wwPDB X-ray Structure Validation Summary Report (i)

Dec 2, 2021 – 01:14 pm GMT

PDB ID	:	6ZBO
Title	:	HIF Prolyl Hydroxylase 2 (PHD2/EGLN1) in Complex with 1-(6-morpholino pyrimidin-4-yl)-4-(1H-1,2,3-triazol-1-yl)-1H-pyrazol-5-ol (Molidustat)
Authors	:	Figg Jr, W.D.; McDonough, M.A.; Nakashima, Y.; Holt-Martyn, J.P.; Schofield, C.J.
Deposited on	:	2020-06-08
Resolution	:	1.79 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

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

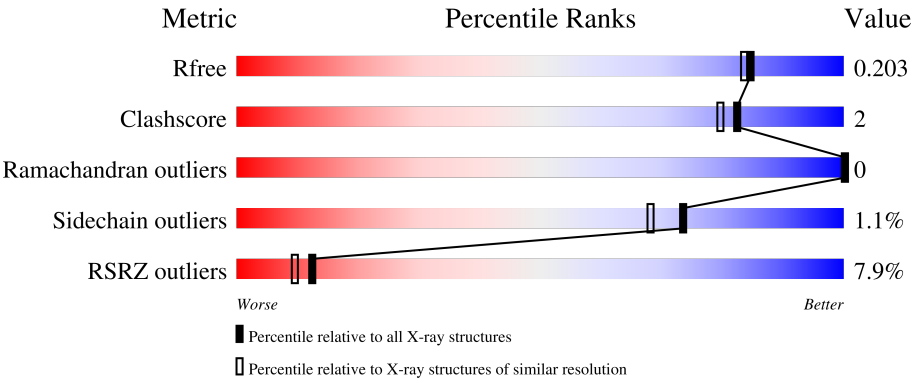
MolProbity	:	4.02b-467
Mogul	:	1.8.4 (270009), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

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 1.79 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.



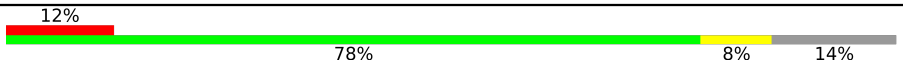
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)
RSRZ outliers	127900	5850 (1.80-1.80)

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	227	5% 84% 6% 10%
1	B	227	4% 85% 5% 10%
1	C	227	5% 84% 5% 11%
1	D	227	6% 86% • 11%
1	E	227	10% 78% 8% 13%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	227	

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 19772 atoms, of which 9394 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.

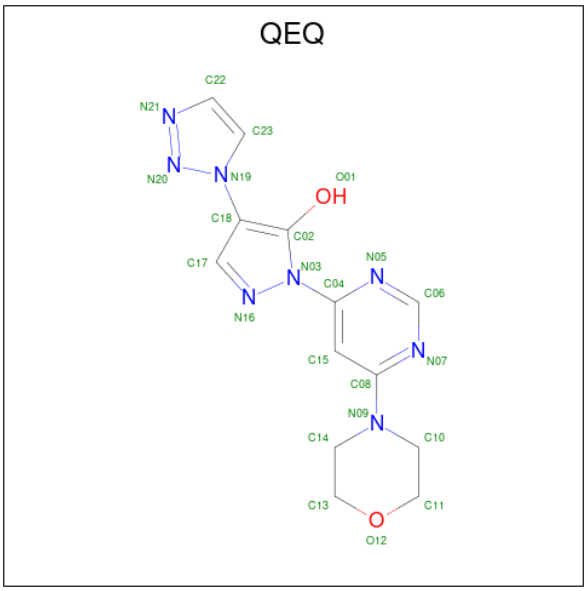
- Molecule 1 is a protein called Egl nine homolog 1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	204	Total	C	H	N	O	S	0	4	0
			3172	1032	1552	277	299	12			
1	B	205	Total	C	H	N	O	S	0	4	0
			3189	1031	1567	278	301	12			
1	C	202	Total	C	H	N	O	S	0	4	0
			3186	1026	1571	281	296	12			
1	D	202	Total	C	H	N	O	S	0	3	0
			3170	1021	1568	279	290	12			
1	E	197	Total	C	H	N	O	S	0	4	0
			3115	1000	1540	277	286	12			
1	F	196	Total	C	H	N	O	S	0	4	0
			3074	994	1512	271	285	12			

- Molecule 2 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Mn	0	0
			1	1		
2	B	1	Total	Mn	0	0
			1	1		
2	C	1	Total	Mn	0	0
			1	1		
2	D	1	Total	Mn	0	0
			1	1		
2	E	1	Total	Mn	0	0
			1	1		
2	F	1	Total	Mn	0	0
			1	1		

- Molecule 3 is 2-(6-morpholin-4-ylpyrimidin-4-yl)-4-(1,2,3-triazol-1-yl)pyrazol-3-ol (three-letter code: QEQ) (formula: $C_{13}H_{14}N_8O_2$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	H	N	O	0	0
			37	13	14	8	2		
3	B	1	Total	C	H	N	O	0	0
			37	13	14	8	2		
3	C	1	Total	C	H	N	O	0	0
			37	13	14	8	2		
3	D	1	Total	C	H	N	O	0	0
			37	13	14	8	2		
3	E	1	Total	C	H	N	O	0	0
			37	13	14	8	2		
3	F	1	Total	C	H	N	O	0	0
			37	13	14	8	2		

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	3	Total	Cl	0	0
			3	3		
4	B	2	Total	Cl	0	0
			2	2		
4	C	1	Total	Cl	0	0
			1	1		
4	D	2	Total	Cl	0	0
			2	2		
4	E	2	Total	Cl	0	0
			2	2		
4	F	3	Total	Cl	0	0
			3	3		

- Molecule 5 is water.

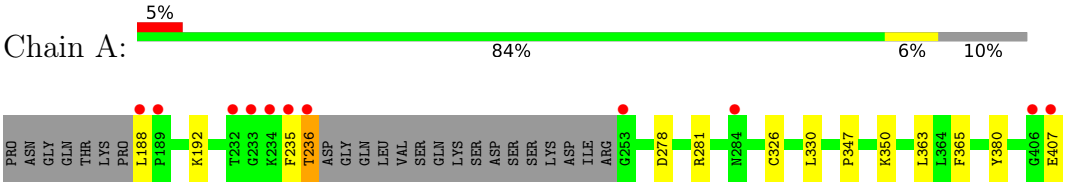
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	139	Total O 141 141	0	3
5	B	125	Total O 125 125	0	0
5	C	95	Total O 96 96	0	1
5	D	102	Total O 102 102	0	0
5	E	70	Total O 70 70	0	1
5	F	91	Total O 91 91	0	0

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.

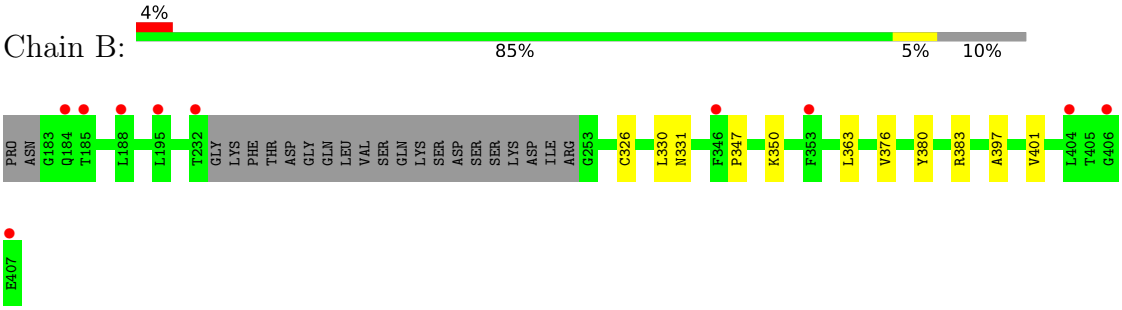
• Molecule 1: Egl nine homolog 1

Chain A:



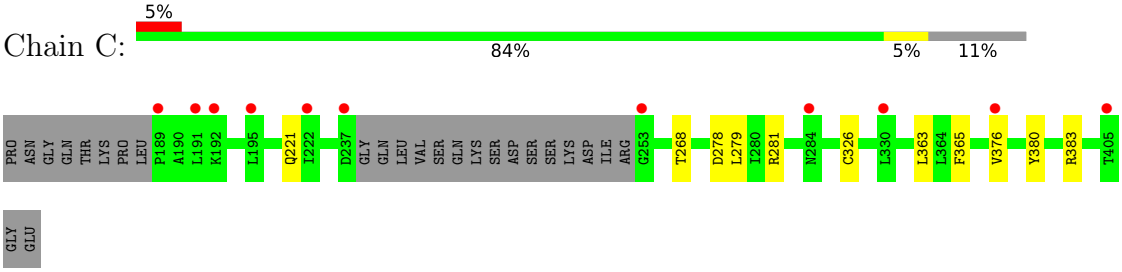
• Molecule 1: Egl nine homolog 1

Chain B:



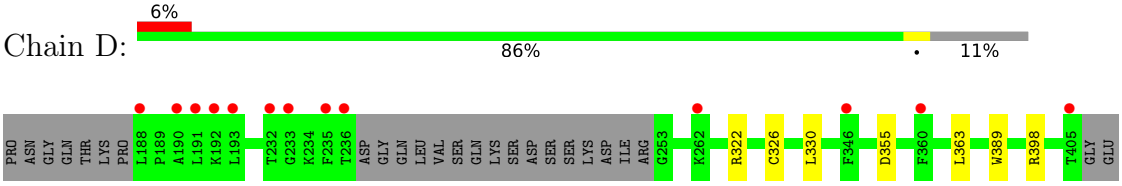
• Molecule 1: Egl nine homolog 1

Chain C:



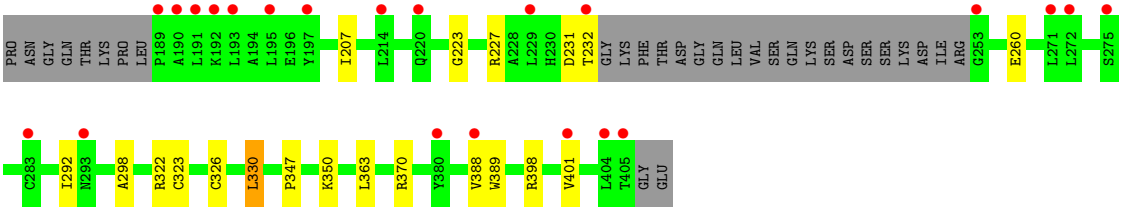
• Molecule 1: Egl nine homolog 1

Chain D:



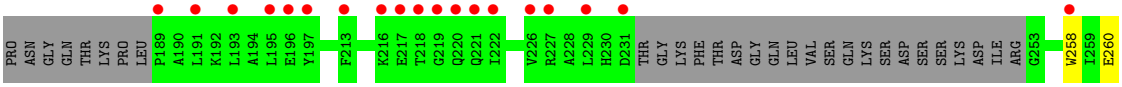
• Molecule 1: Egl nine homolog 1

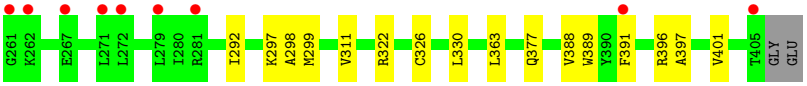
Chain E:



• Molecule 1: Egl nine homolog 1

Chain F:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	77.12Å 75.15Å 127.22Å 90.00° 95.31° 90.00°	Depositor
Resolution (Å)	53.71 – 1.79 53.71 – 1.79	Depositor EDS
% Data completeness (in resolution range)	98.2 (53.71-1.79) 98.2 (53.71-1.79)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.50 (at 1.80Å)	Xtriage
Refinement program	PHENIX 1.19.1	Depositor
R, R_{free}	0.180 , 0.206 0.178 , 0.203	Depositor DCC
R_{free} test set	6820 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å^2)	37.0	Xtriage
Anisotropy	0.497	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	19772	wwPDB-VP
Average B, all atoms (Å^2)	58.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.75% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: QEQ, MN, CL

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	0.42	0/1661	0.60	0/2251
1	B	0.39	0/1668	0.57	0/2257
1	C	0.38	0/1661	0.58	0/2245
1	D	0.33	0/1648	0.55	0/2227
1	E	0.34	0/1623	0.56	0/2193
1	F	0.35	0/1607	0.57	0/2173
All	All	0.37	0/9868	0.57	0/13346

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

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	1620	1552	1545	8	0
1	B	1622	1567	1560	6	0
1	C	1615	1571	1564	8	0
1	D	1602	1568	1562	3	0
1	E	1575	1540	1534	11	0
1	F	1562	1512	1505	9	0
2	A	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
2	E	1	0	0	0	0
2	F	1	0	0	0	0
3	A	23	14	0	0	0
3	B	23	14	0	0	0
3	C	23	14	0	0	0
3	D	23	14	0	0	0
3	E	23	14	0	0	0
3	F	23	14	0	0	0
4	A	3	0	0	0	0
4	B	2	0	0	1	0
4	C	1	0	0	0	0
4	D	2	0	0	1	0
4	E	2	0	0	0	0
4	F	3	0	0	0	0
5	A	141	0	0	2	0
5	B	125	0	0	3	0
5	C	96	0	0	2	0
5	D	102	0	0	1	0
5	E	70	0	0	2	0
5	F	91	0	0	1	0
All	All	10378	9394	9270	47	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 2.

The worst 5 of 47 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)
4:B:604:CL:CL	5:B:806:HOH:O	2.41	0.76
1:C:380:TYR:OH	5:C:701:HOH:O	2.09	0.70
1:B:380:TYR:OH	5:B:701:HOH:O	1.97	0.69
4:D:603:CL:CL	5:D:790:HOH:O	2.47	0.68
1:C:221:GLN:CG	1:C:268:THR:HG21	2.26	0.65

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	204/227 (90%)	197 (97%)	7 (3%)	0	100	100
1	B	205/227 (90%)	201 (98%)	4 (2%)	0	100	100
1	C	202/227 (89%)	198 (98%)	4 (2%)	0	100	100
1	D	201/227 (88%)	195 (97%)	6 (3%)	0	100	100
1	E	197/227 (87%)	194 (98%)	3 (2%)	0	100	100
1	F	196/227 (86%)	194 (99%)	2 (1%)	0	100	100
All	All	1205/1362 (88%)	1179 (98%)	26 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	164/190 (86%)	161 (98%)	3 (2%)	59	48
1	B	167/190 (88%)	166 (99%)	1 (1%)	86	84
1	C	168/190 (88%)	168 (100%)	0	100	100
1	D	167/190 (88%)	165 (99%)	2 (1%)	71	65
1	E	164/190 (86%)	160 (98%)	4 (2%)	49	36
1	F	160/190 (84%)	158 (99%)	2 (1%)	69	62
All	All	990/1140 (87%)	978 (99%)	12 (1%)	73	65

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	227[B]	ARG
1	E	260	GLU
1	F	396	ARG
1	E	330	LEU
1	B	330	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)

Of 25 ligands modelled in this entry, 19 are monoatomic - leaving 6 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

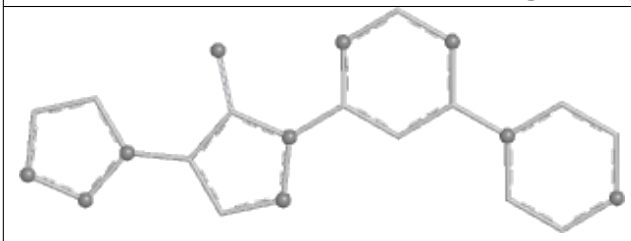
There are no ring outliers.

No monomer is involved in short contacts.

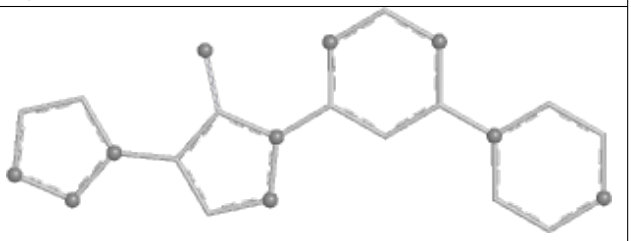
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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

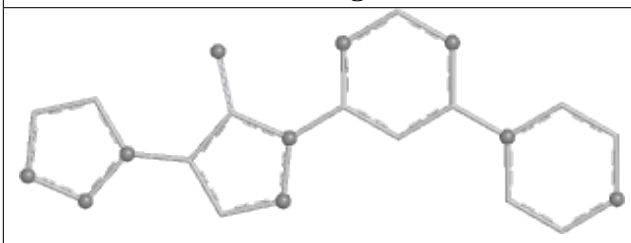
Ligand QEQ A 602



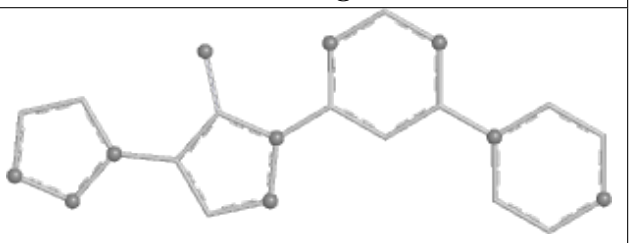
Bond lengths



Bond angles

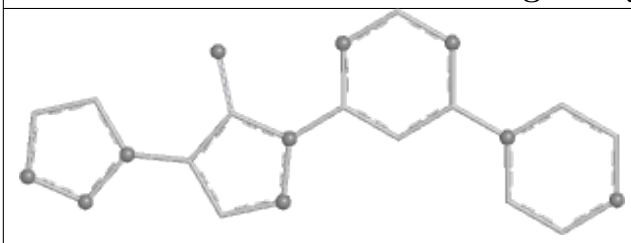


Torsions

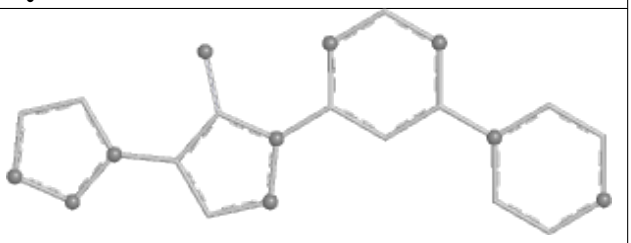


Rings

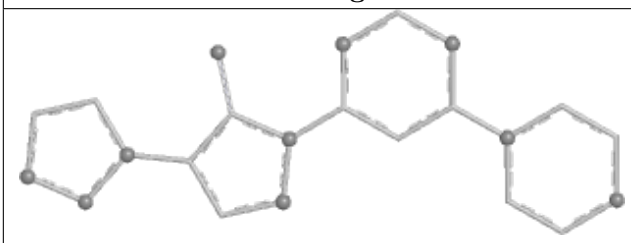
Ligand QEQ B 602



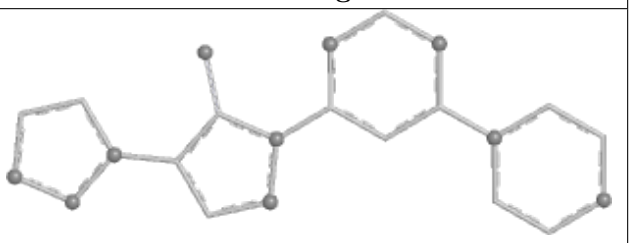
Bond lengths



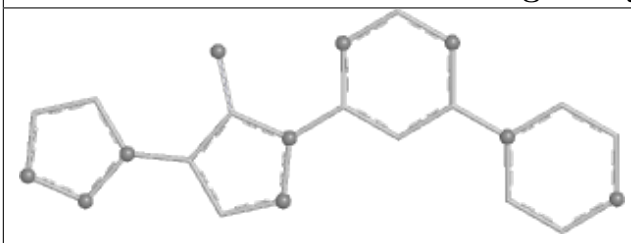
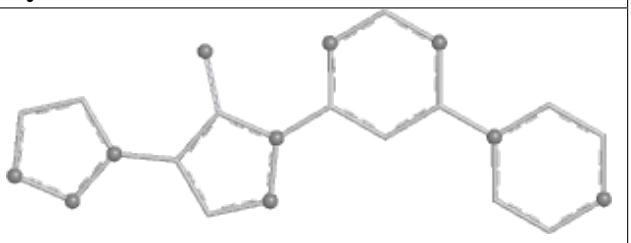
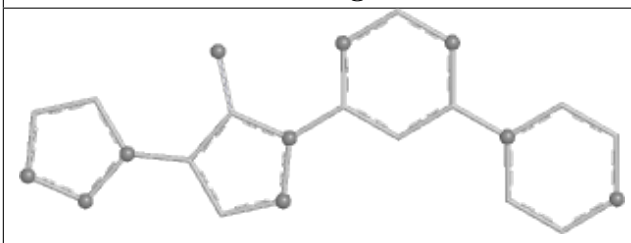
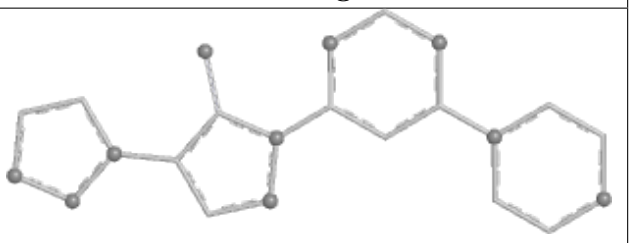
Bond angles

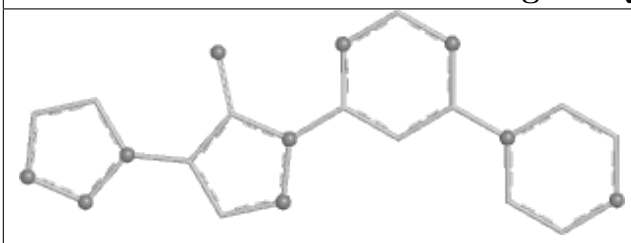
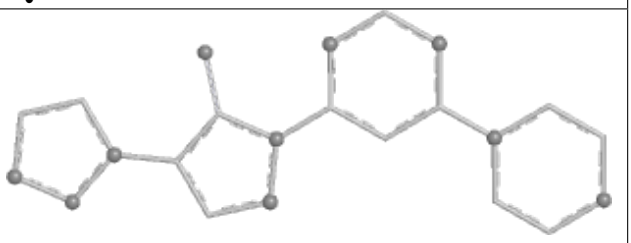
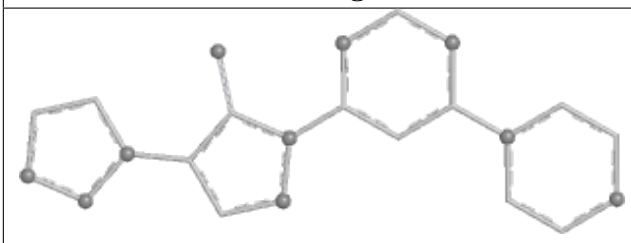
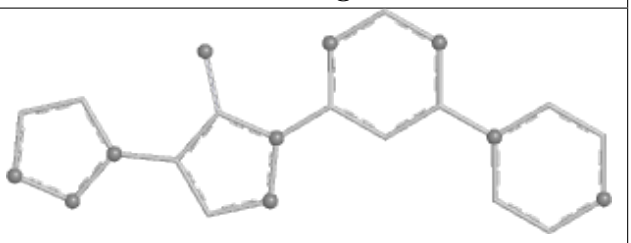


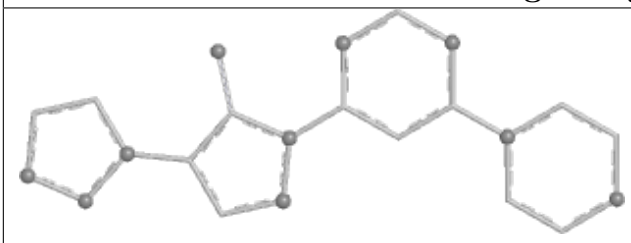
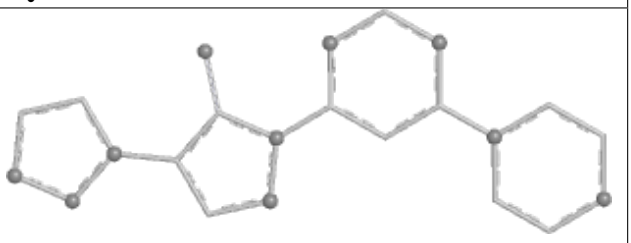
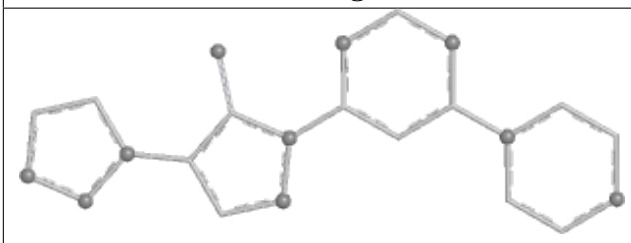
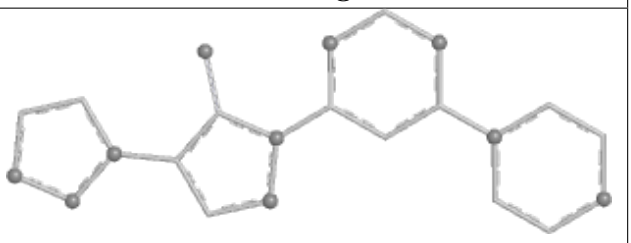
Torsions

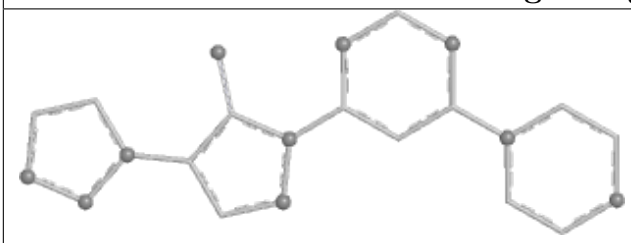
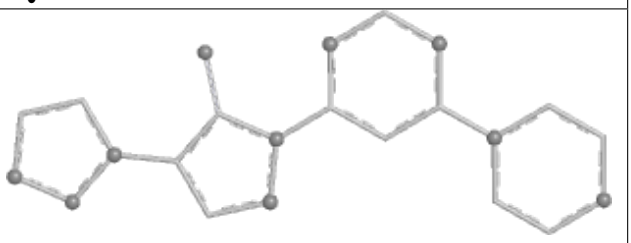
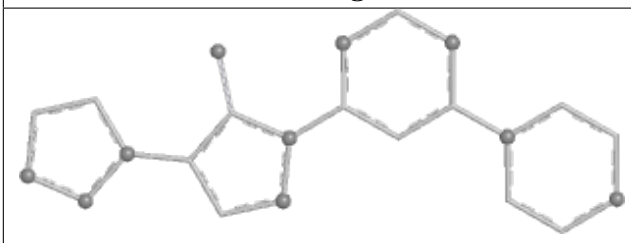
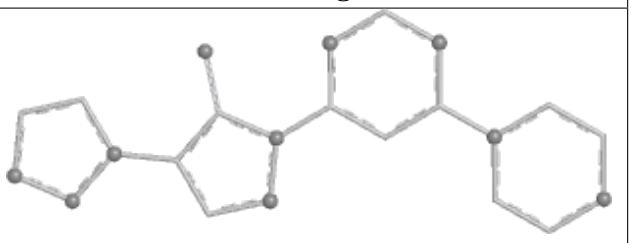


Rings

Ligand QEQ C 602	
	
Bond lengths	**Bond angles**
	
Torsions	**Rings**

Ligand QEQ D 602	
	
Bond lengths	**Bond angles**
	
Torsions	**Rings**

Ligand QEQ E 602	
 Bond lengths	 Bond angles
 Torsions	 Rings

Ligand QEQ F 602	
 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, 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(Å^2)	Q<0.9
1	A	204/227 (89%)	0.71	11 (5%)	25	20	29, 43, 77, 120	0
1	B	205/227 (90%)	0.57	10 (4%)	29	24	32, 44, 69, 97	0
1	C	202/227 (88%)	0.55	11 (5%)	25	20	34, 52, 72, 94	0
1	D	202/227 (88%)	0.67	13 (6%)	19	15	35, 52, 78, 105	0
1	E	197/227 (86%)	0.86	22 (11%)	5	4	41, 61, 83, 103	0
1	F	196/227 (86%)	0.85	28 (14%)	2	1	33, 57, 84, 111	0
All	All	1206/1362 (88%)	0.70	95 (7%)	12	9	29, 51, 81, 120	0

The worst 5 of 95 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	188	LEU	10.1
1	A	236	THR	8.8
1	D	233	GLY	8.2
1	E	405	THR	6.5
1	A	233	GLY	5.9

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 (i)

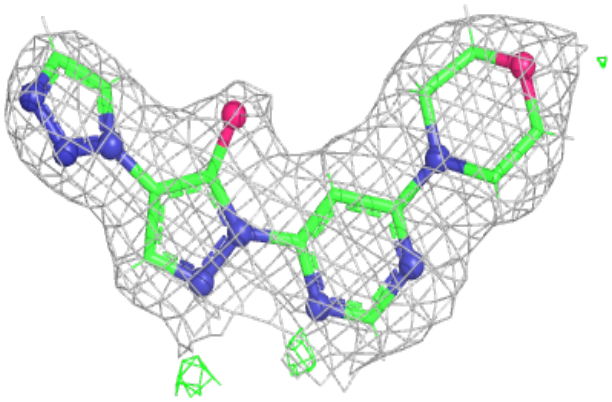
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

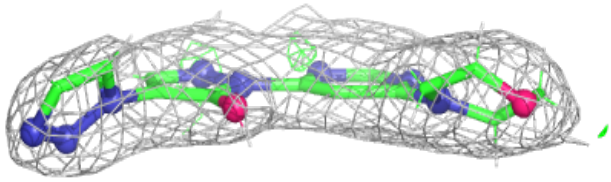
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	CL	D	603	1/1	0.61	0.16	69,69,69,69	0
4	CL	F	604	1/1	0.66	0.10	70,70,70,70	0
4	CL	E	604	1/1	0.68	0.10	69,69,69,69	0
4	CL	A	605	1/1	0.81	0.05	70,70,70,70	0
4	CL	A	604	1/1	0.82	0.07	70,70,70,70	0
4	CL	B	604	1/1	0.83	0.08	67,67,67,67	0
4	CL	C	603	1/1	0.88	0.22	70,70,70,70	0
4	CL	B	603	1/1	0.90	0.12	50,50,50,50	0
4	CL	D	604	1/1	0.90	0.06	77,77,77,77	0
4	CL	E	603	1/1	0.92	0.13	58,58,58,58	0
4	CL	F	603	1/1	0.93	0.20	51,51,51,51	0
3	QEQ	E	602	23/23	0.94	0.12	41,50,63,65	0
4	CL	F	605	1/1	0.94	0.04	76,76,76,76	0
3	QEQ	A	602	23/23	0.95	0.12	32,38,54,57	0
3	QEQ	F	602	23/23	0.95	0.14	38,46,64,66	0
4	CL	A	603	1/1	0.95	0.15	50,50,50,50	0
3	QEQ	D	602	23/23	0.96	0.10	33,44,54,57	0
3	QEQ	B	602	23/23	0.97	0.10	28,38,51,52	0
3	QEQ	C	602	23/23	0.97	0.13	34,41,55,56	0
2	MN	E	601	1/1	0.98	0.18	39,39,39,39	0
2	MN	D	601	1/1	0.99	0.17	33,33,33,33	0
2	MN	B	601	1/1	0.99	0.19	30,30,30,30	0
2	MN	C	601	1/1	0.99	0.19	33,33,33,33	0
2	MN	F	601	1/1	1.00	0.20	35,35,35,35	0
2	MN	A	601	1/1	1.00	0.18	28,28,28,28	0

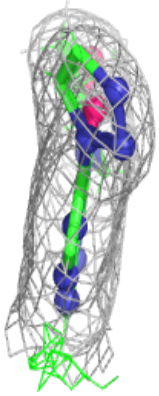
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around QEQ E 602:

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

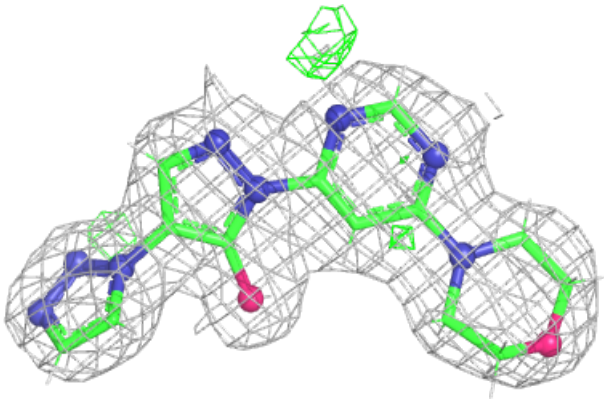


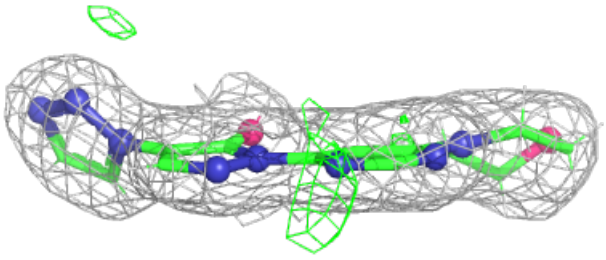


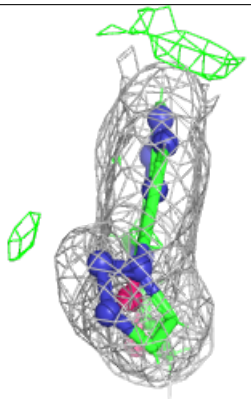


Electron density around QEQ A 602:

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

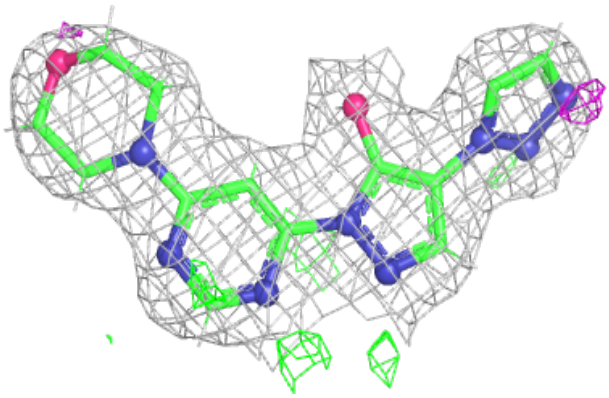


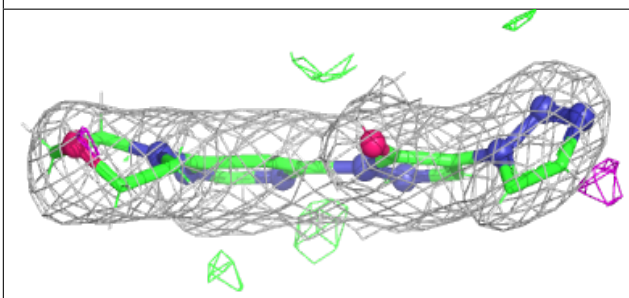


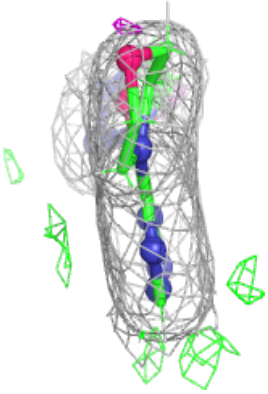


Electron density around QEQ F 602:

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

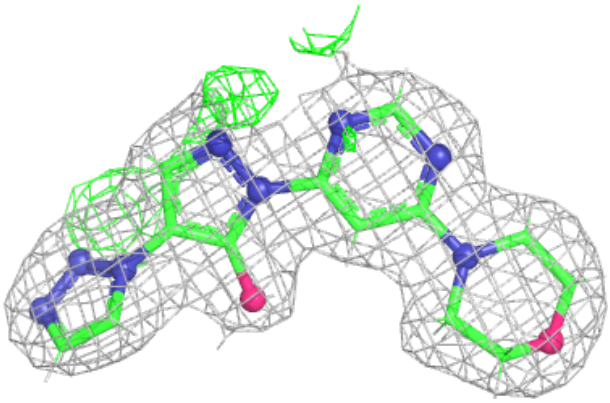


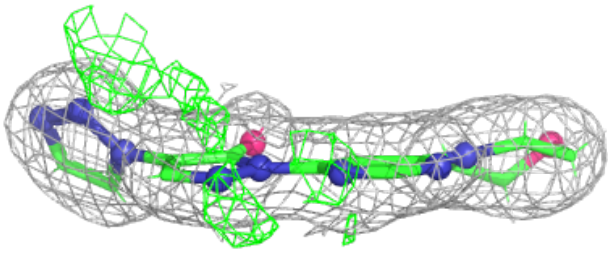


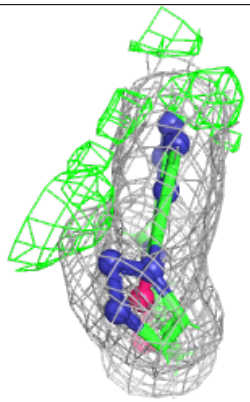


Electron density around QEQ D 602:

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

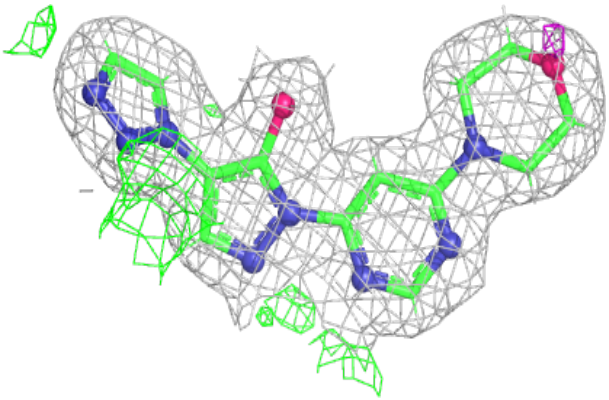


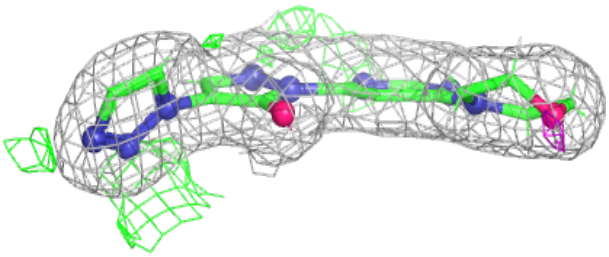


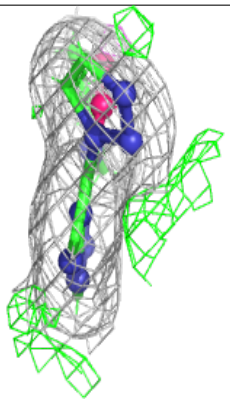


Electron density around QEQ B 602:

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

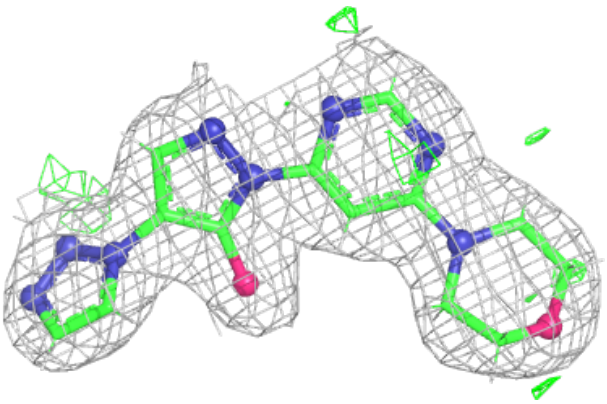


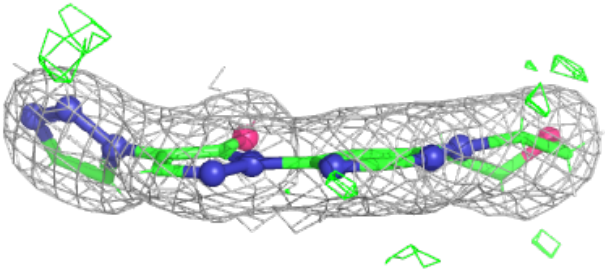


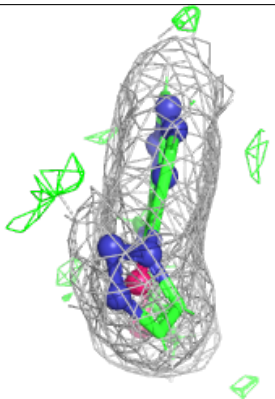


Electron density around QEQ C 602:

$2mF_o$-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 (i)

There are no such residues in this entry.