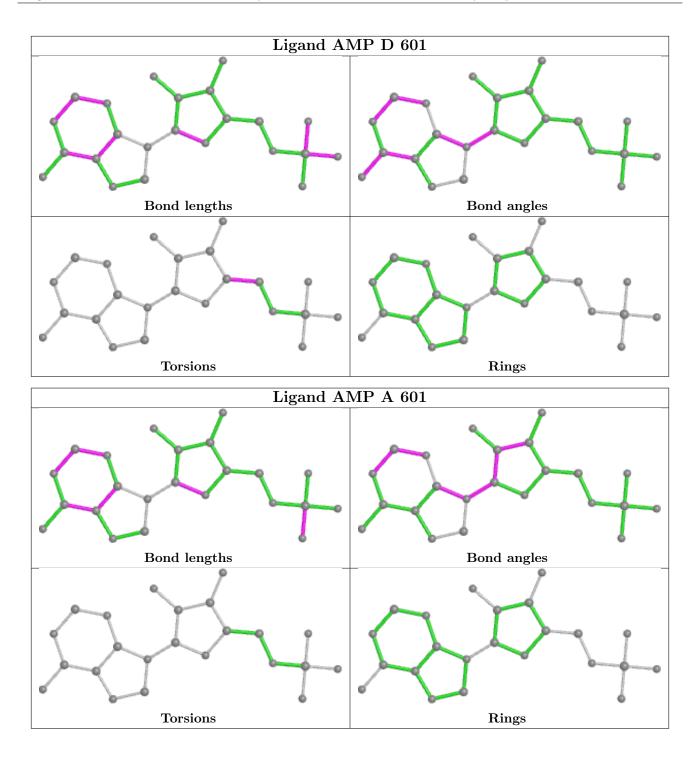




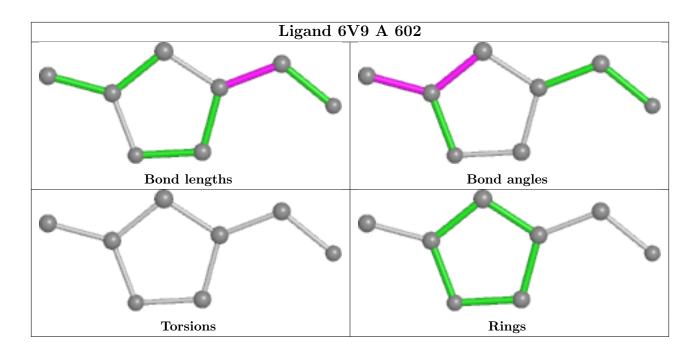












5.7 Other polymers (i)

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^{th} 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	$OWAB(A^2)$	Q<0.9
1	А	518/524~(98%)	-0.18	13 (2%) 57 64	11, 21, 38, 59	0
1	В	518/524~(98%)	-0.19	6 (1%) 79 83	12, 21, 38, 57	0
1	С	518/524~(98%)	-0.04	26 (5%) 28 35	12, 22, 47, 70	0
1	D	518/524~(98%)	0.03	32 (6%) 20 25	14, 24, 43, 73	0
All	All	2072/2096~(98%)	-0.10	77 (3%) 41 49	11, 22, 42, 73	0

The worst 5 of 77 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	508	THR	9.5
1	D	476	THR	8.3
1	С	474	GLY	7.4
1	С	522	VAL	7.1
1	D	462	LEU	6.4

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 monosaccharides in this entry.

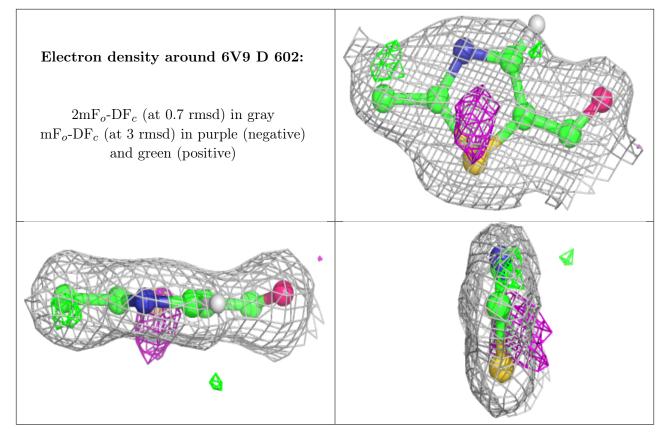
6.4 Ligands (i)

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^{th} 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.

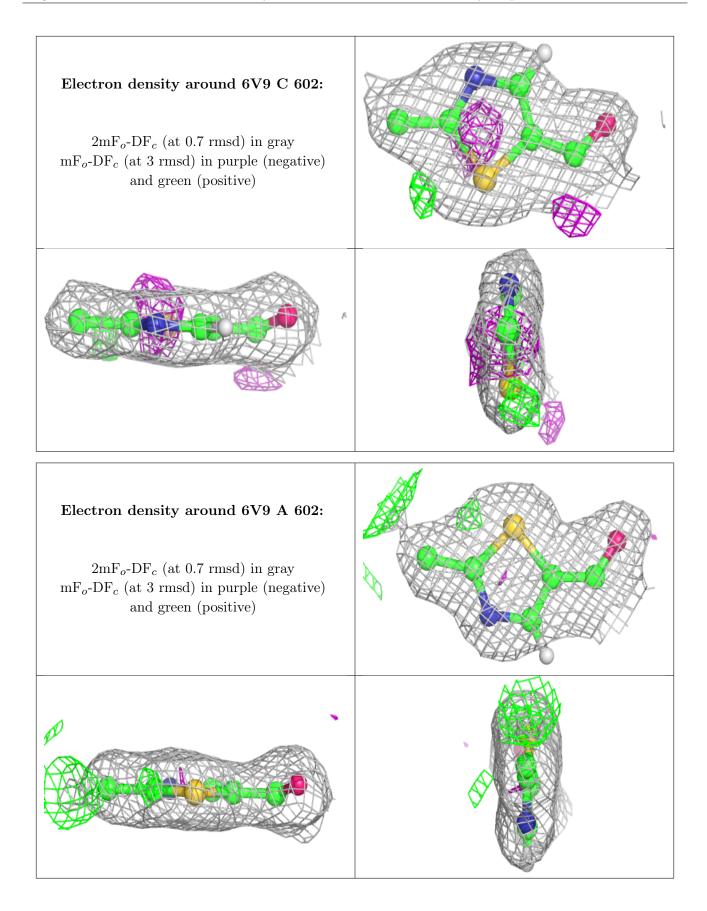


Mol	Type	Chain	Res	Atoms	RSCC	RSR	$\mathbf{B} ext{-factors}(\mathrm{\AA}^2)$	Q < 0.9
3	6V9	D	602	8/9	0.83	0.15	$18,\!25,\!32,\!39$	0
3	6V9	С	602	8/9	0.87	0.13	$17,\!22,\!29,\!38$	0
3	6V9	А	602	8/9	0.87	0.14	12,20,25,40	0
3	6V9	В	602	8/9	0.89	0.15	17,20,28,44	0
2	AMP	С	601	23/23	0.94	0.11	21,26,31,34	0
2	AMP	D	601	23/23	0.95	0.10	23,28,34,37	0
2	AMP	А	601	23/23	0.96	0.09	18,23,26,30	0
2	AMP	В	601	23/23	0.96	0.09	20,25,30,36	0

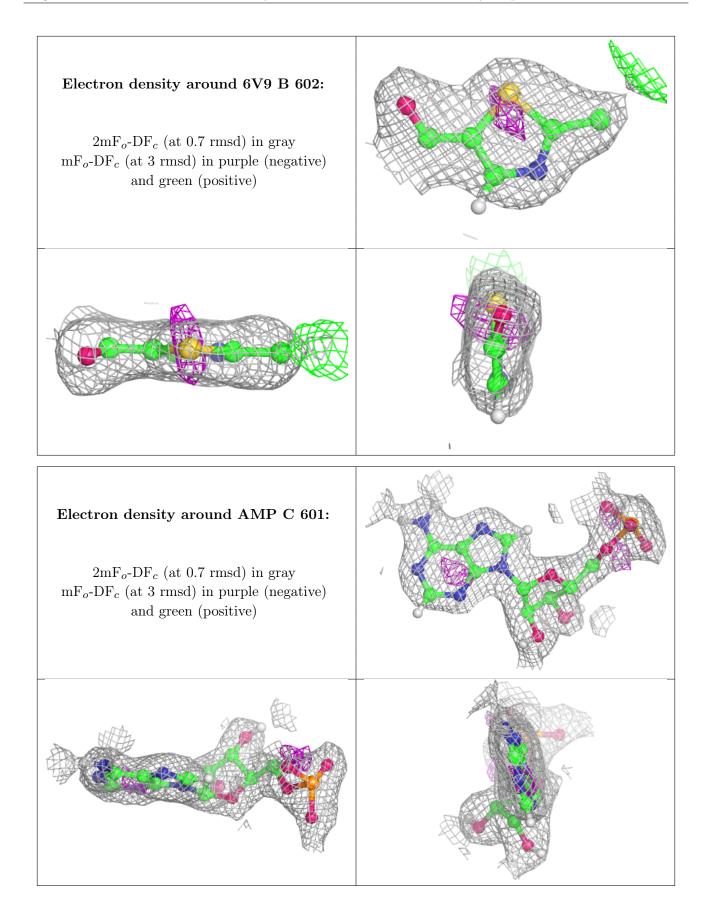
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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



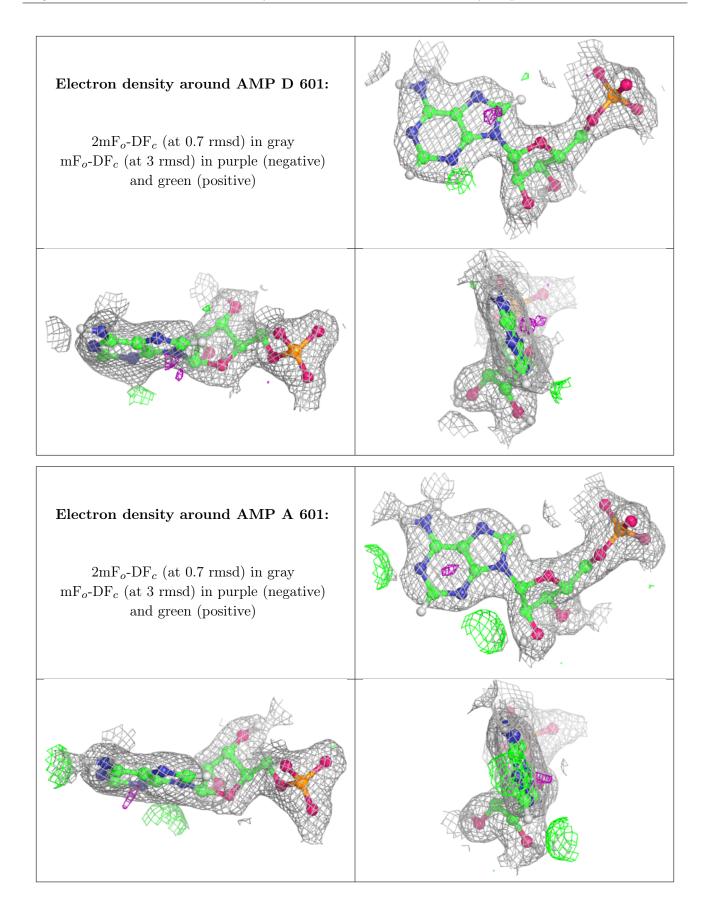




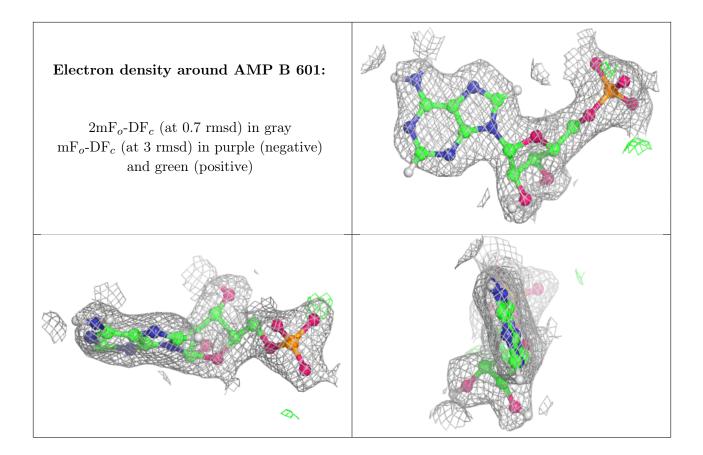












6.5 Other polymers (i)

There are no such residues in this entry.

