wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 12, 2024 – 11:05 am BST

PDB ID	:	9F99
Title	:	Crystal structure of MUS81-EME1 bound by compound 10.
Authors	:	Collie, G.W.
Deposited on	:	2024-05-07
Resolution	:	2.80 Å(reported)

The following versions of software and data (see references ⓘ) were used in the production of this report:

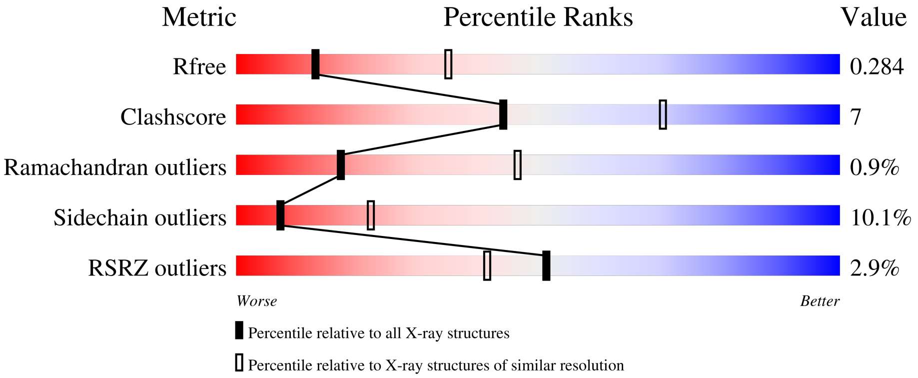
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36.2
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

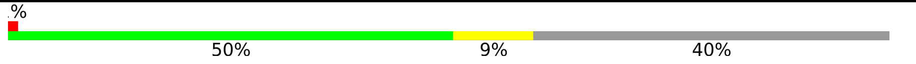
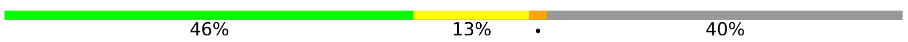
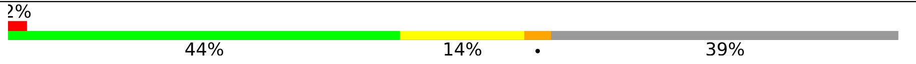
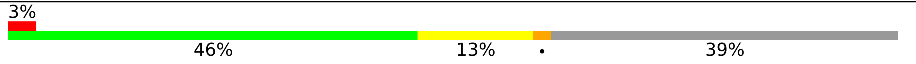
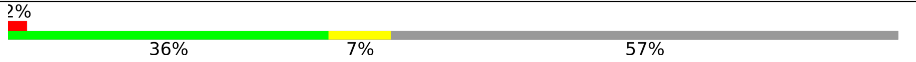
The reported resolution of this entry is 2.80 Å.

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.



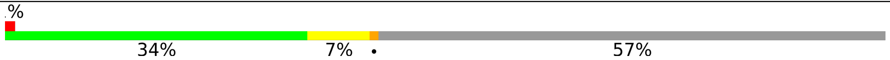
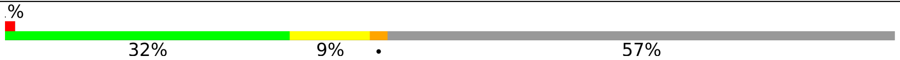
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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	A	308	
1	C	308	
1	E	308	
1	G	308	
2	B	326	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	D	326	
2	F	326	
2	H	326	

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 10145 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 Crossover junction endonuclease MUS81.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	185	1465	928	266	267	4	0	0	0
1	C	186	1467	928	267	268	4	0	0	0
1	E	187	1474	933	268	269	4	0	0	0
1	G	187	1469	931	264	270	4	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	244	GLY	-	expression tag	UNP Q96NY9
A	245	SER	-	expression tag	UNP Q96NY9
C	244	GLY	-	expression tag	UNP Q96NY9
C	245	SER	-	expression tag	UNP Q96NY9
E	244	GLY	-	expression tag	UNP Q96NY9
E	245	SER	-	expression tag	UNP Q96NY9
G	244	GLY	-	expression tag	UNP Q96NY9
G	245	SER	-	expression tag	UNP Q96NY9

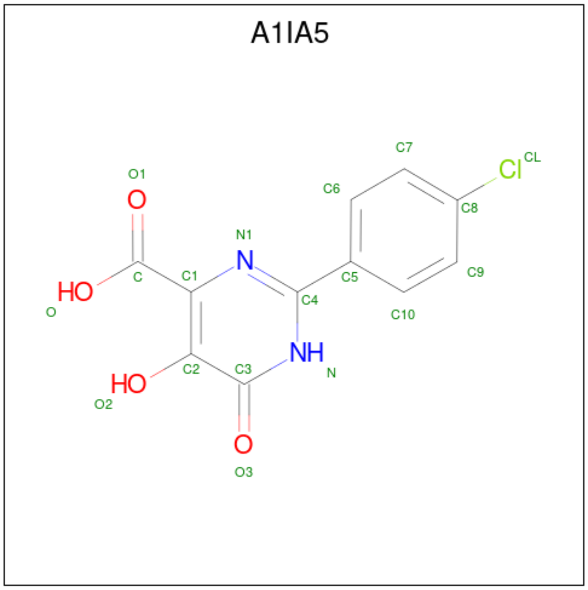
- Molecule 2 is a protein called Crossover junction endonuclease EME1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	139	1048	666	173	201	8	0	0	0
2	D	139	1048	666	173	201	8	0	0	0
2	F	139	1050	666	175	201	8	0	0	0
2	H	139	1044	663	172	201	8	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	245	GLY	-	expression tag	UNP Q96AY2
D	245	GLY	-	expression tag	UNP Q96AY2
F	245	GLY	-	expression tag	UNP Q96AY2
H	245	GLY	-	expression tag	UNP Q96AY2

- Molecule 3 is 2-(4-chlorophenyl)-5-oxidanyl-6-oxidanylidene-1H-pyrimidine-4-carboxylic acid (three-letter code: A1IA5) (formula: $C_{11}H_7ClN_2O_4$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C Cl N O 18 11 1 2 4	0	0
3	C	1	Total C Cl N O 18 11 1 2 4	0	0
3	E	1	Total C Cl N O 18 11 1 2 4	0	0
3	G	1	Total C Cl N O 18 11 1 2 4	0	0

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	2	Total Mg 2 2	0	0
4	C	2	Total Mg 2 2	0	0

Continued on next page...

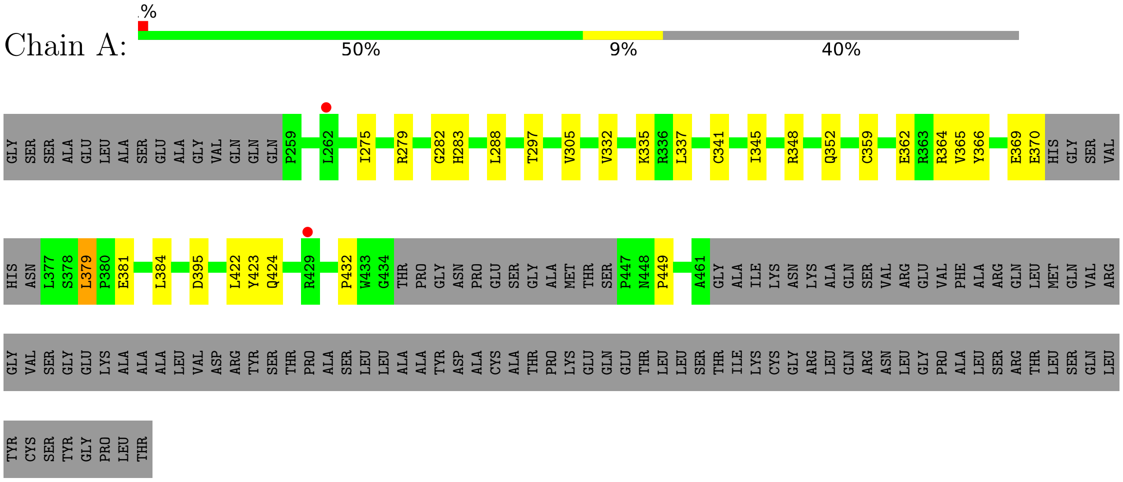
Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	E	2	Total Mg 2 2	0	0
4	G	2	Total Mg 2 2	0	0

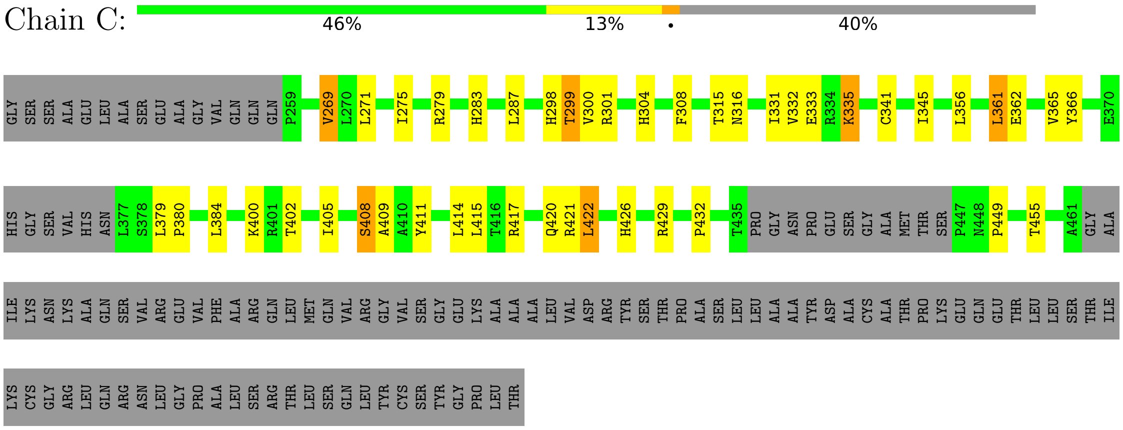
3 Residue-property plots ⓘ

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.

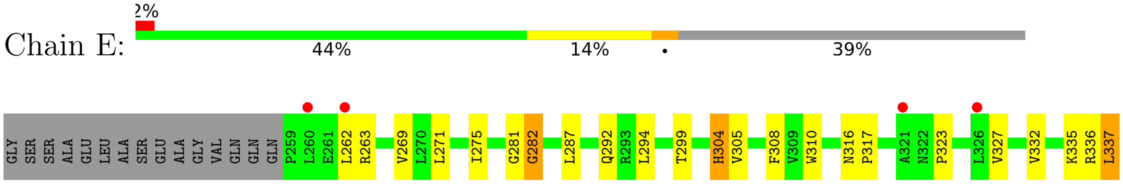
- Molecule 1: Crossover junction endonuclease MUS81

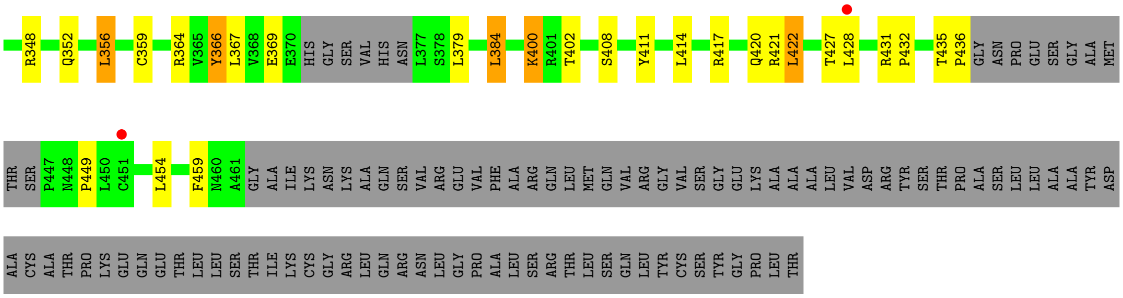


- Molecule 1: Crossover junction endonuclease MUS81



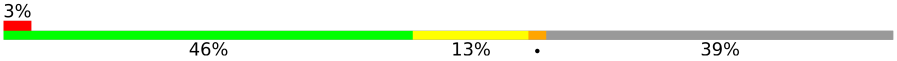
- Molecule 1: Crossover junction endonuclease MUS81

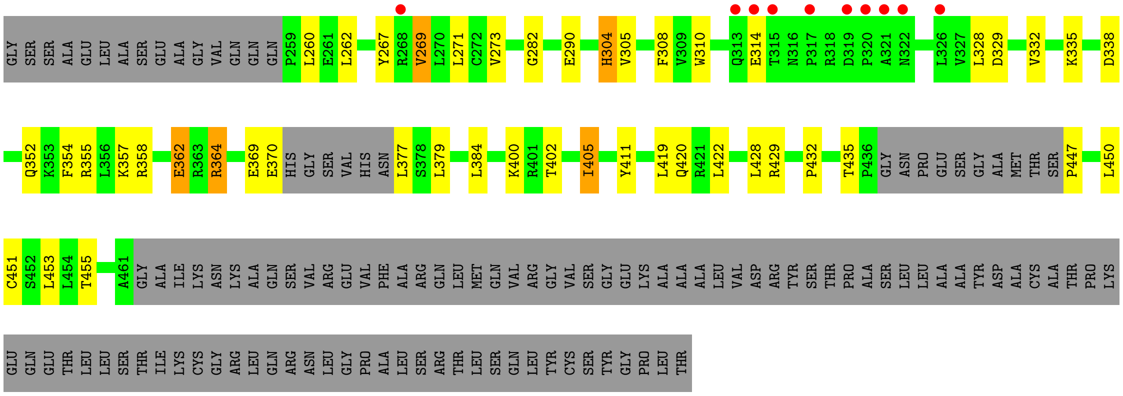




- Molecule 1: Crossover junction endonuclease MUS81

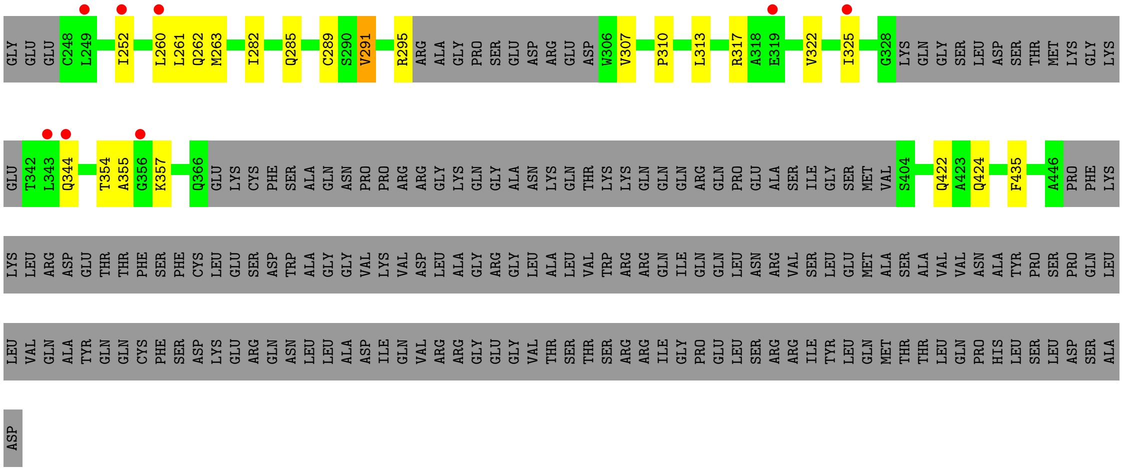
Chain G:





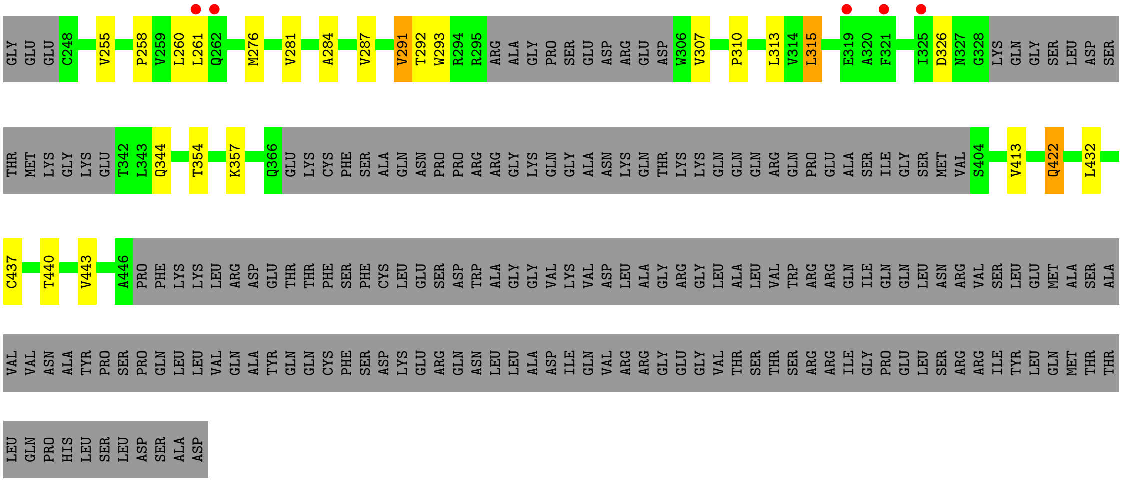
- Molecule 2: Crossover junction endonuclease EME1

Chain B:



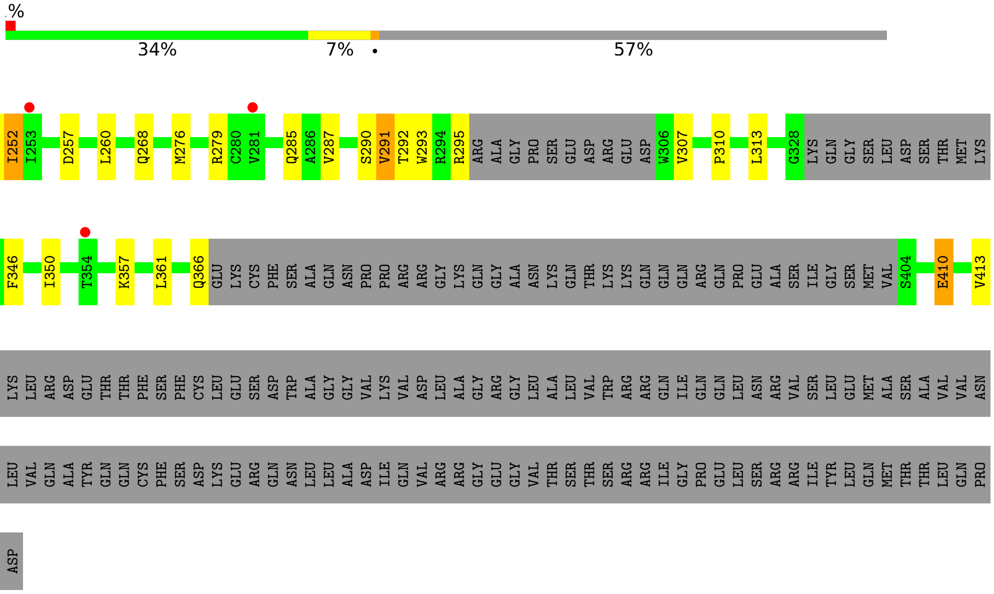
- Molecule 2: Crossover junction endonuclease EME1

Chain D:



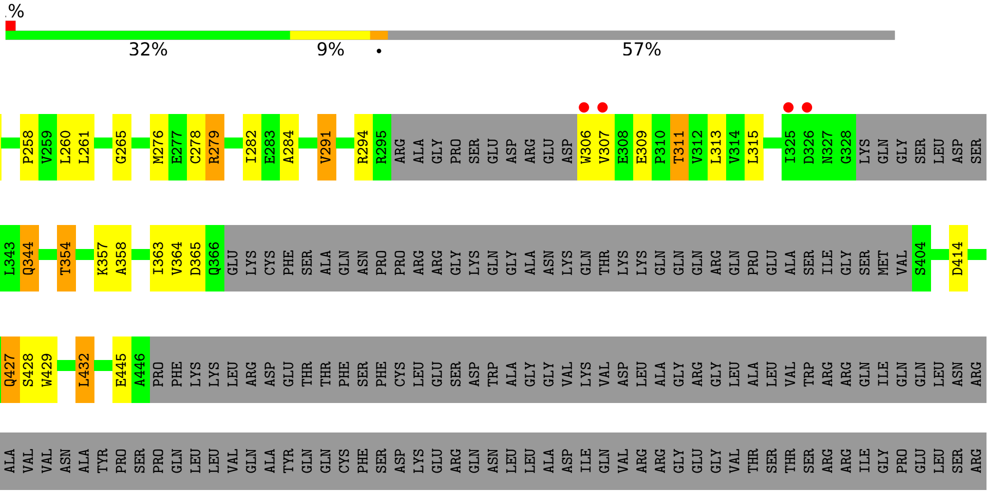
- Molecule 2: Crossover junction endonuclease EME1

Chain F:



- Molecule 2: Crossover junction endonuclease EME1

Chain H:




ARG
ILE
TYR
LEU
GLN
MET
THR
THR
LEU
GLN
PRO
HIS
LEU
SER
LEU
ASP
SER
ALA
ASP

4 Data and refinement statistics ⓘ

Property	**Value**	**Source**
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	111.35Å 125.10Å 218.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	38.95 – 2.80 38.95 – 2.80	Depositor EDS
% Data completeness (in resolution range)	51.1 (38.95-2.80) 51.2 (38.95-2.80)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.66 (at 2.81Å)	Xtriage
Refinement program	BUSTER 2.11.8	Depositor
R, R_{free}	0.240 , 0.280 0.240 , 0.284	Depositor DCC
R_{free} test set	1932 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	78.0	Xtriage
Anisotropy	0.111	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 43.6	EDS
L-test for twinning[2]	$< \lvert L \rvert > = 0.46$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	10145	wwPDB-VP
Average B, all atoms (Å^2)	84.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.78% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< \lvert L \rvert >$, $< L^2 >$ 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: A1IA5, MG

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	0.38	0/1493	0.62	1/2022 (0.0%)
1	C	0.40	0/1495	0.63	0/2027
1	E	0.36	0/1503	0.60	0/2039
1	G	0.36	0/1498	0.57	0/2033
2	B	0.31	0/1059	0.50	0/1439
2	D	0.33	0/1059	0.53	0/1439
2	F	0.32	0/1061	0.50	0/1442
2	H	0.32	0/1055	0.48	0/1435
All	All	0.35	0/10223	0.57	1/13876 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	369	GLU	C-N-CA	5.35	135.09	121.70

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	A	1465	0	1459	15	0
1	C	1467	0	1455	20	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	E	1474	0	1462	30	0
1	G	1469	0	1451	26	0
2	B	1048	0	1055	9	0
2	D	1048	0	1055	15	0
2	F	1050	0	1055	10	0
2	H	1044	0	1044	16	0
3	A	18	0	0	0	0
3	C	18	0	0	0	0
3	E	18	0	0	0	0
3	G	18	0	0	0	0
4	A	2	0	0	0	0
4	C	2	0	0	0	0
4	E	2	0	0	0	0
4	G	2	0	0	0	0
All	All	10145	0	10036	133	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 7.

The worst 5 of 133 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:335:LYS:NZ	1:A:352:GLN:HE22	1.58	1.00
1:A:335:LYS:HZ1	1:A:352:GLN:HE22	1.02	0.96
1:E:431:ARG:HD3	1:E:435:THR:HA	1.56	0.86
1:A:335:LYS:HZ1	1:A:352:GLN:NE2	1.71	0.86
2:D:258:PRO:HD2	2:D:284:ALA:HA	1.60	0.83

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	179/308 (58%)	172 (96%)	6 (3%)	1 (1%)	25	56
1	C	180/308 (58%)	172 (96%)	7 (4%)	1 (1%)	25	56
1	E	181/308 (59%)	176 (97%)	3 (2%)	2 (1%)	14	41
1	G	181/308 (59%)	170 (94%)	9 (5%)	2 (1%)	14	41
2	B	131/326 (40%)	127 (97%)	2 (2%)	2 (2%)	10	33
2	D	131/326 (40%)	120 (92%)	10 (8%)	1 (1%)	19	49
2	F	131/326 (40%)	122 (93%)	8 (6%)	1 (1%)	19	49
2	H	131/326 (40%)	122 (93%)	8 (6%)	1 (1%)	19	49
All	All	1245/2536 (49%)	1181 (95%)	53 (4%)	11 (1%)	17	46

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	432	PRO
2	D	307	VAL
1	E	282	GLY
1	E	432	PRO
1	G	282	GLY

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	158/260 (61%)	147 (93%)	11 (7%)	15	40
1	C	158/260 (61%)	138 (87%)	20 (13%)	4	14
1	E	159/260 (61%)	142 (89%)	17 (11%)	6	20
1	G	158/260 (61%)	141 (89%)	17 (11%)	6	19
2	B	113/275 (41%)	106 (94%)	7 (6%)	18	47
2	D	113/275 (41%)	105 (93%)	8 (7%)	14	39
2	F	113/275 (41%)	99 (88%)	14 (12%)	4	14
2	H	112/275 (41%)	97 (87%)	15 (13%)	4	12

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	1084/2140 (51%)	975 (90%)	109 (10%)	7 22

5 of 109 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	379	LEU
2	F	291	VAL
2	H	344	GLN
1	E	400	LYS
2	F	252	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
2	H	327	ASN
2	H	251	HIS
1	E	352	GLN
1	G	352	GLN
1	C	352	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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 12 ligands modelled in this entry, 8 are monoatomic - leaving 4 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	A1IA5	E	601	4	17,19,19	0.25	0	20,27,27	0.44	0
3	A1IA5	C	601	4	17,19,19	0.22	0	20,27,27	0.50	0
3	A1IA5	A	601	4	17,19,19	0.45	0	20,27,27	0.50	0
3	A1IA5	G	601	4	17,19,19	0.33	0	20,27,27	0.45	0

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	A1IA5	E	601	4	-	0/8/8/8	0/2/2/2
3	A1IA5	C	601	4	-	6/8/8/8	0/2/2/2
3	A1IA5	A	601	4	-	3/8/8/8	0/2/2/2
3	A1IA5	G	601	4	-	4/8/8/8	0/2/2/2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 13 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	601	A1IA5	O-C-C1-C2
3	A	601	A1IA5	O-C-C1-N1
3	C	601	A1IA5	O-C-C1-C2
3	C	601	A1IA5	O-C-C1-N1
3	C	601	A1IA5	O1-C-C1-C2

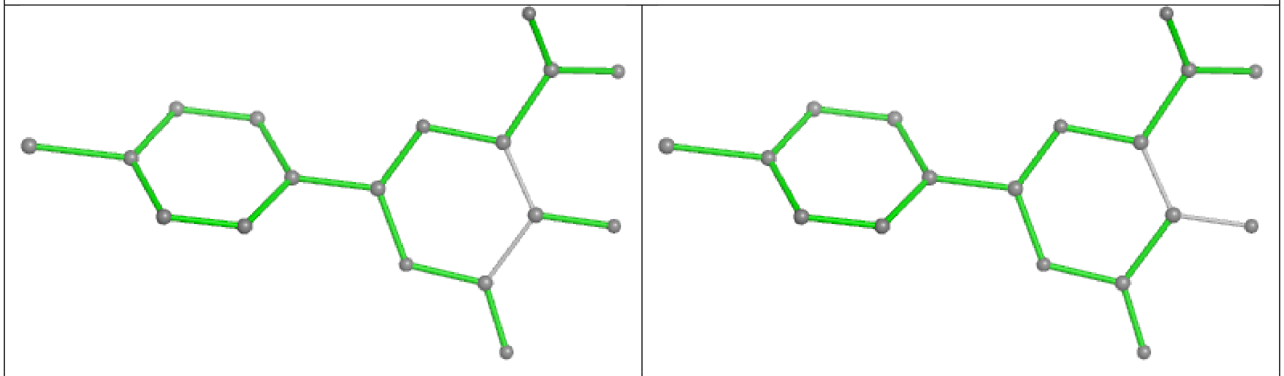
There are no ring outliers.

No monomer is involved in short contacts.

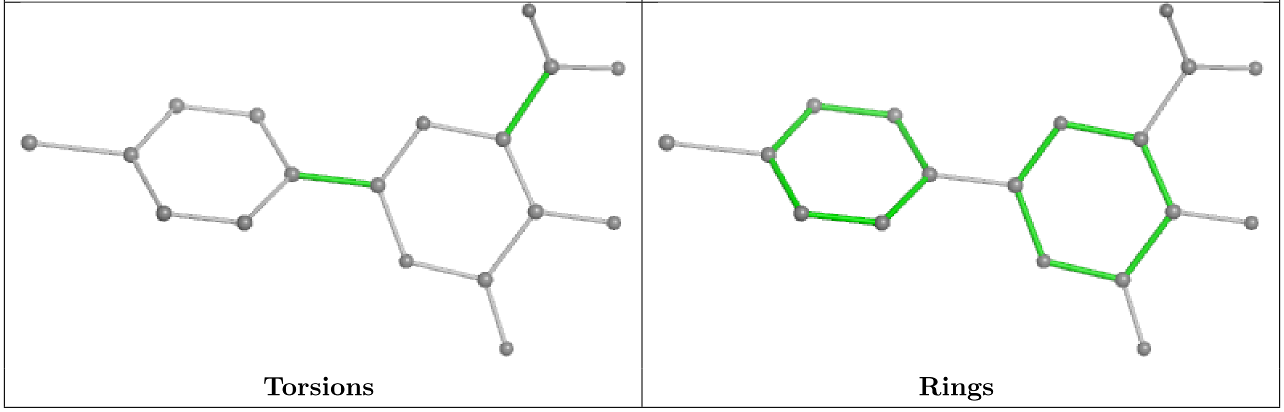
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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 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.

Ligand A1IA5 E 601

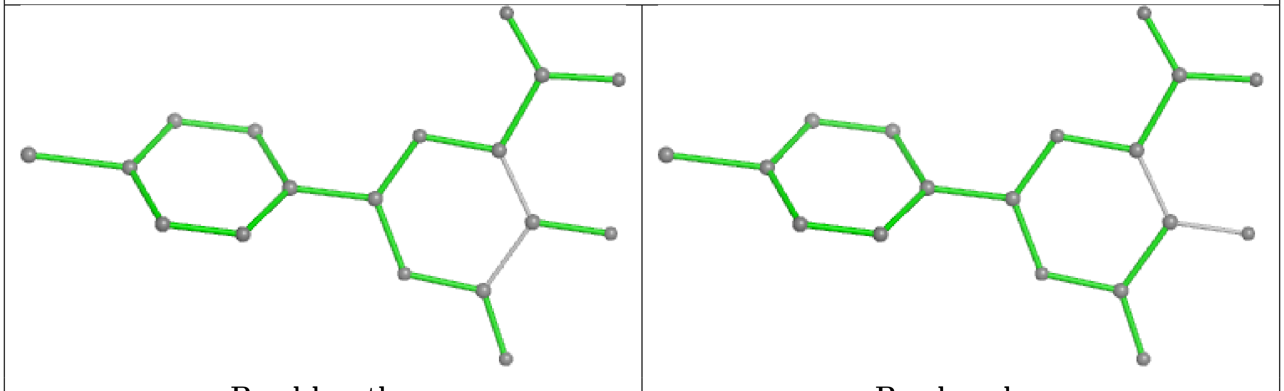


Bond lengths **Bond angles**



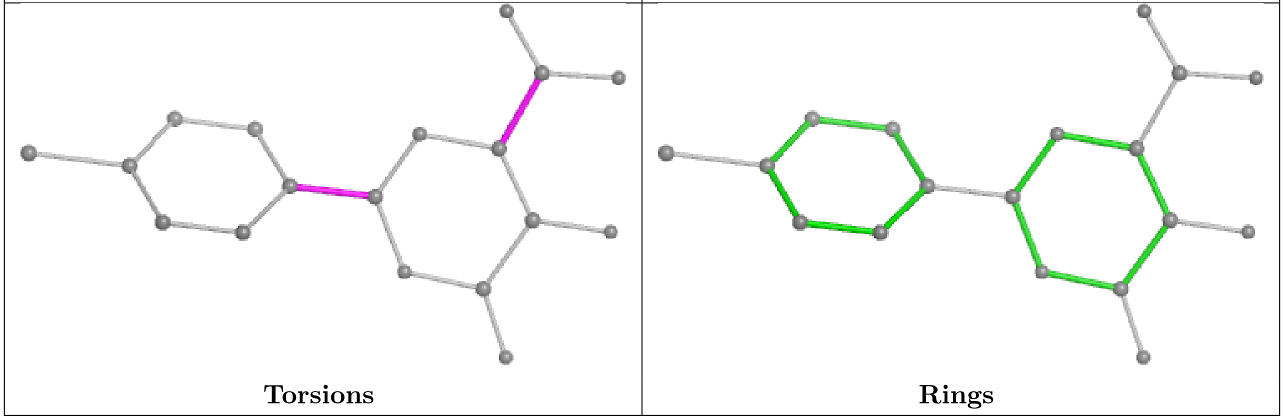
Torsions **Rings**

Ligand A1IA5 C 601



Bond lengths

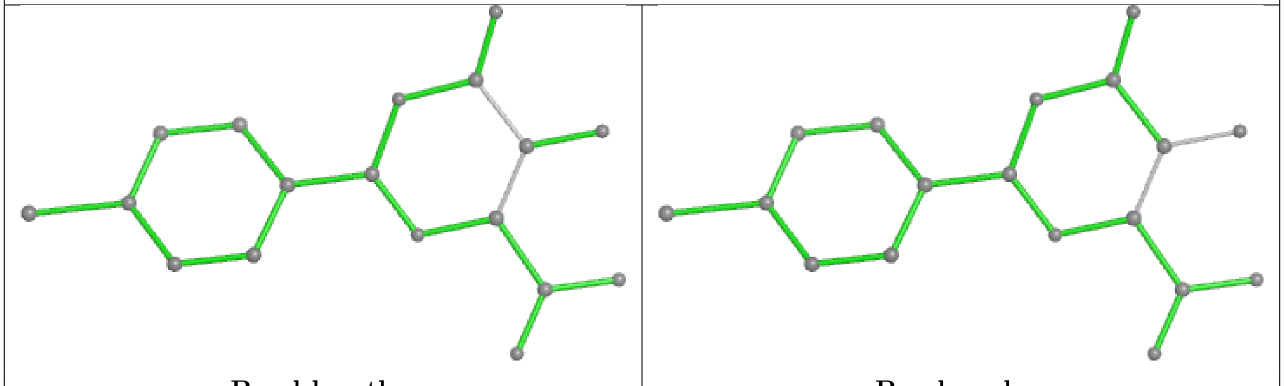
Bond angles



Torsions

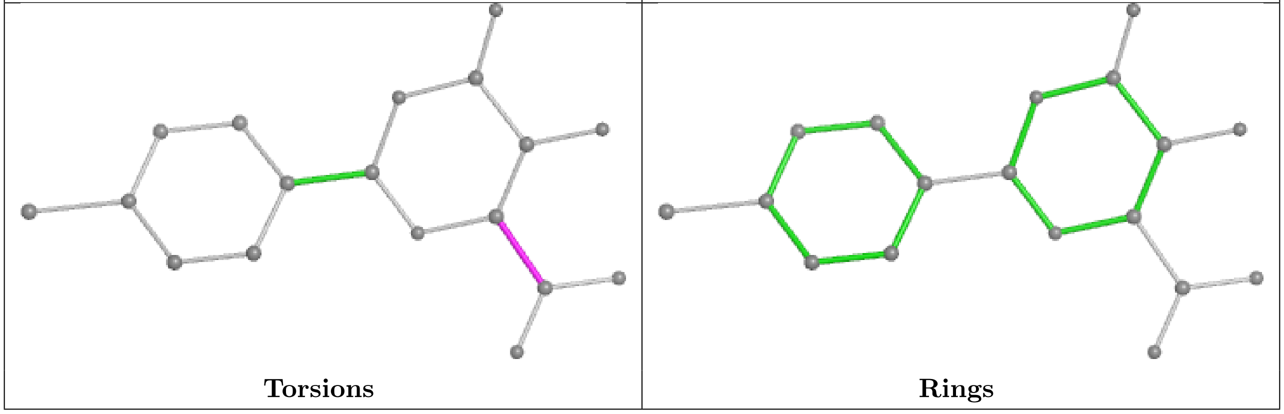
Rings

Ligand A1IA5 A 601



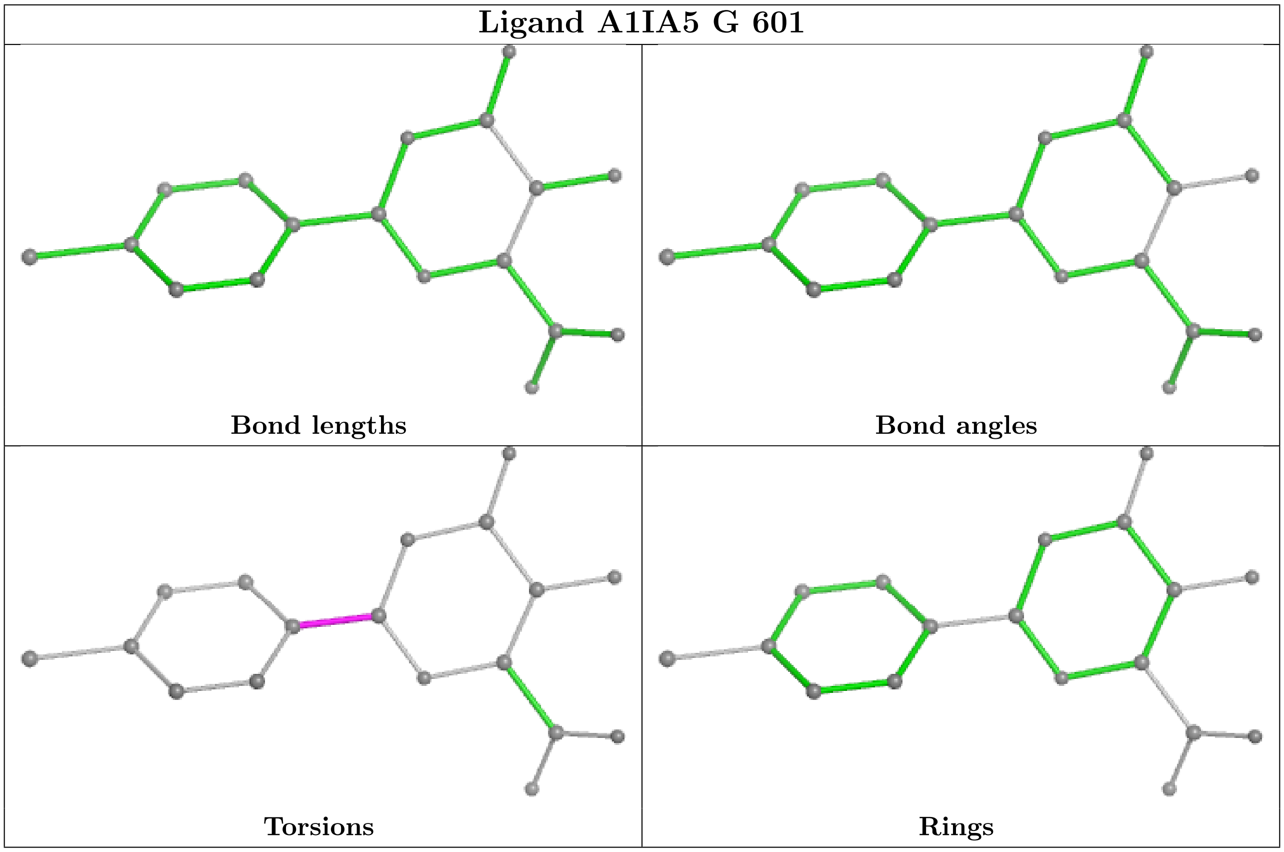
Bond lengths

Bond angles



Torsions

Rings



Ligand A1IA5 G 601

Bond lengths

Bond angles

Torsions

Rings

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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	A	185/308 (60%)	-0.28	2 (1%)	80	75	46, 68, 88, 97	0
1	C	186/308 (60%)	-0.41	0	100	100	49, 66, 88, 100	0
1	E	187/308 (60%)	-0.00	6 (3%)	47	37	54, 79, 104, 113	0
1	G	187/308 (60%)	0.08	10 (5%)	26	17	53, 82, 105, 111	0
2	B	139/326 (42%)	0.39	8 (5%)	23	15	74, 99, 125, 130	0
2	D	139/326 (42%)	0.22	5 (3%)	42	32	70, 97, 118, 121	0
2	F	139/326 (42%)	0.07	3 (2%)	62	52	70, 95, 112, 116	0
2	H	139/326 (42%)	-0.01	4 (2%)	51	41	75, 97, 117, 125	0
All	All	1301/2536 (51%)	-0.02	38 (2%)	51	41	46, 85, 114, 130	0

The worst 5 of 38 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	320	PRO	4.2
1	G	319	ASP	3.9
1	G	321	ALA	3.7
2	B	260	LEU	3.6
1	G	314	GLU	3.4

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

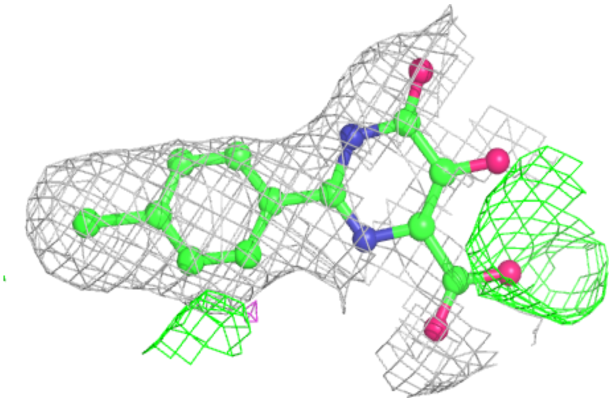
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

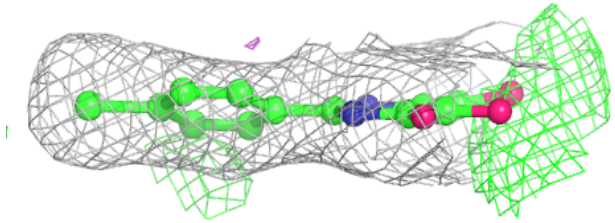
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($Å^2$)	Q<0.9
3	A1IA5	A	601	18/18	0.93	0.20	66,67,67,67	0
3	A1IA5	C	601	18/18	0.93	0.19	72,73,74,74	0
4	MG	E	603	1/1	0.93	0.18	51,51,51,51	0
4	MG	E	602	1/1	0.94	0.25	54,54,54,54	0
3	A1IA5	G	601	18/18	0.94	0.15	73,75,76,76	0
3	A1IA5	E	601	18/18	0.95	0.17	63,64,64,65	0
4	MG	C	602	1/1	0.95	0.24	57,57,57,57	0
4	MG	G	603	1/1	0.96	0.12	49,49,49,49	0
4	MG	A	602	1/1	0.97	0.35	57,57,57,57	0
4	MG	A	603	1/1	0.98	0.14	34,34,34,34	0
4	MG	G	602	1/1	0.99	0.16	43,43,43,43	0
4	MG	C	603	1/1	0.99	0.16	35,35,35,35	0

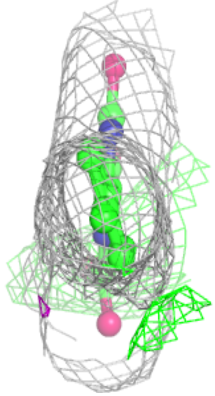
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around A1IA5 A 601:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

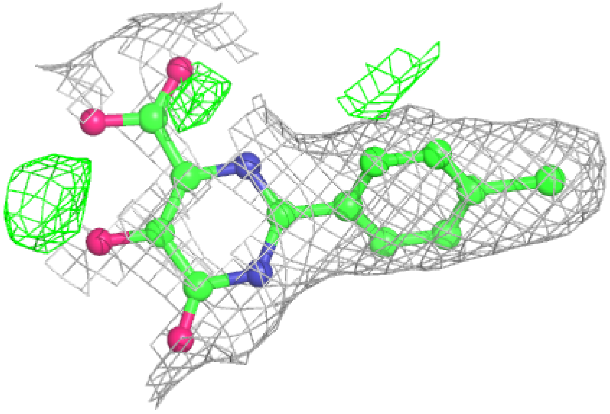


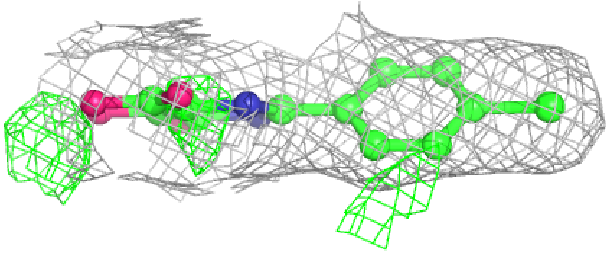


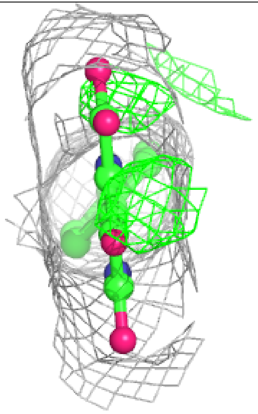


Electron density around A1IA5 C 601:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

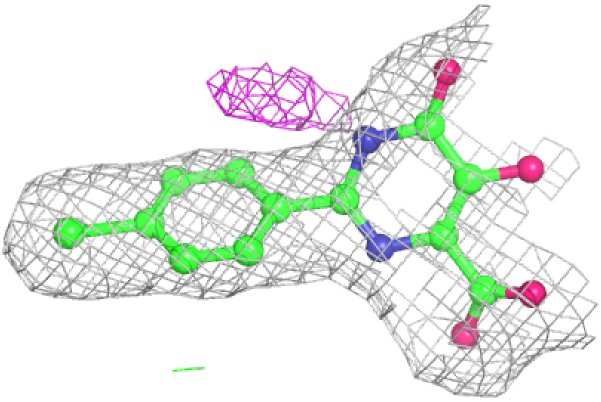


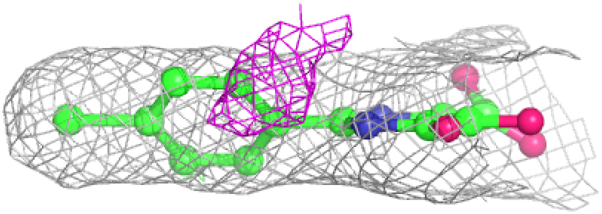


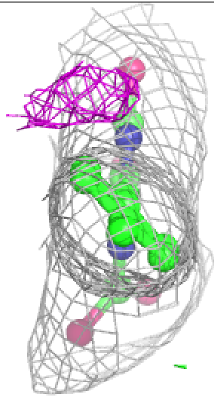


Electron density around A1IA5 G 601:

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
and green (positive)

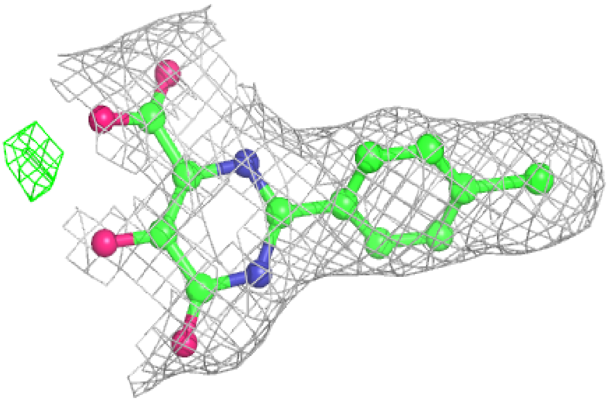


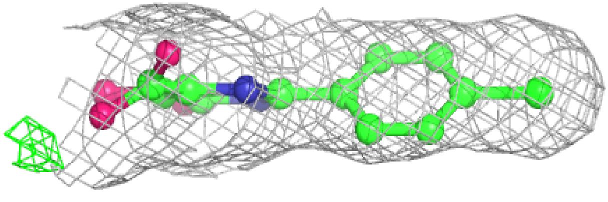


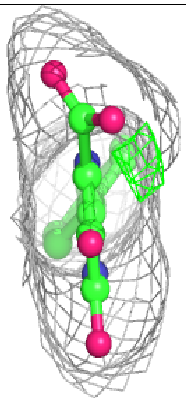


Electron density around A1IA5 E 601:

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
and green (positive)







6.5 Other polymers ⓘ

There are no such residues in this entry.