Full wwPDB X-ray Structure Validation Report (i)

Apr 28, 2024 – 02:17 am BST

PDB ID	:	4D2W
Title	:	Structure of MELK in complex with inhibitors
Authors	:	Johnson, C.N.; Berdini, V.; Beke, L.; Bonnet, P.; Brehmer, D.; Coyle, J.E.; Day, P.J.; Frederickson, M.; Freyne, E.J.E.; Gilissen, R.A.H.J.; Hamlett, C.C.F.; Howard, S.; Meerpoel, L.; McMenamin, R.; Patel, S.; Rees, D.C.; Sharff, A.; Sommen, F.; Wu, T.; Linders, J.T.M.
Deposited on	:	2014-05-13
Resolution	:	1.92 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 (i)) were used in the production of this report:

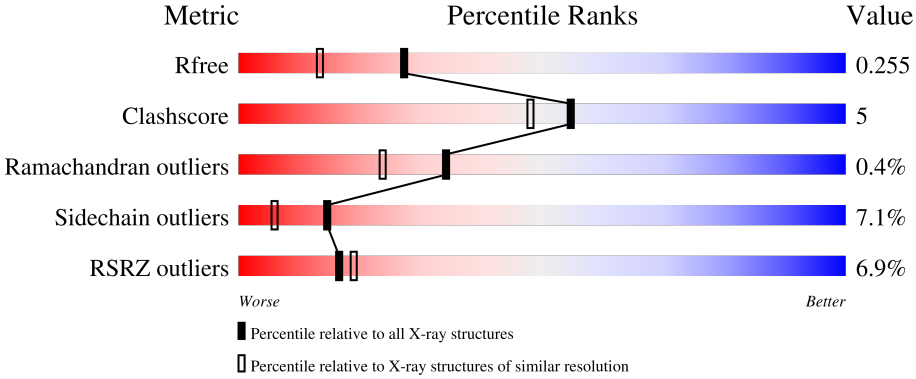
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36.2
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
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 1.92 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	7937 (1.94-1.90)
Clashscore	141614	8644 (1.94-1.90)
Ramachandran outliers	138981	8530 (1.94-1.90)
Sidechain outliers	138945	8530 (1.94-1.90)
RSRZ outliers	127900	7793 (1.94-1.90)

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	356	5% 75% 12% • 12%
1	B	356	5% 76% 12% • 9%
1	C	356	8% 72% 14% • 12%
1	D	356	6% 80% 9% • 8%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 11529 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 MATERNAL EMBRYONIC LEUCINE ZIPPER KINASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	314	2547	1641	429	459	18	0	1	0
1	B	323	2618	1688	442	471	17	0	2	0
1	C	313	2542	1640	429	456	17	0	0	0
1	D	326	2661	1716	452	475	18	0	3	0

There are 108 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	expression tag	UNP Q14680
A	-18	GLY	-	expression tag	UNP Q14680
A	-17	SER	-	expression tag	UNP Q14680
A	-16	SER	-	expression tag	UNP Q14680
A	-15	HIS	-	expression tag	UNP Q14680
A	-14	HIS	-	expression tag	UNP Q14680
A	-13	HIS	-	expression tag	UNP Q14680
A	-12	HIS	-	expression tag	UNP Q14680
A	-11	HIS	-	expression tag	UNP Q14680
A	-10	HIS	-	expression tag	UNP Q14680
A	-9	SER	-	expression tag	UNP Q14680
A	-8	SER	-	expression tag	UNP Q14680
A	-7	GLY	-	expression tag	UNP Q14680
A	-6	LEU	-	expression tag	UNP Q14680
A	-5	VAL	-	expression tag	UNP Q14680
A	-4	PRO	-	expression tag	UNP Q14680
A	-3	ARG	-	expression tag	UNP Q14680
A	-2	GLY	-	expression tag	UNP Q14680
A	-1	SER	-	expression tag	UNP Q14680
A	0	HIS	-	expression tag	UNP Q14680
A	167	ALA	THR	engineered mutation	UNP Q14680

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	171	ALA	SER	engineered mutation	UNP Q14680
A	213	THR	ASN	engineered mutation	UNP Q14680
A	214	ALA	VAL	engineered mutation	UNP Q14680
A	215	ALA	MET	engineered mutation	UNP Q14680
A	218	VAL	TYR	engineered mutation	UNP Q14680
A	219	ALA	LYS	engineered mutation	UNP Q14680
B	-19	MET	-	expression tag	UNP Q14680
B	-18	GLY	-	expression tag	UNP Q14680
B	-17	SER	-	expression tag	UNP Q14680
B	-16	SER	-	expression tag	UNP Q14680
B	-15	HIS	-	expression tag	UNP Q14680
B	-14	HIS	-	expression tag	UNP Q14680
B	-13	HIS	-	expression tag	UNP Q14680
B	-12	HIS	-	expression tag	UNP Q14680
B	-11	HIS	-	expression tag	UNP Q14680
B	-10	HIS	-	expression tag	UNP Q14680
B	-9	SER	-	expression tag	UNP Q14680
B	-8	SER	-	expression tag	UNP Q14680
B	-7	GLY	-	expression tag	UNP Q14680
B	-6	LEU	-	expression tag	UNP Q14680
B	-5	VAL	-	expression tag	UNP Q14680
B	-4	PRO	-	expression tag	UNP Q14680
B	-3	ARG	-	expression tag	UNP Q14680
B	-2	GLY	-	expression tag	UNP Q14680
B	-1	SER	-	expression tag	UNP Q14680
B	0	HIS	-	expression tag	UNP Q14680
B	167	ALA	THR	engineered mutation	UNP Q14680
B	171	ALA	SER	engineered mutation	UNP Q14680
B	213	THR	ASN	engineered mutation	UNP Q14680
B	214	ALA	VAL	engineered mutation	UNP Q14680
B	215	ALA	MET	engineered mutation	UNP Q14680
B	218	VAL	TYR	engineered mutation	UNP Q14680
B	219	ALA	LYS	engineered mutation	UNP Q14680
C	-19	MET	-	expression tag	UNP Q14680
C	-18	GLY	-	expression tag	UNP Q14680
C	-17	SER	-	expression tag	UNP Q14680
C	-16	SER	-	expression tag	UNP Q14680
C	-15	HIS	-	expression tag	UNP Q14680
C	-14	HIS	-	expression tag	UNP Q14680
C	-13	HIS	-	expression tag	UNP Q14680
C	-12	HIS	-	expression tag	UNP Q14680
C	-11	HIS	-	expression tag	UNP Q14680

Continued on next page...

Continued from previous page...

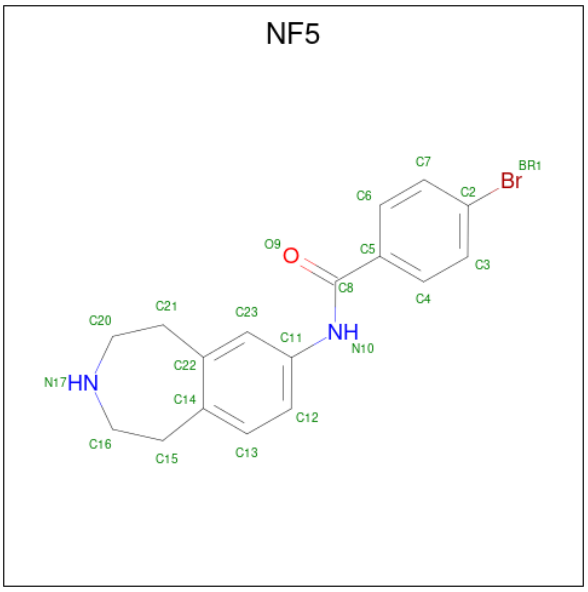
Chain	Residue	Modelled	Actual	Comment	Reference
C	-10	HIS	-	expression tag	UNP Q14680
C	-9	SER	-	expression tag	UNP Q14680
C	-8	SER	-	expression tag	UNP Q14680
C	-7	GLY	-	expression tag	UNP Q14680
C	-6	LEU	-	expression tag	UNP Q14680
C	-5	VAL	-	expression tag	UNP Q14680
C	-4	PRO	-	expression tag	UNP Q14680
C	-3	ARG	-	expression tag	UNP Q14680
C	-2	GLY	-	expression tag	UNP Q14680
C	-1	SER	-	expression tag	UNP Q14680
C	0	HIS	-	expression tag	UNP Q14680
C	167	ALA	THR	engineered mutation	UNP Q14680
C	171	ALA	SER	engineered mutation	UNP Q14680
C	213	THR	ASN	engineered mutation	UNP Q14680
C	214	ALA	VAL	engineered mutation	UNP Q14680
C	215	ALA	MET	engineered mutation	UNP Q14680
C	218	VAL	TYR	engineered mutation	UNP Q14680
C	219	ALA	LYS	engineered mutation	UNP Q14680
D	-19	MET	-	expression tag	UNP Q14680
D	-18	GLY	-	expression tag	UNP Q14680
D	-17	SER	-	expression tag	UNP Q14680
D	-16	SER	-	expression tag	UNP Q14680
D	-15	HIS	-	expression tag	UNP Q14680
D	-14	HIS	-	expression tag	UNP Q14680
D	-13	HIS	-	expression tag	UNP Q14680
D	-12	HIS	-	expression tag	UNP Q14680
D	-11	HIS	-	expression tag	UNP Q14680
D	-10	HIS	-	expression tag	UNP Q14680
D	-9	SER	-	expression tag	UNP Q14680
D	-8	SER	-	expression tag	UNP Q14680
D	-7	GLY	-	expression tag	UNP Q14680
D	-6	LEU	-	expression tag	UNP Q14680
D	-5	VAL	-	expression tag	UNP Q14680
D	-4	PRO	-	expression tag	UNP Q14680
D	-3	ARG	-	expression tag	UNP Q14680
D	-2	GLY	-	expression tag	UNP Q14680
D	-1	SER	-	expression tag	UNP Q14680
D	0	HIS	-	expression tag	UNP Q14680
D	167	ALA	THR	engineered mutation	UNP Q14680
D	171	ALA	SER	engineered mutation	UNP Q14680
D	213	THR	ASN	engineered mutation	UNP Q14680
D	214	ALA	VAL	engineered mutation	UNP Q14680

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	215	ALA	MET	engineered mutation	UNP Q14680
D	218	VAL	TYR	engineered mutation	UNP Q14680
D	219	ALA	LYS	engineered mutation	UNP Q14680

- Molecule 2 is 4-bromo-N-(2,3,4,5-tetrahydro-1H-3-benzazepin-7-yl)benzamide (three-letter code: NF5) (formula: $C_{17}H_{17}BrN_2O$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	Br	C	N	O	0	0
			21	1	17	2	1		
2	B	1	Total	Br	C	N	O	0	0
			21	1	17	2	1		
2	C	1	Total	Br	C	N	O	0	0
			21	1	17	2	1		
2	D	1	Total	Br	C	N	O	0	0
			21	1	17	2	1		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	312	Total	O	0	0
			312	312		
3	B	270	Total	O	0	0
			270	270		
3	C	237	Total	O	0	0
			237	237		

Continued on next page...

Continued from previous page...

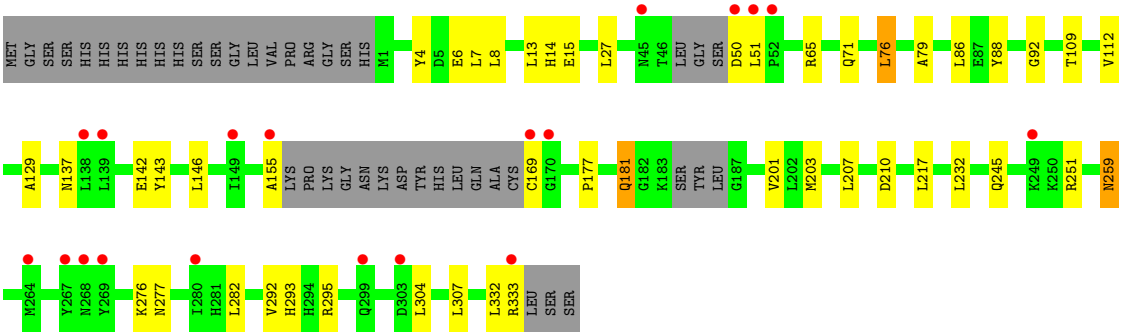
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	D	258	258	258	0	0

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.

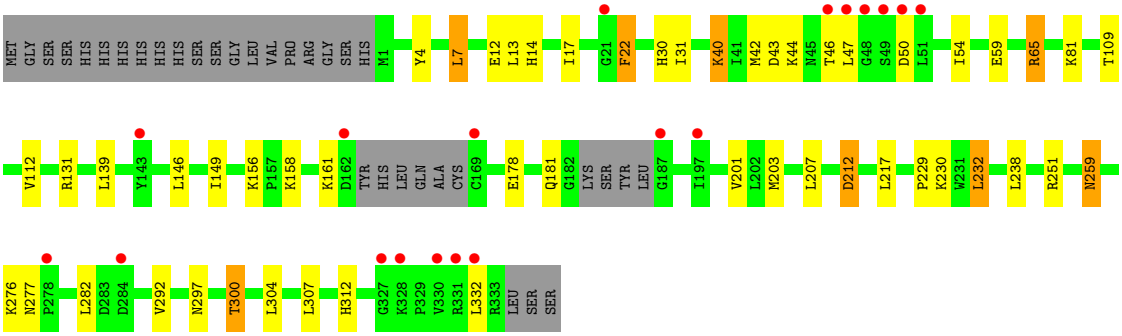
• Molecule 1: MATERNAL EMBRYONIC LEUCINE ZIPPER KINASE

Chain A:



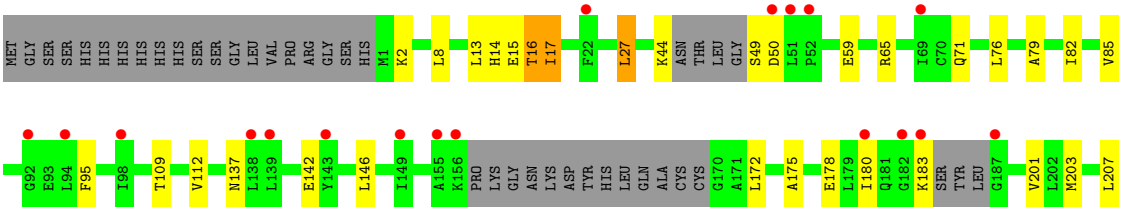
• Molecule 1: MATERNAL EMBRYONIC LEUCINE ZIPPER KINASE

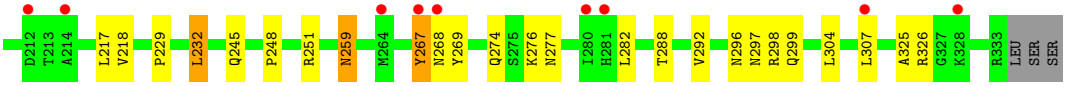
Chain B:



• Molecule 1: MATERNAL EMBRYONIC LEUCINE ZIPPER KINASE

Chain C:





- Molecule 1: MATERNAL EMBRYONIC LEUCINE ZIPPER KINASE

Chain D:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	65.96Å 74.95Å 78.22Å 85.74° 70.07° 89.94°	Depositor
Resolution (Å)	35.01 – 1.92 34.57 – 1.92	Depositor EDS
% Data completeness (in resolution range)	94.8 (35.01-1.92) 94.8 (34.57-1.92)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.44 (at 1.92Å)	Xtriage
Refinement program	BUSTER 2.11.5	Depositor
R, R_{free}	0.201 , 0.247 0.210 , 0.255	Depositor DCC
R_{free} test set	5074 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	32.4	Xtriage
Anisotropy	0.224	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 60.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	11529	wwPDB-VP
Average B, all atoms (Å^2)	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.99% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: NF5

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	0.65	0/2606	0.66	0/3523
1	B	0.62	0/2682	0.68	0/3626
1	C	0.58	0/2598	0.66	0/3509
1	D	0.68	0/2728	0.69	0/3687
All	All	0.63	0/10614	0.68	0/14345

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	169	CYS	Peptide

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2547	0	2561	25	0
1	B	2618	0	2648	27	0
1	C	2542	0	2567	28	0
1	D	2661	0	2698	27	0
2	A	21	0	17	3	0
2	B	21	0	17	2	0
2	C	21	0	17	0	0
2	D	21	0	17	4	0
3	A	312	0	0	7	0
3	B	270	0	0	3	0
3	C	237	0	0	3	0
3	D	258	0	0	1	0
All	All	11529	0	10542	109	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 5.

All (109) 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)
2:D:1334:NF5:BR1	3:D:2038:HOH:O	2.30	1.03
1:D:220:LYS:HG2	1:D:223[A]:ARG:HH21	1.40	0.85
1:B:212:ASP:O	1:C:79:ALA:HB2	1.83	0.78
1:A:92:GLY:HA2	3:A:2118:HOH:O	1.86	0.76
1:A:177:PRO:O	1:A:181:GLN:HG2	1.86	0.75
1:C:245:GLN:O	1:C:251:ARG:HD3	1.86	0.75
1:C:288:THR:HG23	1:C:298:ARG:HH22	1.50	0.75
1:B:17:ILE:O	1:C:17:ILE:HD12	1.89	0.71
1:C:16:THR:O	1:C:17:ILE:HB	1.90	0.69
1:A:6:GLU:CD	1:A:76:LEU:HD11	2.15	0.67
1:B:40:LYS:HE2	1:B:42:MET:CE	2.24	0.67
1:C:175:ALA:HB3	1:C:180:ILE:HD11	1.77	0.67
1:C:85:VAL:HG23	3:C:2059:HOH:O	1.96	0.65
1:C:248:PRO:HA	1:C:251:ARG:HG3	1.78	0.65
1:C:175:ALA:CB	1:C:180:ILE:HD11	2.26	0.65
1:D:57:GLU:HG3	1:D:61:LEU:HD23	1.79	0.65
1:A:79:ALA:HB2	1:D:212:ASP:O	1.98	0.64
2:B:1334:NF5:H23	3:B:2269:HOH:O	1.95	0.64
1:B:12:GLU:HB2	1:B:31:ILE:HD11	1.81	0.62
1:D:2:LYS:HD3	1:D:3:ASP:H	1.65	0.62
1:D:290:LEU:CD1	1:D:320:LEU:HD12	2.30	0.61
1:B:40:LYS:HE2	1:B:42:MET:HE1	1.83	0.60

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:22:PHE:HE1	1:B:42:MET:HE2	1.67	0.59
1:C:65:ARG:HG3	1:C:71:GLN:HE22	1.67	0.57
1:C:325:ALA:HB2	3:C:2226:HOH:O	2.03	0.57
1:B:201:VAL:CG1	1:B:207:LEU:HD23	2.34	0.57
1:B:7[A]:LEU:HD23	1:B:7[A]:LEU:H	1.69	0.57
1:B:4:TYR:CE2	1:C:95:PHE:HZ	2.22	0.57
1:A:142:GLU:HG2	3:A:2112:HOH:O	2.05	0.56
1:C:109:THR:HG21	1:C:203:MET:HG3	1.88	0.56
2:D:1334:NF5:O9	2:D:1334:NF5:H12	2.06	0.55
1:B:139[A]:LEU:HD21	1:B:149:ILE:HD13	1.88	0.55
1:C:245:GLN:HB2	1:C:251:ARG:HG2	1.88	0.54
1:C:76:LEU:HD12	3:C:2059:HOH:O	2.07	0.54
1:B:4:TYR:HA	1:B:7[A]:LEU:HD21	1.90	0.54
1:C:201:VAL:HG11	1:C:207:LEU:HD23	1.90	0.53
1:D:65:ARG:HG2	1:D:277:ASN:HB3	1.89	0.53
1:A:129:ALA:H	1:A:155:ALA:HB3	1.73	0.53
1:D:57:GLU:HG3	1:D:61:LEU:CD2	2.37	0.53
1:B:297:ASN:OD1	1:B:300:THR:HG23	2.09	0.53
1:C:14:HIS:HB2	1:C:27:LEU:HB3	1.90	0.53
1:A:65:ARG:HG3	1:A:71:GLN:HE22	1.75	0.52
1:C:201:VAL:CG1	1:C:207:LEU:HD23	2.39	0.52
1:B:40:LYS:HE2	1:B:42:MET:HE3	1.92	0.51
1:B:50:ASP:O	1:B:54:ILE:HG12	2.11	0.51
1:D:8:LEU:HD11	1:D:13:LEU:HD22	1.92	0.51
1:C:178:GLU:HB2	1:C:183:LYS:HB3	1.93	0.50
1:D:17[B]:ILE:HD11	1:D:27:LEU:HD11	1.94	0.50
1:A:27:LEU:HD13	1:A:88:TYR:CE1	2.47	0.49
2:A:1334:NF5:H23	3:A:2118:HOH:O	2.13	0.49
1:B:65:ARG:HD3	3:B:2057:HOH:O	2.12	0.49
3:A:2041:HOH:O	2:D:1334:NF5:H21	2.12	0.48
1:A:109:THR:HG21	1:A:203:MET:HG3	1.96	0.48
2:B:1334:NF5:O9	2:B:1334:NF5:H12	2.14	0.48
1:D:229:PRO:HD2	1:D:232:LEU:HD22	1.96	0.47
1:A:245:GLN:O	1:A:251:ARG:HD3	2.14	0.47
1:D:259:ASN:HD22	1:D:259:ASN:H	1.62	0.47
1:A:7:LEU:O	1:A:7:LEU:HD13	2.15	0.46
1:A:142:GLU:HG3	1:A:143:TYR:CE2	2.50	0.46
1:B:229:PRO:HD2	1:B:232:LEU:HD22	1.97	0.46
1:D:2:LYS:HG2	1:D:3:ASP:N	2.30	0.46
1:D:47:LEU:HD11	1:D:51:LEU:HD23	1.96	0.46
1:D:280:ILE:HG22	1:D:281:HIS:H	1.80	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:43:ASP:HB3	1:B:46:THR:HG22	1.97	0.46
1:C:259:ASN:HD22	1:C:259:ASN:H	1.63	0.46
1:A:6:GLU:OE1	1:A:76:LEU:HD11	2.16	0.46
1:B:109:THR:HG21	1:B:203:MET:HG3	1.98	0.45
1:B:201:VAL:HG11	1:B:207:LEU:HD23	1.97	0.45
1:A:27:LEU:CD1	1:A:88:TYR:CE1	3.00	0.45
1:D:2:LYS:CG	1:D:3:ASP:N	2.79	0.45
1:D:47:LEU:HD21	1:D:51:LEU:HB2	1.98	0.45
1:A:210:ASP:OD2	1:D:2:LYS:HG3	2.16	0.45
1:D:278:PRO:HG3	1:D:281:HIS:HD2	1.82	0.44
1:B:112:VAL:HG13	1:B:146:LEU:HD11	1.99	0.44
1:B:178:GLU:HA	1:B:181:GLN:HG2	1.99	0.44
1:C:267:TYR:C	1:C:269:TYR:H	2.20	0.44
3:A:2031:HOH:O	1:D:99:ILE:HD11	2.17	0.43
1:D:290:LEU:HD11	1:D:320:LEU:HD12	2.00	0.43
1:C:137:ASN:HD22	1:C:137:ASN:HA	1.64	0.43
1:B:259:ASN:H	1:B:259:ASN:HD22	1.66	0.43
1:D:2:LYS:CD	1:D:3:ASP:H	2.30	0.43
1:A:86:LEU:HD13	2:A:1334:NF5:BR1	2.74	0.43
2:A:1334:NF5:H21	3:A:2118:HOH:O	2.18	0.43
1:A:112:VAL:HG13	1:A:146:LEU:HD11	1.99	0.43
1:A:293:HIS:HB2	3:A:2272:HOH:O	2.18	0.43
1:D:172:LEU:HD21	1:D:218:VAL:HG23	2.01	0.43
1:B:131:ARG:HA	3:B:2129:HOH:O	2.18	0.42
1:D:44:LYS:HA	1:D:47:LEU:HD22	2.00	0.42
1:D:2:LYS:CG	1:D:3:ASP:H	2.31	0.42
1:A:4:TYR:O	1:A:8:LEU:HD23	2.20	0.42
1:A:7:LEU:HD12	1:A:8:LEU:HD22	2.00	0.42
1:C:229:PRO:HD2	1:C:232:LEU:HD22	2.00	0.42
1:C:112:VAL:HG13	1:C:146:LEU:HD11	2.01	0.42
1:B:44:LYS:HA	1:B:47:LEU:HD12	2.02	0.42
1:B:276:LYS:C	1:B:277:ASN:HD22	2.23	0.42
1:D:57:GLU:O	1:D:61:LEU:HD23	2.20	0.42
1:B:30:HIS:HE2	1:B:312:HIS:CE1	2.38	0.41
1:D:198:LEU:HA	1:D:201:VAL:HG12	2.02	0.41
1:C:44:LYS:HD2	1:C:82:ILE:HG13	2.02	0.41
1:C:267:TYR:CD1	1:C:267:TYR:N	2.88	0.41
1:A:201:VAL:HG11	1:A:207:LEU:HD23	2.02	0.41
1:A:259:ASN:HD22	1:A:259:ASN:H	1.68	0.41
1:A:7:LEU:HD13	1:A:7:LEU:C	2.41	0.41
1:C:276:LYS:C	1:C:277:ASN:HD22	2.24	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:88:TYR:CE1	2:D:1334:NF5:H13	2.56	0.41
1:B:139[A]:LEU:CD2	1:B:149:ILE:HD13	2.51	0.41
1:C:172:LEU:HD21	1:C:218:VAL:HG23	2.04	0.40
1:A:276:LYS:C	1:A:277:ASN:HD22	2.24	0.40
1:A:137:ASN:HD22	1:A:137:ASN:HA	1.66	0.40

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	307/356 (86%)	296 (96%)	11 (4%)	0	100	100
1	B	319/356 (90%)	303 (95%)	15 (5%)	1 (0%)	41	31
1	C	305/356 (86%)	292 (96%)	10 (3%)	3 (1%)	15	6
1	D	325/356 (91%)	312 (96%)	12 (4%)	1 (0%)	41	31
All	All	1256/1424 (88%)	1203 (96%)	48 (4%)	5 (0%)	34	24

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	22	PHE
1	C	16	THR
1	C	17	ILE
1	C	268	ASN
1	D	280	ILE

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	278/314 (88%)	261 (94%)	17 (6%)	18 8
1	B	286/314 (91%)	262 (92%)	24 (8%)	11 4
1	C	277/314 (88%)	255 (92%)	22 (8%)	12 4
1	D	291/314 (93%)	271 (93%)	20 (7%)	15 6
All	All	1132/1256 (90%)	1049 (93%)	83 (7%)	14 5

All (83) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	13	LEU
1	A	14	HIS
1	A	15	GLU
1	A	50	ASP
1	A	51	LEU
1	A	76	LEU
1	A	181	GLN
1	A	217	LEU
1	A	232	LEU
1	A	259	ASN
1	A	282	LEU
1	A	292	VAL
1	A	295	ARG
1	A	304	LEU
1	A	307	LEU
1	A	332	LEU
1	A	333	ARG
1	B	7[A]	LEU
1	B	7[B]	LEU
1	B	13	LEU
1	B	14	HIS
1	B	40	LYS
1	B	59	GLU
1	B	65	ARG
1	B	81	LYS
1	B	156	LYS
1	B	158	LYS
1	B	161	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	212	ASP
1	B	217	LEU
1	B	230	LYS
1	B	232	LEU
1	B	238	LEU
1	B	251	ARG
1	B	259	ASN
1	B	282	LEU
1	B	292	VAL
1	B	300	THR
1	B	304	LEU
1	B	307	LEU
1	B	332	LEU
1	C	2	LYS
1	C	8	LEU
1	C	13	LEU
1	C	15	GLU
1	C	27	LEU
1	C	49	SER
1	C	50	ASP
1	C	59	GLU
1	C	142	GLU
1	C	217	LEU
1	C	232	LEU
1	C	259	ASN
1	C	267	TYR
1	C	274	GLN
1	C	282	LEU
1	C	292	VAL
1	C	296	ASN
1	C	297	ASN
1	C	299	GLN
1	C	304	LEU
1	C	307	LEU
1	C	326	ARG
1	D	2	LYS
1	D	8	LEU
1	D	13	LEU
1	D	14	HIS
1	D	15	GLU
1	D	17[A]	ILE
1	D	17[B]	ILE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	42	MET
1	D	47	LEU
1	D	51	LEU
1	D	185	TYR
1	D	217	LEU
1	D	232	LEU
1	D	259	ASN
1	D	264	MET
1	D	292	VAL
1	D	294[A]	HIS
1	D	294[B]	HIS
1	D	332	LEU
1	D	333	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (35) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	71	GLN
1	A	137	ASN
1	A	245	GLN
1	A	256	ASN
1	A	259	ASN
1	A	265	GLN
1	A	274	GLN
1	A	277	ASN
1	A	312	HIS
1	B	71	GLN
1	B	137	ASN
1	B	245	GLN
1	B	259	ASN
1	B	277	ASN
1	B	312	HIS
1	C	14	HIS
1	C	71	GLN
1	C	137	ASN
1	C	241	GLN
1	C	245	GLN
1	C	256	ASN
1	C	259	ASN
1	C	274	GLN
1	C	277	ASN
1	C	312	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	71	GLN
1	D	137	ASN
1	D	160	ASN
1	D	245	GLN
1	D	256	ASN
1	D	259	ASN
1	D	265	GLN
1	D	274	GLN
1	D	281	HIS
1	D	312	HIS

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

4 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
2	NF5	A	1334	-	21,23,23	0.33	0	30,31,31	0.63	0
2	NF5	B	1334	-	21,23,23	0.35	0	30,31,31	0.39	0
2	NF5	D	1334	-	21,23,23	0.36	0	30,31,31	0.61	0
2	NF5	C	1334	-	21,23,23	0.32	0	30,31,31	0.69	1 (3%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	NF5	A	1334	-	-	0/8/16/16	0/3/3/3
2	NF5	B	1334	-	-	0/8/16/16	0/3/3/3
2	NF5	D	1334	-	-	0/8/16/16	0/3/3/3
2	NF5	C	1334	-	-	0/8/16/16	0/3/3/3

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	1334	NF5	C11-N10-C8	2.48	133.01	126.58

There are no chirality outliers.

There are no torsion outliers.

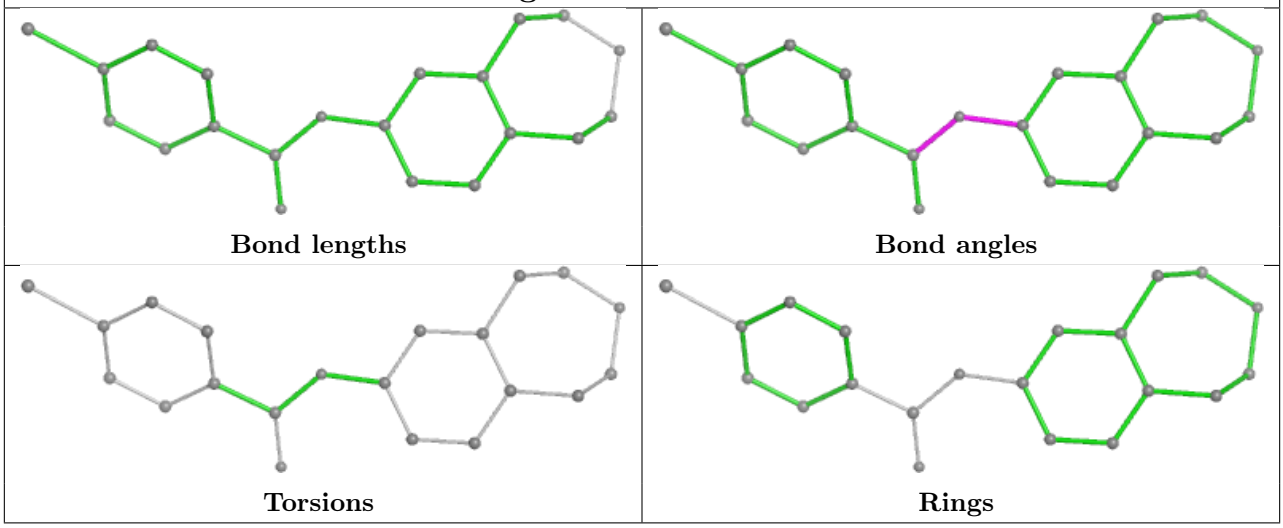
There are no ring outliers.

3 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1334	NF5	3	0
2	B	1334	NF5	2	0
2	D	1334	NF5	4	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 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 NF5 C 1334



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, 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	314/356 (88%)	0.25	19 (6%)	21	24	23, 37, 66, 97	0
1	B	323/356 (90%)	0.42	19 (5%)	22	25	22, 41, 79, 106	0
1	C	313/356 (87%)	0.50	27 (8%)	10	12	27, 48, 78, 135	0
1	D	326/356 (91%)	0.33	23 (7%)	16	18	20, 36, 72, 103	0
All	All	1276/1424 (89%)	0.37	88 (6%)	16	19	20, 40, 74, 135	0

All (88) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	186	LEU	8.9
1	B	169	CYS	8.1
1	B	330	VAL	7.5
1	B	327	GLY	6.8
1	C	267	TYR	6.4
1	A	169	CYS	6.3
1	D	184	SER	5.1
1	D	281	HIS	4.6
1	B	47	LEU	4.5
1	D	169	CYS	4.3
1	D	182	GLY	4.1
1	C	51	LEU	4.1
1	C	268	ASN	3.8
1	A	269	TYR	3.6
1	D	327	GLY	3.5
1	A	267	TYR	3.5
1	C	22	PHE	3.5
1	C	143	TYR	3.4
1	C	52	PRO	3.3
1	D	185	TYR	3.2
1	A	303	ASP	3.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	162	ASP	3.2
1	B	49	SER	3.2
1	A	155	ALA	3.2
1	B	332	LEU	3.1
1	C	280	ILE	3.1
1	A	50	ASP	3.1
1	D	332	LEU	3.1
1	D	328	LYS	3.1
1	C	156	LYS	3.1
1	D	183	LYS	3.0
1	D	326	ARG	3.0
1	A	299	GLN	3.0
1	D	187	GLY	3.0
1	C	138	LEU	2.9
1	B	331	ARG	2.9
1	C	264	MET	2.9
1	D	201	VAL	2.8
1	C	214	ALA	2.7
1	C	281	HIS	2.7
1	C	149	ILE	2.7
1	B	21	GLY	2.7
1	D	197	ILE	2.7
1	D	333	ARG	2.7
1	A	52	PRO	2.7
1	D	331	ARG	2.7
1	D	280	ILE	2.6
1	A	139	LEU	2.6
1	B	187	GLY	2.6
1	B	328	LYS	2.6
1	A	45	ASN	2.6
1	B	51	LEU	2.5
1	A	51	LEU	2.5
1	C	187	GLY	2.5
1	C	155	ALA	2.5
1	D	329	PRO	2.5
1	D	47	LEU	2.5
1	A	249	LYS	2.5
1	C	180	ILE	2.5
1	A	149	ILE	2.4
1	A	264	MET	2.4
1	D	49	SER	2.4
1	C	98	ILE	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	143	TYR	2.4
1	B	48	GLY	2.4
1	D	288	THR	2.4
1	C	307	LEU	2.3
1	C	94	LEU	2.3
1	A	268	ASN	2.3
1	B	197	ILE	2.3
1	C	183	LYS	2.3
1	A	138	LEU	2.3
1	D	102	ASP	2.3
1	B	46	THR	2.2
1	C	182	GLY	2.2
1	C	212	ASP	2.2
1	D	45	ASN	2.1
1	B	50	ASP	2.1
1	B	284	ASP	2.1
1	B	278	PRO	2.1
1	A	280	ILE	2.1
1	A	170	GLY	2.1
1	C	139	LEU	2.1
1	C	50	ASP	2.1
1	C	69	ILE	2.1
1	C	92	GLY	2.1
1	A	333	ARG	2.0
1	C	328	LYS	2.0

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)

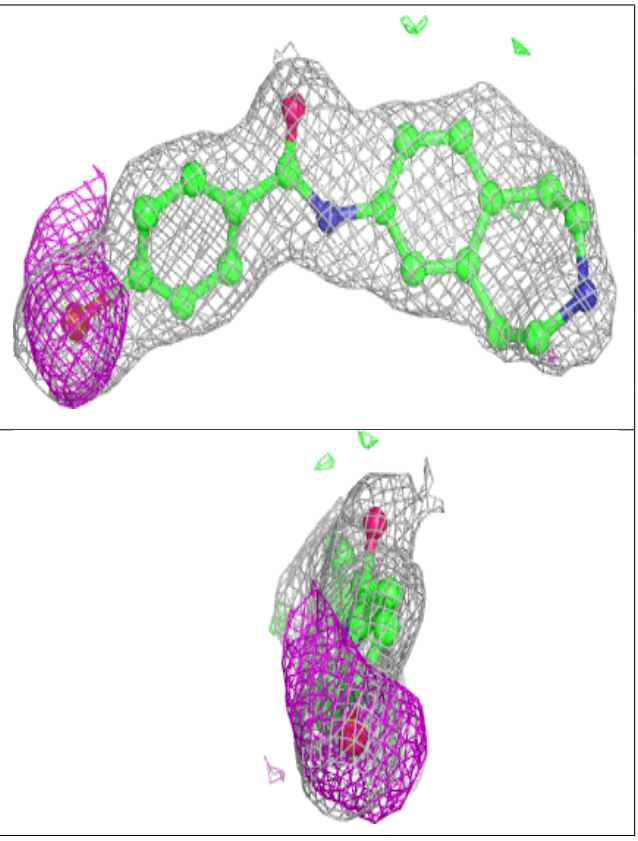
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	NF5	D	1334	21/21	0.89	0.14	26,35,50,60	0
2	NF5	B	1334	21/21	0.92	0.12	32,49,56,65	0
2	NF5	C	1334	21/21	0.93	0.14	24,48,56,58	0
2	NF5	A	1334	21/21	0.93	0.12	22,34,41,51	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 NF5 C 1334:

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.