



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 30, 2026 – 03:54 AM UTC

PDB ID : 6ZH0 / pdb_00006zh0
Title : Structure of human galactokinase 1 bound with 2-(4-chlorophenyl)-N-(pyrimidin-2-yl)acetamide
Authors : Mackinnon, S.R.; Bezerra, G.A.; Zhang, M.; Foster, W.; Krojer, T.; Brandao-Neto, J.; Douangamath, A.; Arrowsmith, C.; Edwards, A.; Bountra, C.; Brennan, P.; Lai, K.; Yue, W.W.
Deposited on : 2020-06-20
Resolution : 2.50 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 ⓘ 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	:	FAILED
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Buster-report	:	wwPDB partial adaption of 1.1.7 (2018)
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

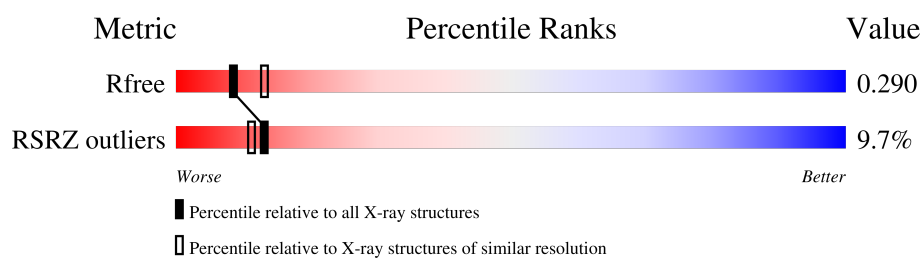
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.50 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}	180053	5829 (2.50-2.50)
RSRZ outliers	180081	5833 (2.50-2.50)

MolProbity failed to run properly - the sequence quality summary graphics cannot be shown.

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	WNP	A	403	-	-	-	X

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 11258 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 Galactokinase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	390	Total	C	N	O	S	0	0	0
			2704	1697	480	512	15			
1	B	390	Total	C	N	O	S	0	0	0
			2750	1722	483	530	15			
1	E	391	Total	C	N	O	S	0	1	0
			2821	1766	499	540	16			
1	D	371	Total	C	N	O	S	0	0	0
			2582	1617	457	495	13			

There are 40 discrepancies between the modelled and reference sequences:

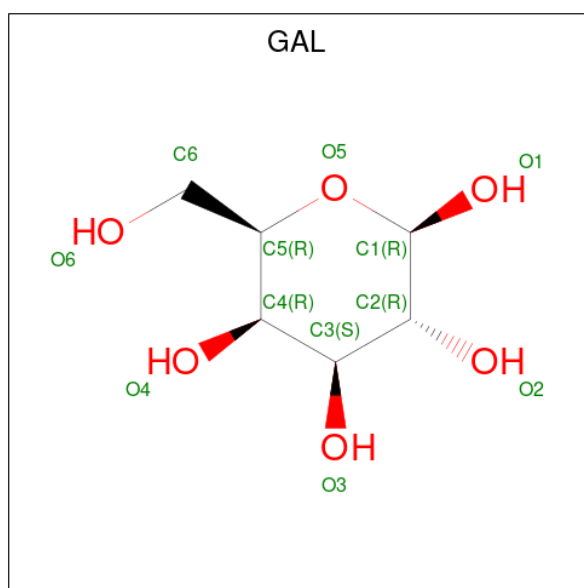
Chain	Residue	Modelled	Actual	Comment	Reference
A	-6	MET	-	initiating methionine	UNP P51570
A	-5	ALA	-	expression tag	UNP P51570
A	-4	HIS	-	expression tag	UNP P51570
A	-3	HIS	-	expression tag	UNP P51570
A	-2	HIS	-	expression tag	UNP P51570
A	-1	HIS	-	expression tag	UNP P51570
A	0	HIS	-	expression tag	UNP P51570
A	1	HIS	-	expression tag	UNP P51570
A	252	ALA	LYS	engineered mutation	UNP P51570
A	253	ALA	GLU	engineered mutation	UNP P51570
B	-6	MET	-	initiating methionine	UNP P51570
B	-5	ALA	-	expression tag	UNP P51570
B	-4	HIS	-	expression tag	UNP P51570
B	-3	HIS	-	expression tag	UNP P51570
B	-2	HIS	-	expression tag	UNP P51570
B	-1	HIS	-	expression tag	UNP P51570
B	0	HIS	-	expression tag	UNP P51570
B	1	HIS	-	expression tag	UNP P51570
B	252	ALA	LYS	engineered mutation	UNP P51570
B	253	ALA	GLU	engineered mutation	UNP P51570
E	-6	MET	-	initiating methionine	UNP P51570

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	-5	ALA	-	expression tag	UNP P51570
E	-4	HIS	-	expression tag	UNP P51570
E	-3	HIS	-	expression tag	UNP P51570
E	-2	HIS	-	expression tag	UNP P51570
E	-1	HIS	-	expression tag	UNP P51570
E	0	HIS	-	expression tag	UNP P51570
E	1	HIS	-	expression tag	UNP P51570
E	252	ALA	LYS	engineered mutation	UNP P51570
E	253	ALA	GLU	engineered mutation	UNP P51570
D	-6	MET	-	initiating methionine	UNP P51570
D	-5	ALA	-	expression tag	UNP P51570
D	-4	HIS	-	expression tag	UNP P51570
D	-3	HIS	-	expression tag	UNP P51570
D	-2	HIS	-	expression tag	UNP P51570
D	-1	HIS	-	expression tag	UNP P51570
D	0	HIS	-	expression tag	UNP P51570
D	1	HIS	-	expression tag	UNP P51570
D	252	ALA	LYS	engineered mutation	UNP P51570
D	253	ALA	GLU	engineered mutation	UNP P51570

- Molecule 2 is beta-D-galactopyranose (CCD ID: GAL) (formula: C₆H₁₂O₆).



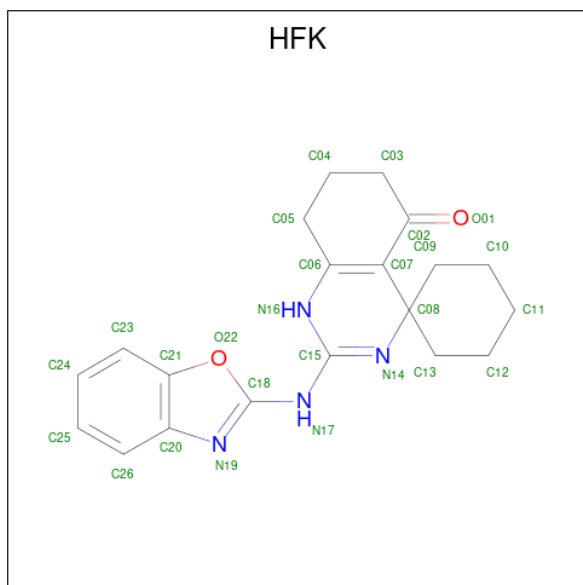
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			12	6	6		
2	B	1	Total	C	O	0	0
			12	6	6		

Continued on next page...

Continued from previous page...

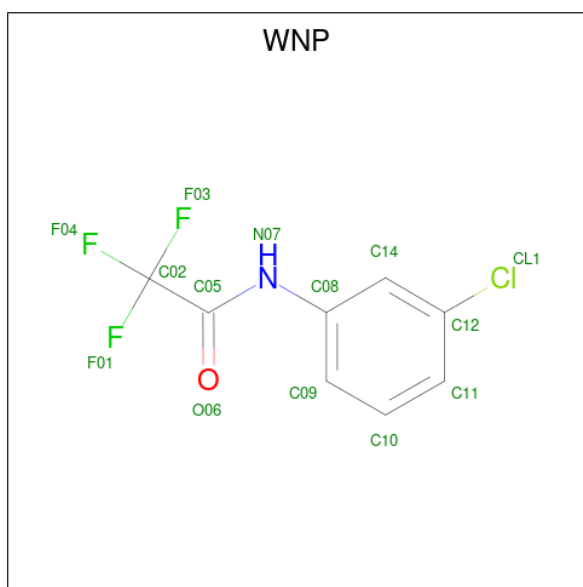
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	E	1	Total	C	O	0	0
			12	6	6		

- Molecule 3 is 2-(1,3-benzoxazol-2-ylamino)spiro[1,6,7,8-tetrahydroquinazoline-4,1'-cyclohexane]-5-one (CCD ID: HFK) (formula: $C_{20}H_{22}N_4O_2$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			26	20	4	2		
3	B	1	Total	C	N	O	0	0
			26	20	4	2		
3	E	1	Total	C	N	O	1	0
			26	20	4	2		

- Molecule 4 is N-(3-chlorophenyl)-2,2,2-trifluoroacetamide (CCD ID: WNP) (formula: $C_8H_5ClF_3NO$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
4	A	1	Total 14	C 8	Cl 1	F 3	N 1	O 1	0	0
4	B	1	Total 14	C 8	Cl 1	F 3	N 1	O 1	0	0
4	E	1	Total 14	C 8	Cl 1	F 3	N 1	O 1	0	0
4	D	1	Total 14	C 8	Cl 1	F 3	N 1	O 1	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	71	Total	O	0	0
			71	71		
5	B	47	Total	O	0	0
			47	47		
5	E	63	Total	O	0	0
			63	63		
5	D	50	Total	O	0	0
			50	50		

MolProbity failed to run properly - this section is therefore empty.

3 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	73.53Å 115.00Å 120.85Å 90.00° 100.65° 90.00°	Depositor
Resolution (Å)	118.77 – 2.50 118.77 – 2.50	Depositor EDS
% Data completeness (in resolution range)	99.7 (118.77-2.50) 99.7 (118.77-2.50)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.16 (at 2.27Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.239 , 0.290 0.241 , 0.290	Depositor DCC
R_{free} test set	4466 reflections (4.90%)	wwPDB-VP
Wilson B-factor (Å ²)	41.0	Xtriage
Anisotropy	0.581	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 55.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	11258	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 4.69% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

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.

4 Model quality [i](#)

4.1 Standard geometry [i](#)

MolProbity failed to run properly - this section is therefore empty.

4.2 Too-close contacts [i](#)

MolProbity failed to run properly - this section is therefore empty.

4.3 Torsion angles [i](#)

4.3.1 Protein backbone [i](#)

MolProbity failed to run properly - this section is therefore empty.

4.3.2 Protein sidechains [i](#)

MolProbity failed to run properly - this section is therefore empty.

4.3.3 RNA [i](#)

MolProbity failed to run properly - this section is therefore empty.

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

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

4.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

4.6 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$
3	HFK	A	402	-	29,30,30	3.30	13 (44%)	32,44,44	6.08	15 (46%)
2	GAL	A	401	-	12,12,12	0.53	0	17,17,17	1.75	4 (23%)
4	WNP	A	403	-	14,14,14	0.34	0	20,20,20	0.57	0
3	HFK	E	402	-	29,30,30	1047.39	13 (44%)	32,44,44	10.66	9 (28%)
4	WNP	B	403	-	14,14,14	0.46	0	20,20,20	0.81	1 (5%)
2	GAL	E	401	-	12,12,12	1.15	1 (8%)	17,17,17	2.59	6 (35%)
4	WNP	D	401	-	14,14,14	0.39	0	20,20,20	0.67	1 (5%)
4	WNP	E	403	-	14,14,14	0.37	0	20,20,20	0.51	0
3	HFK	B	402	-	29,30,30	2.90	13 (44%)	32,44,44	2.15	9 (28%)
2	GAL	B	401	-	12,12,12	0.75	0	17,17,17	1.92	6 (35%)

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	HFK	A	402	-	-	1/4/43/43	0/5/5/5
2	GAL	A	401	-	-	2/2/22/22	0/1/1/1
4	WNP	A	403	-	-	0/10/10/10	0/1/1/1
3	HFK	E	402	-	-	2/4/43/43	0/5/5/5
4	WNP	B	403	-	-	0/10/10/10	0/1/1/1
2	GAL	E	401	-	-	2/2/22/22	0/1/1/1
4	WNP	D	401	-	-	2/10/10/10	0/1/1/1
4	WNP	E	403	-	-	0/10/10/10	0/1/1/1
3	HFK	B	402	-	-	2/4/43/43	0/5/5/5
2	GAL	B	401	-	-	1/2/22/22	0/1/1/1

The worst 5 of 40 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	E	402	HFK	O01-C02	5640.35	115.51	1.23
3	A	402	HFK	C06-C07	12.44	1.53	1.37
3	E	402	HFK	C06-C07	10.98	1.51	1.37
3	B	402	HFK	C06-C07	9.87	1.49	1.37
3	B	402	HFK	C15-N14	6.14	1.46	1.30

The worst 5 of 51 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	E	402	HFK	O01-C02-C03	-43.62	49.38	120.87
3	E	402	HFK	O01-C02-C07	-40.26	67.57	121.53
3	A	402	HFK	N17-C18-N19	-21.62	67.25	128.58
3	A	402	HFK	O22-C18-N19	-14.20	90.05	115.57
3	A	402	HFK	C20-N19-C18	10.83	125.61	104.13

There are no chirality outliers.

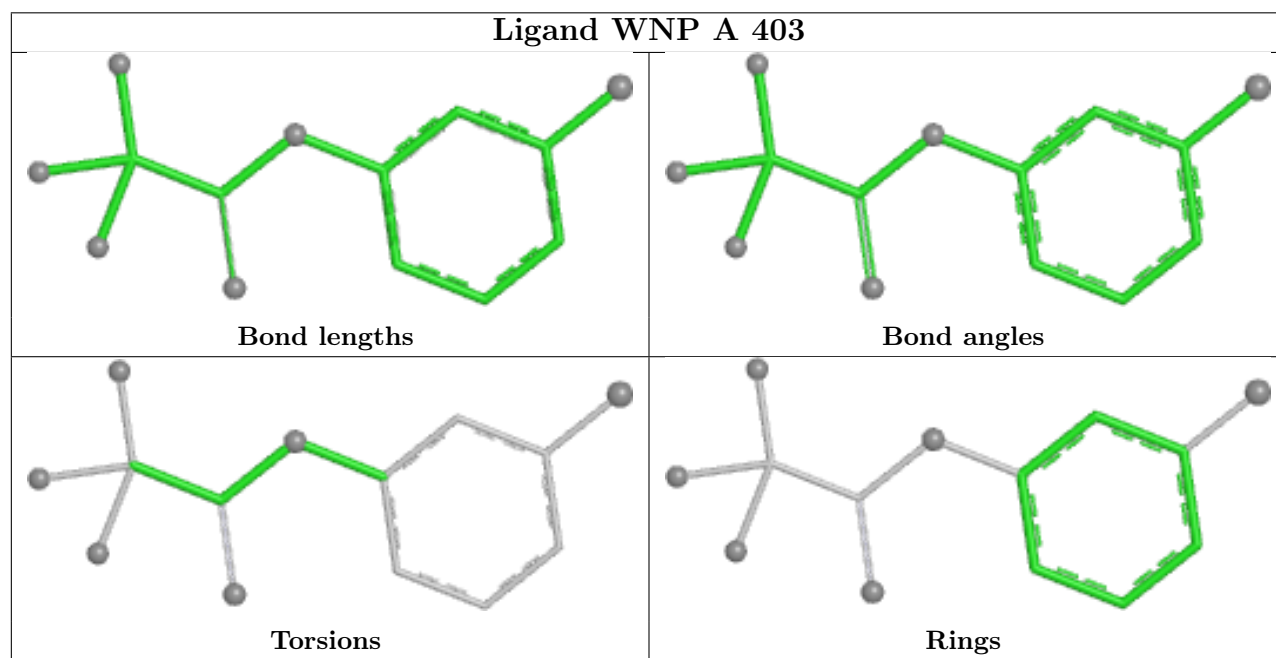
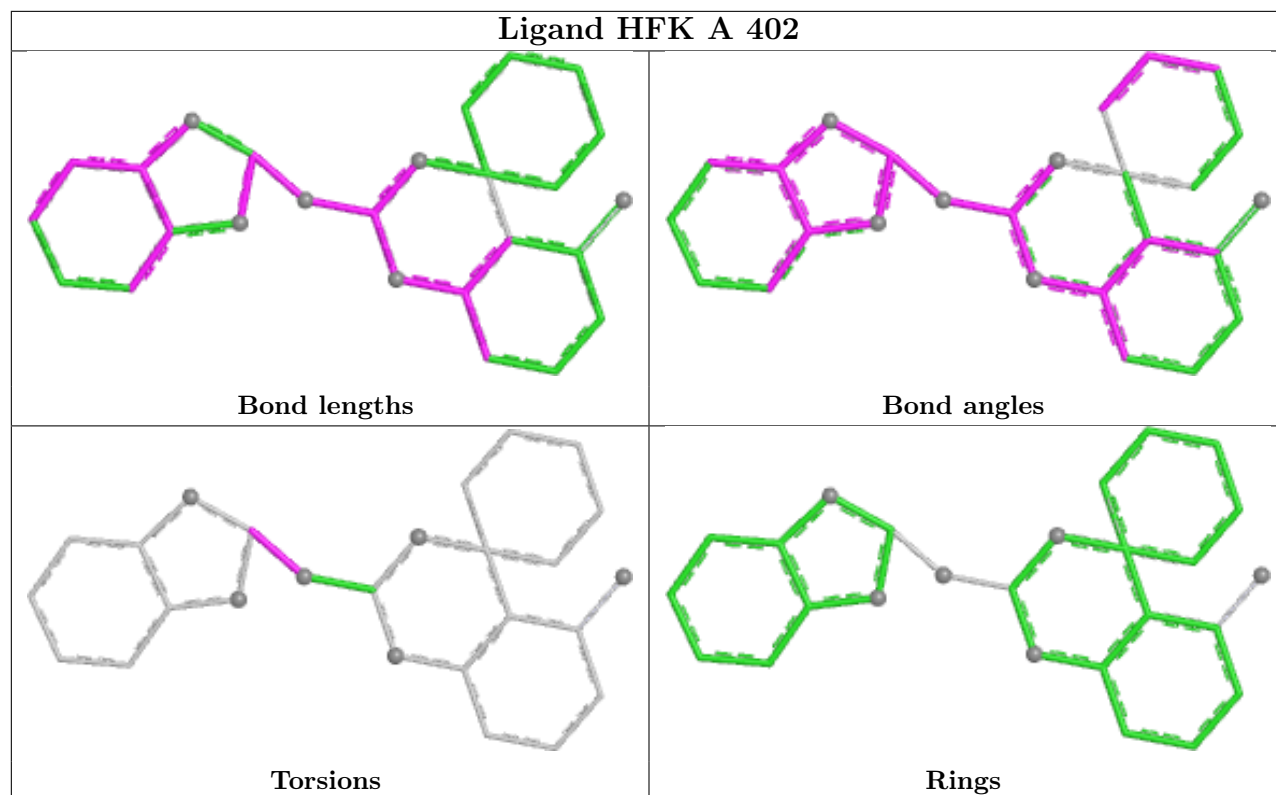
5 of 12 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	401	GAL	O5-C5-C6-O6
2	E	401	GAL	O5-C5-C6-O6
2	E	401	GAL	C4-C5-C6-O6
2	A	401	GAL	C4-C5-C6-O6
2	B	401	GAL	O5-C5-C6-O6

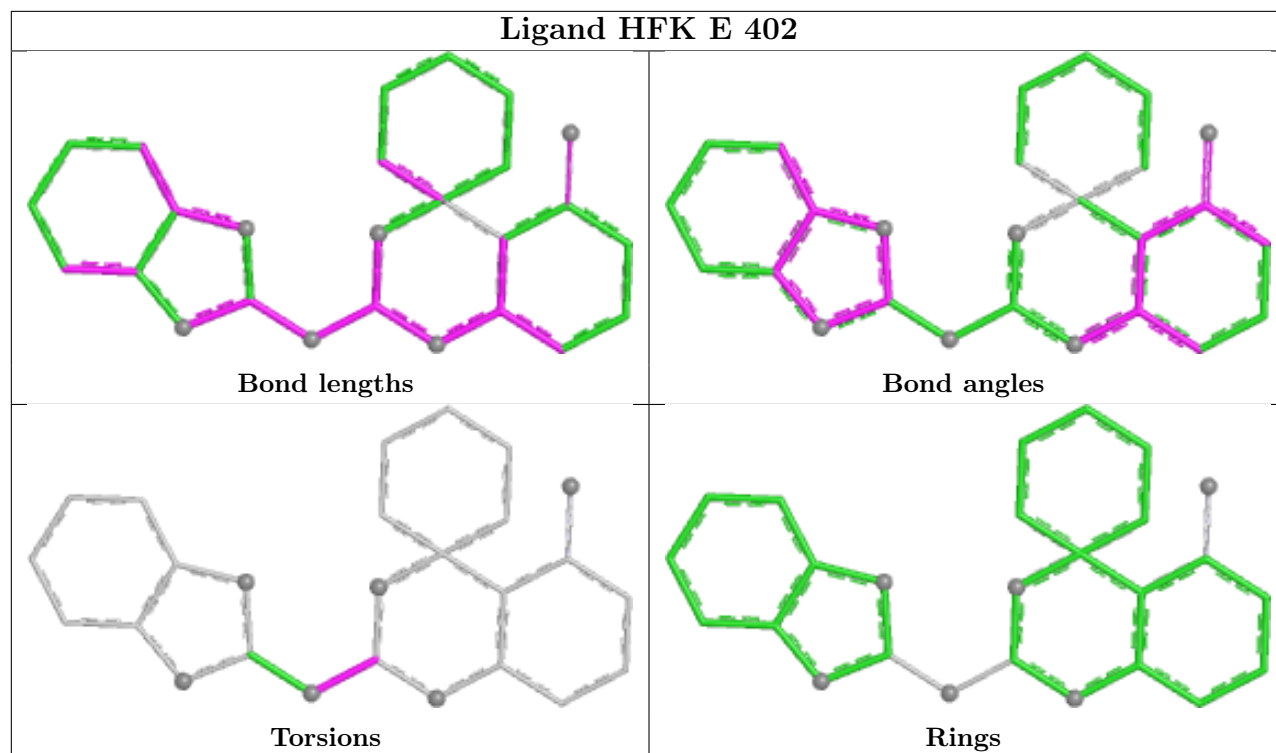
There are no ring outliers.

No monomer is involved in short contacts.

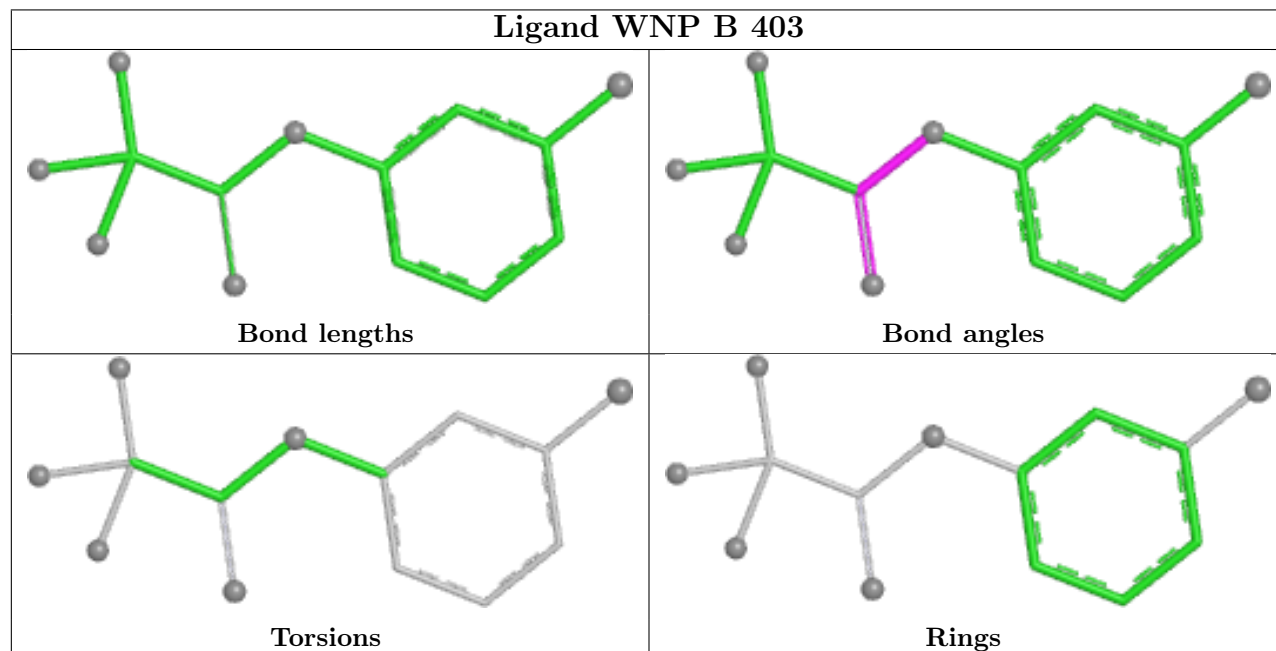
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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

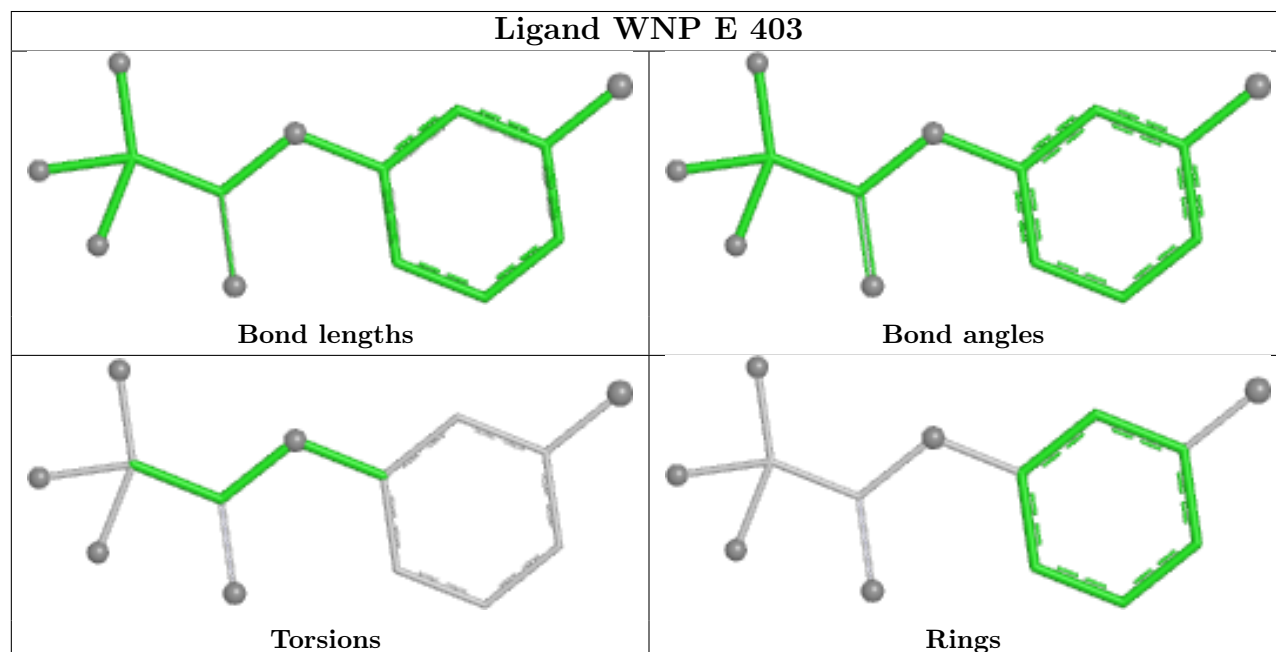
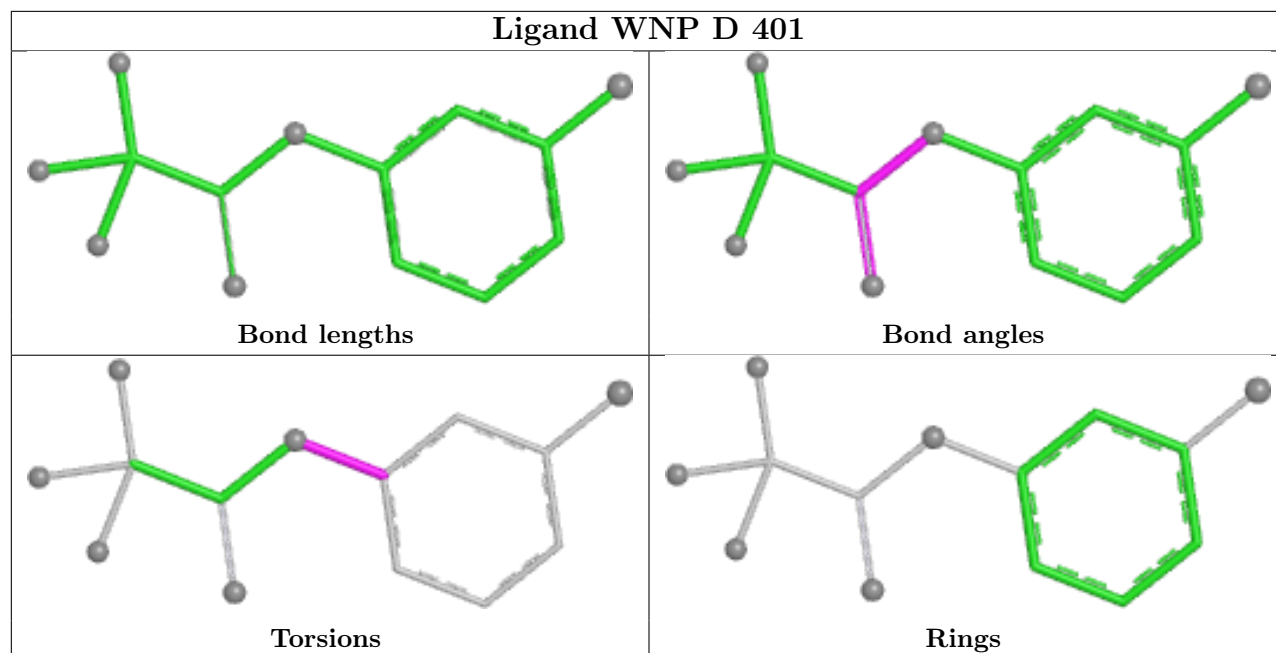


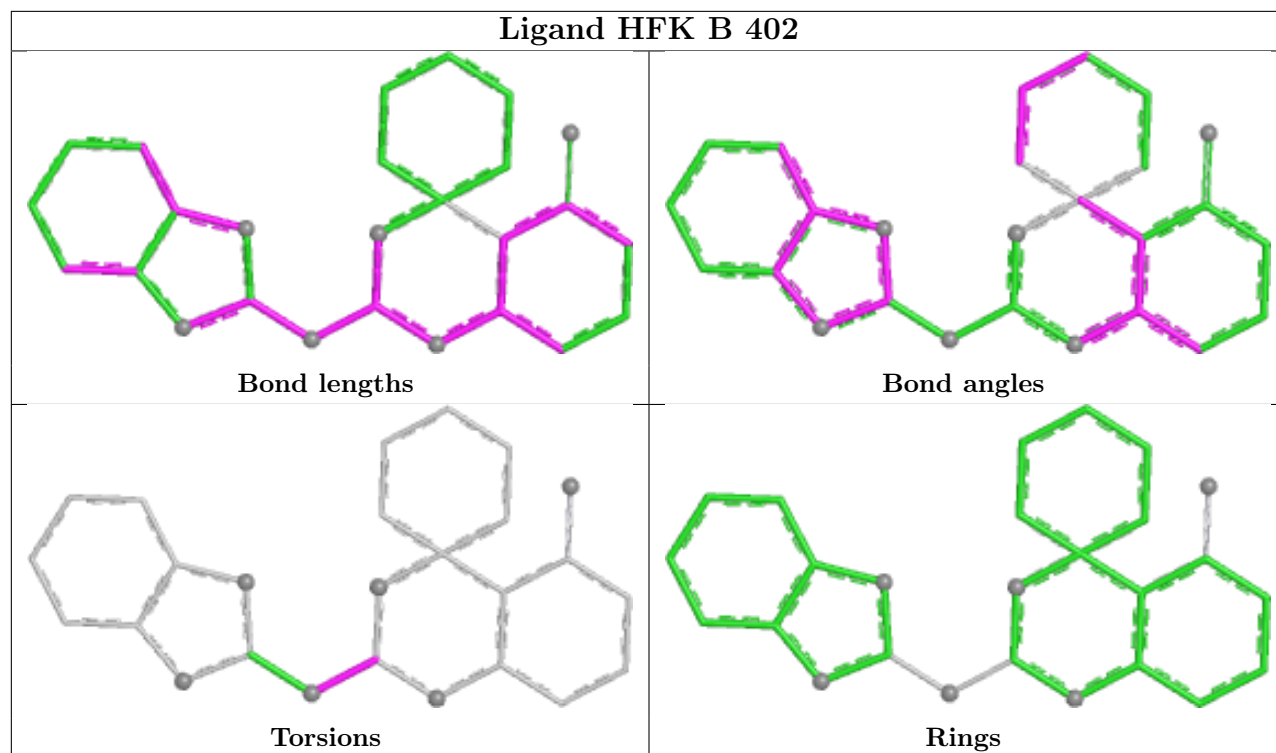
Ligand HFK E 402



Ligand WNP B 403







4.7 Other polymers [i](#)

There are no such residues in this entry.

4.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

5 Fit of model and data [i](#)

5.1 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, 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	390/399 (97%)	0.86	51 (13%) 7 6	39, 61, 99, 115	0
1	B	390/399 (97%)	0.65	30 (7%) 19 17	38, 55, 86, 102	0
1	D	371/399 (92%)	1.06	64 (17%) 4 3	36, 60, 109, 141	0
1	E	391/399 (97%)	0.35	4 (1%) 79 76	23, 45, 64, 73	1 (0%)
All	All	1542/1596 (96%)	0.73	149 (9%) 13 11	23, 54, 95, 141	1 (0%)

The worst 5 of 149 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	25	GLY	5.0
1	A	233	SER	4.9
1	D	373	GLY	4.9
1	D	7	PRO	4.5
1	D	253	ALA	4.5

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

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

5.3 Carbohydrates [i](#)

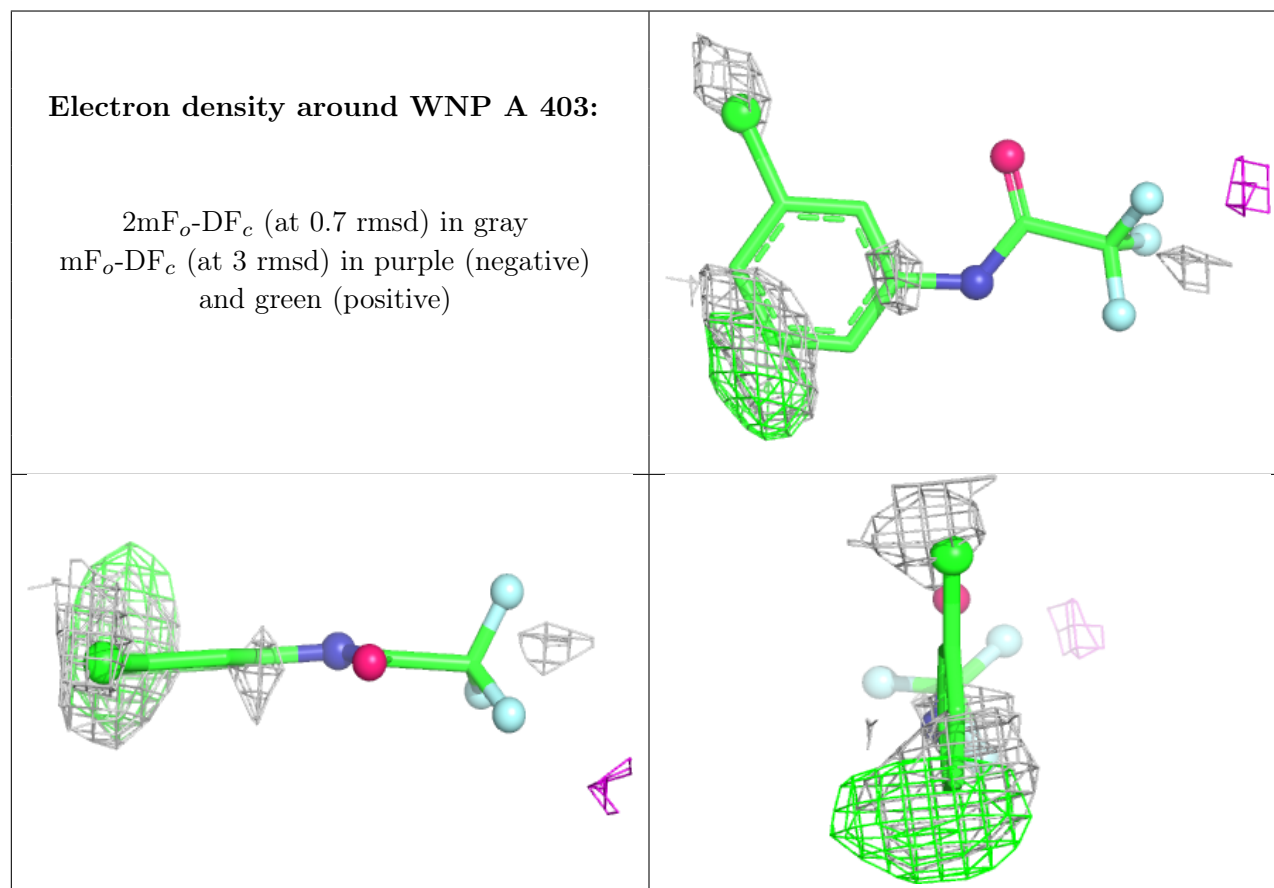
There are no oligosaccharides in this entry.

5.4 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, 95th percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

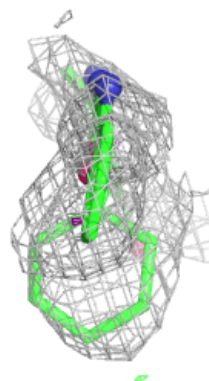
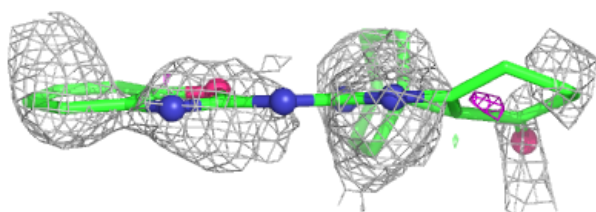
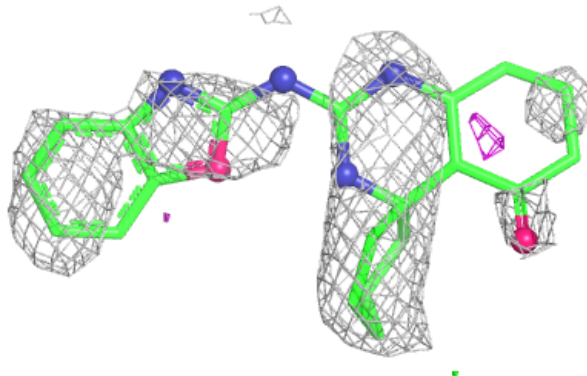
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($\text{\AA}^2$)	Q<0.9
4	WNP	A	403	14/14	0.54	0.49	70,76,82,83	14
3	HFK	B	402	26/26	0.77	0.21	84,87,91,92	0
4	WNP	E	403	14/14	0.80	0.33	49,51,59,61	14
3	HFK	A	402	26/26	0.83	0.19	70,77,80,81	0
4	WNP	D	401	14/14	0.83	0.29	38,40,43,43	14
3	HFK	E	402	26/26	0.84	0.15	20,54,58,58	1
4	WNP	B	403	14/14	0.85	0.27	51,53,55,55	14
2	GAL	B	401	12/12	0.88	0.10	44,47,48,48	0
2	GAL	A	401	12/12	0.91	0.09	48,50,54,55	0
2	GAL	E	401	12/12	0.93	0.08	35,36,37,37	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.

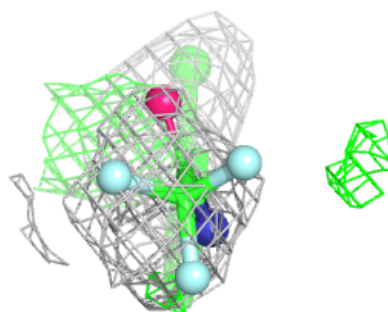
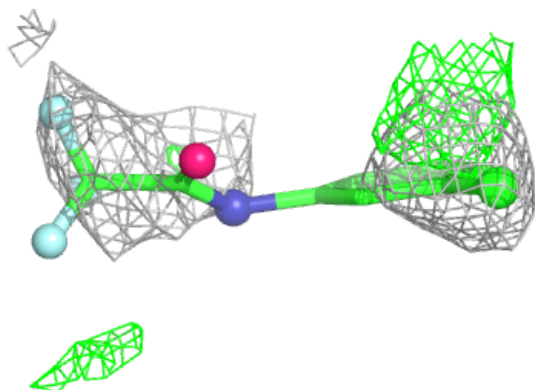
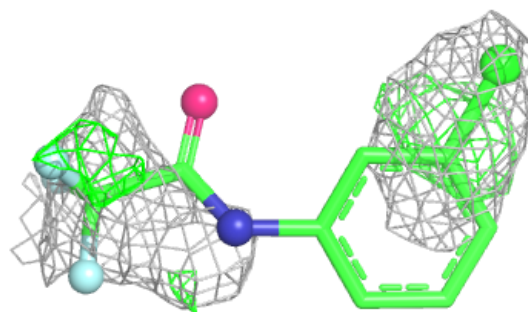


Electron density around HFK B 402:

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

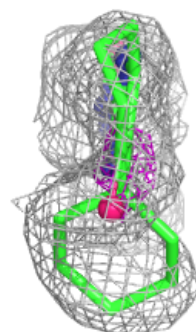
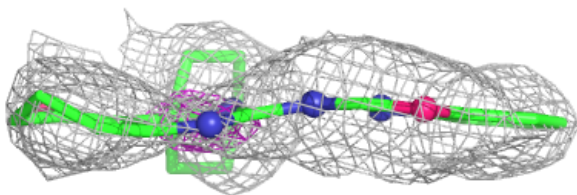
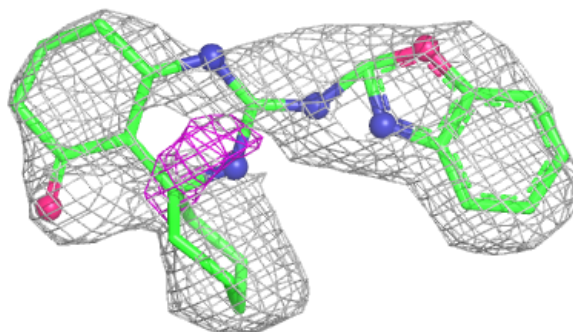
**Electron density around WNP E 403:**

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



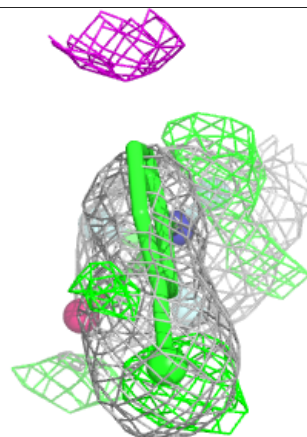
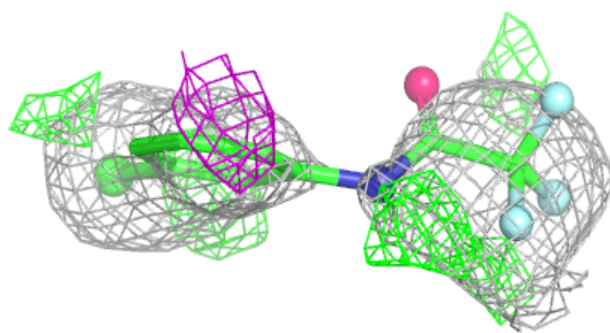
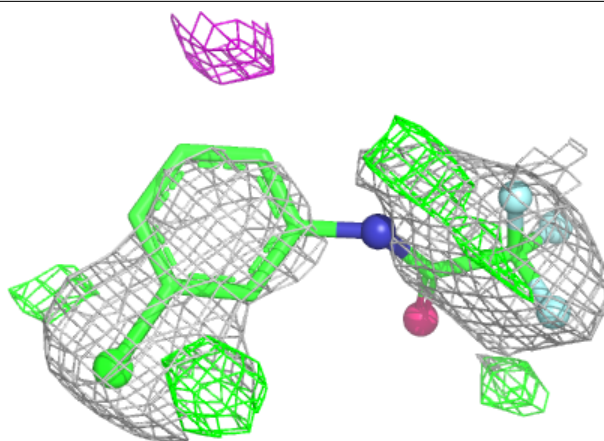
Electron density around HFK A 402:

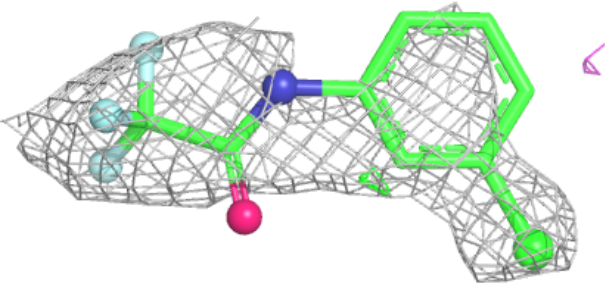
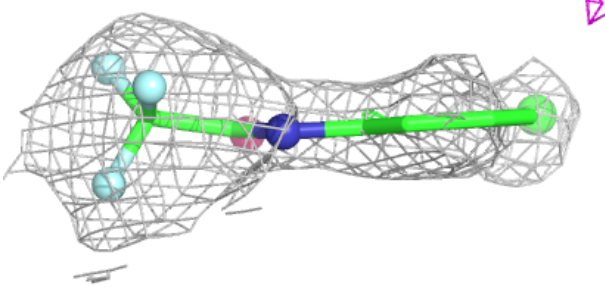
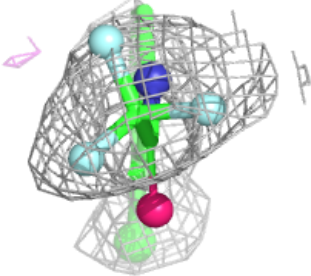
$2mF_o - DF_c$ (at 0.7 rmsd) in gray
 $mF_o - DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)



Electron density around WNP D 401:

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



Electron density around HFK E 402: $2mF_o-DF_c$ (at 0.7 rmsd) in gray mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)	 A small, localized electron density map showing a single peak of positive density (green) and a small negative density (purple) on a white background.
 A small, localized electron density map showing a single peak of positive density (green) and a small negative density (purple) on a white background.	 A small, localized electron density map showing a single peak of positive density (green) and a small negative density (purple) on a white background.
Electron density around WNP B 403: $2mF_o-DF_c$ (at 0.7 rmsd) in gray mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)	 A 3D molecular model of a ligand bound to a protein site. The ligand is shown in green sticks, with blue spheres for nitrogen and red for oxygen. It is surrounded by a gray mesh representing the $2mF_o-DF_c$ electron density at 0.7 rmsd. A purple arrow points to a specific feature on the right side of the molecule.
 A 3D molecular model of a ligand bound to a protein site, viewed from a different angle than the top-right panel. The ligand is shown in green sticks, with blue spheres for nitrogen and red for oxygen. It is surrounded by a gray mesh representing the $2mF_o-DF_c$ electron density at 0.7 rmsd. A purple arrow points to a specific feature on the right side of the molecule.	 A 3D molecular model of a ligand bound to a protein site, viewed from a different angle than the top-right panel. The ligand is shown in green sticks, with blue spheres for nitrogen and red for oxygen. It is surrounded by a gray mesh representing the $2mF_o-DF_c$ electron density at 0.7 rmsd. A purple arrow points to a specific feature on the left side of the molecule.

5.5 Other polymers [i](#)

There are no such residues in this entry.