Full wwPDB X-ray Structure Validation Report ⓘ

Oct 15, 2023 – 03:14 AM EDT

PDB ID	:	8DKB
Title	:	Crystal Structure of human YEATS4 in complex with Pfizer small molecule compound 3b
Authors	:	Dias, J.M.; Byrnes, L.J.; Varghese, A.H.
Deposited on	:	2022-07-05
Resolution	:	2.58 Å(reported)

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

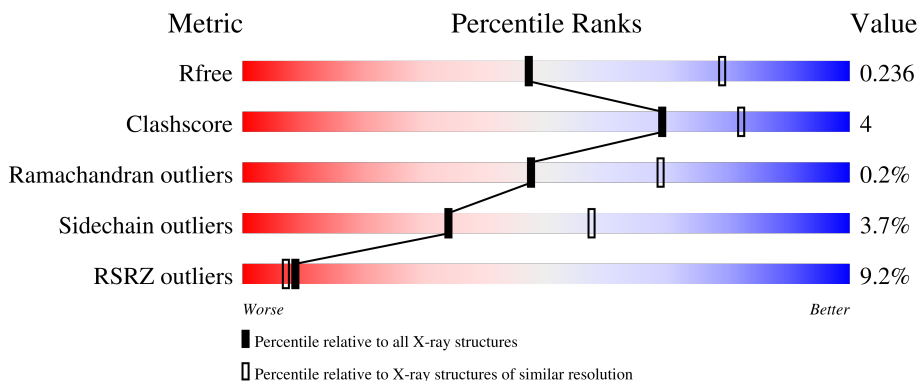|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.58 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.
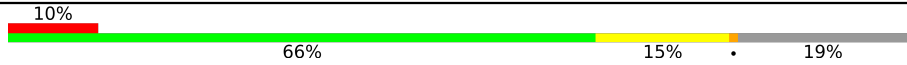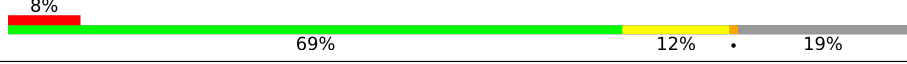
Metric	Percentile Ranks	Value
Rfree		0.236
Clashscore		4
Ramachandran outliers		0.2%
Sidechain outliers		3.7%
RSRZ outliers		9.2%

Worse — *Better*

▌ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3676 (2.60-2.56)
Clashscore	141614	4049 (2.60-2.56)
Ramachandran outliers	138981	3979 (2.60-2.56)
Sidechain outliers	138945	3979 (2.60-2.56)
RSRZ outliers	127900	3614 (2.60-2.56)

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	155	4% / 75% / 6% / 19%
1	B	155	% / 74% / 6% · / 19%
1	C	155	10% / 75% / 7% · / 17%
1	D	155	11% / 72% / 7% · / 19%
1	E	155	6% / 70% / 11% · / 18%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	155	10% 66% 15% • 19%
1	G	155	8% 69% 12% • 19%
1	H	155	10% 73% 8% • 19%

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 8996 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 YEATS domain-containing protein 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	126	Total	C	N	O	S	0	0	0
			1050	691	168	188	3			
1	B	126	Total	C	N	O	S	0	0	0
			1058	694	172	189	3			
1	C	129	Total	C	N	O	S	0	0	0
			1066	699	170	194	3			
1	D	125	Total	C	N	O	S	0	0	0
			1039	683	166	187	3			
1	E	127	Total	C	N	O	S	0	0	0
			1061	696	172	190	3			
1	F	126	Total	C	N	O	S	0	0	0
			1050	691	168	188	3			
1	G	126	Total	C	N	O	S	0	0	0
			1052	691	170	188	3			
1	H	126	Total	C	N	O	S	0	0	0
			1046	687	168	188	3			

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	150	HIS	-	expression tag	UNP O95619
A	151	HIS	-	expression tag	UNP O95619
A	152	HIS	-	expression tag	UNP O95619
A	153	HIS	-	expression tag	UNP O95619
A	154	HIS	-	expression tag	UNP O95619
A	155	HIS	-	expression tag	UNP O95619
B	150	HIS	-	expression tag	UNP O95619
B	151	HIS	-	expression tag	UNP O95619
B	152	HIS	-	expression tag	UNP O95619
B	153	HIS	-	expression tag	UNP O95619
B	154	HIS	-	expression tag	UNP O95619
B	155	HIS	-	expression tag	UNP O95619
C	150	HIS	-	expression tag	UNP O95619

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	151	HIS	-	expression tag	UNP O95619
C	152	HIS	-	expression tag	UNP O95619
C	153	HIS	-	expression tag	UNP O95619
C	154	HIS	-	expression tag	UNP O95619
C	155	HIS	-	expression tag	UNP O95619
D	150	HIS	-	expression tag	UNP O95619
D	151	HIS	-	expression tag	UNP O95619
D	152	HIS	-	expression tag	UNP O95619
D	153	HIS	-	expression tag	UNP O95619
D	154	HIS	-	expression tag	UNP O95619
D	155	HIS	-	expression tag	UNP O95619
E	150	HIS	-	expression tag	UNP O95619
E	151	HIS	-	expression tag	UNP O95619
E	152	HIS	-	expression tag	UNP O95619
E	153	HIS	-	expression tag	UNP O95619
E	154	HIS	-	expression tag	UNP O95619
E	155	HIS	-	expression tag	UNP O95619
F	150	HIS	-	expression tag	UNP O95619
F	151	HIS	-	expression tag	UNP O95619
F	152	HIS	-	expression tag	UNP O95619
F	153	HIS	-	expression tag	UNP O95619
F	154	HIS	-	expression tag	UNP O95619
F	155	HIS	-	expression tag	UNP O95619
G	150	HIS	-	expression tag	UNP O95619
G	151	HIS	-	expression tag	UNP O95619
G	152	HIS	-	expression tag	UNP O95619
G	153	HIS	-	expression tag	UNP O95619
G	154	HIS	-	expression tag	UNP O95619
G	155	HIS	-	expression tag	UNP O95619
H	150	HIS	-	expression tag	UNP O95619
H	151	HIS	-	expression tag	UNP O95619
H	152	HIS	-	expression tag	UNP O95619
H	153	HIS	-	expression tag	UNP O95619
H	154	HIS	-	expression tag	UNP O95619
H	155	HIS	-	expression tag	UNP O95619

- Molecule 2 is N-ethyl-1-{(3S,4S)-1-[(1-hydroxycyclohexyl)methyl]-3-methylpiperidin-4-yl}-2-methyl-1H-benzimidazole-5-carboxamide (three-letter code: SJI) (formula: $C_{24}H_{36}N_4O_2$) (labeled as "Ligand of Interest" by depositor).
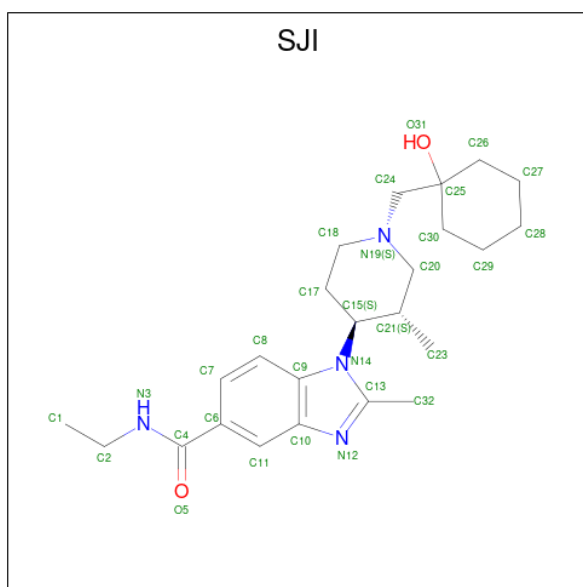
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total 30	C 24	N 4	O 2	0	0
2	B	1	Total 30	C 24	N 4	O 2	0	0
2	C	1	Total 30	C 24	N 4	O 2	0	0
2	D	1	Total 30	C 24	N 4	O 2	0	0
2	E	1	Total 30	C 24	N 4	O 2	0	0
2	F	1	Total 30	C 24	N 4	O 2	0	0
2	G	1	Total 30	C 24	N 4	O 2	0	0
2	H	1	Total 30	C 24	N 4	O 2	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	34	Total 34	O 34	0	0
3	B	49	Total 49	O 49	0	0
3	C	57	Total 57	O 57	0	0
3	D	34	Total 34	O 34	0	0

Continued on next page...

Continued from previous page...

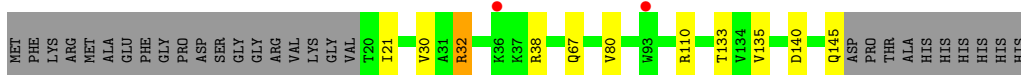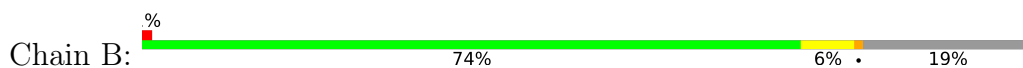| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	E	35	Total 35	O 35	0	0
3	F	39	Total 39	O 39	0	0
3	G	44	Total 44	O 44	0	0
3	H	42	Total 42	O 42	0	0

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.
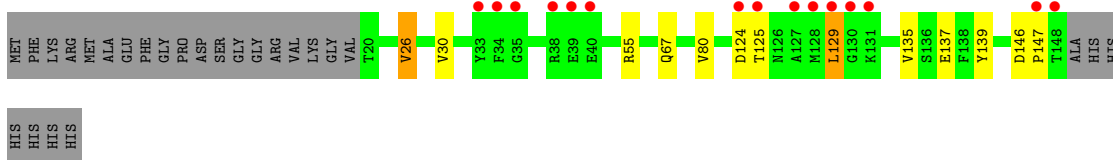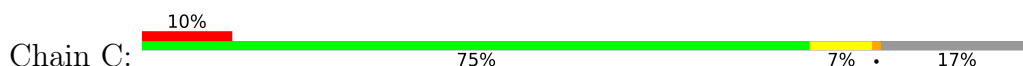
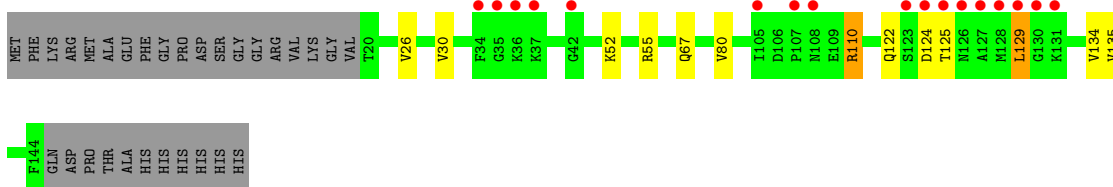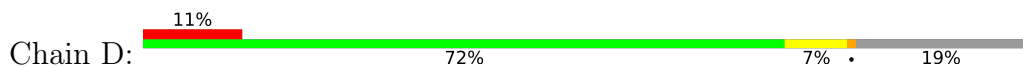• Molecule 1: YEATS domain-containing protein 4

Chain A:

• Molecule 1: YEATS domain-containing protein 4

Chain B:

• Molecule 1: YEATS domain-containing protein 4
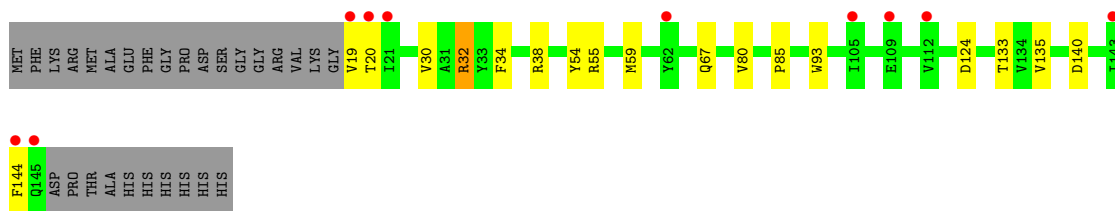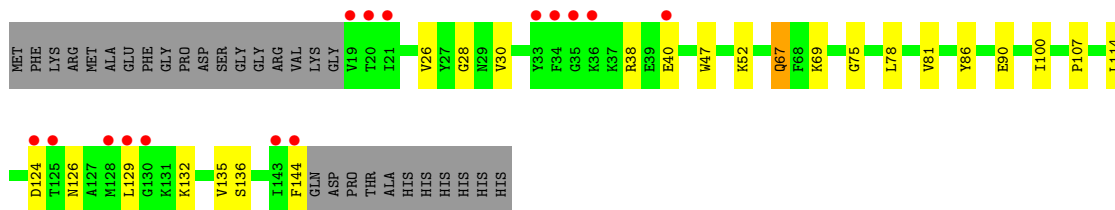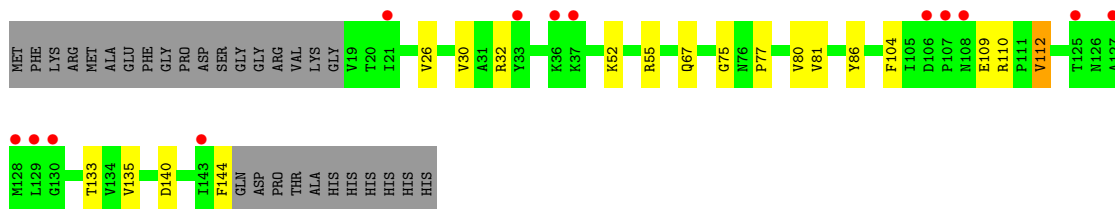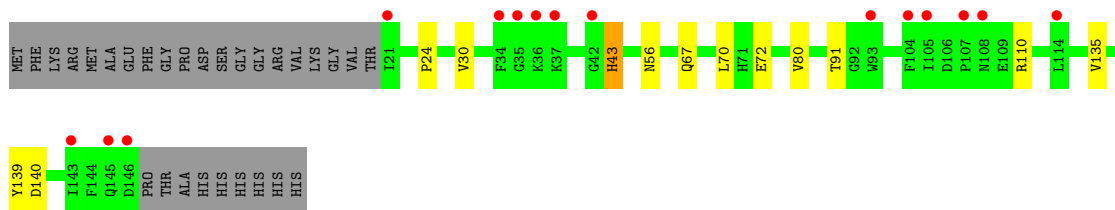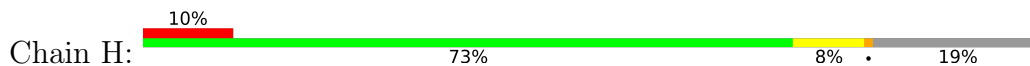
Chain C:

• Molecule 1: YEATS domain-containing protein 4

Chain D:

• Molecule 1: YEATS domain-containing protein 4

Chain E:

MET PHE LYS ARG MET ALA GLU PHE GLY PRO ASP SER GLY GLY ARG VAL LYS GLY V19 T20 I21 V30 A31 R32 Y33 F34 R38 Y54 R55 M59 Y62 Q67 V80 P85 W93 I105 E109 V112 D124 T133 V134 V135 D140 I143

F144 Q145 ASP PRO THR ALA HIS HIS HIS HIS HIS

• Molecule 1: YEATS domain-containing protein 4

Chain F:

10%

66%   15%   •   19%

MET PHE LYS ARG MET ALA GLU PHE GLY PRO ASP SER GLY GLY ARG VAL LYS GLY V19 T20 I21 V26 Y27 G28 N29 V30 Y33 F34 G35 K36 K37 R38 E39 E40 W47 K52 Q67 F68 K69 G75 L78 V81 Y86 E90 I100 P107 L114

D124 T125 N126 A127 M128 L129 G130 K131 K132 V135 S136 I143 F144 GLN ASP PRO THR ALA HIS HIS HIS HIS HIS HIS

• Molecule 1: YEATS domain-containing protein 4

Chain G:

8%

69%   12%   •   19%

MET PHE LYS ARG MET ALA GLU PHE GLY PRO ASP SER GLY GLY ARG VAL LYS GLY V19 T20 I21 V26 V30 A31 R32 Y33 K36 K37 K52 R55 Q67 G75 N76 P77 V80 V81 Y86 F104 I105 D106 P107 N108 E109 R110 P111 V112 T125 N126 A127

M128 L129 G130 T133 V134 V135 D140 I143 F144 GLN ASP PRO THR ALA HIS HIS HIS HIS HIS HIS

• Molecule 1: YEATS domain-containing protein 4

Chain H:

10%

73%   8%   •   19%

MET PHE LYS ARG MET ALA GLU PHE GLY PRO ASP SER GLY GLY ARG VAL LYS GLY VAL THR I21 P24 V30 F34 G35 K36 K37 G42 H43 N56 Q67 L70 H71 E72 V80 T91 G92 W93 F104 I105 D106 P107 N108 E109 R110 L14 V135

Y139 D140 I143 F144 Q145 D146 PRO THR ALA HIS HIS HIS HIS HIS HIS

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	60.60Å 109.30Å 103.83Å 90.00° 95.68° 90.00°	Depositor
Resolution (Å)	41.30 – 2.58 41.30 – 2.58	Depositor EDS
% Data completeness (in resolution range)	62.0 (41.30-2.58) 62.0 (41.30-2.58)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.12	Depositor
$< I/\sigma(I) >$ [1]	1.46 (at 2.58Å)	Xtriage
Refinement program	BUSTER 2.11.7	Depositor
R, R_{free}	0.171 , 0.226 0.182 , 0.236	Depositor DCC
R_{free} test set	1286 reflections (4.89%)	wwPDB-VP
Wilson B-factor (Å^2)	58.7	Xtriage
Anisotropy	0.091	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 66.2	EDS
L-test for twinning[2]	$< |L| > = 0.48, < L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	8996	wwPDB-VP
Average B, all atoms (Å^2)	72.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.19% 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: SJI

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	0.53	0/1081	0.68	0/1465
1	B	0.56	0/1089	0.67	0/1474
1	C	0.53	0/1098	0.70	0/1491
1	D	0.53	0/1070	0.68	0/1451
1	E	0.54	0/1092	0.71	1/1480 (0.1%)
1	F	0.55	0/1081	0.73	0/1465
1	G	0.54	0/1083	0.76	0/1468
1	H	0.50	0/1077	0.67	0/1460
All	All	0.54	0/8671	0.70	1/11754 (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	E	144	PHE	C-N-CA	5.03	134.26	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	1050	0	1040	6	0
1	B	1058	0	1050	6	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	1066	0	1040	12	0
1	D	1039	0	1020	12	0
1	E	1061	0	1048	10	0
1	F	1050	0	1040	15	0
1	G	1052	0	1040	11	0
1	H	1046	0	1023	5	0
2	A	30	0	0	0	0
2	B	30	0	0	0	0
2	C	30	0	0	0	0
2	D	30	0	0	0	0
2	E	30	0	0	0	0
2	F	30	0	0	0	0
2	G	30	0	0	0	0
2	H	30	0	0	0	0
3	A	34	0	0	0	0
3	B	49	0	0	0	0
3	C	57	0	0	0	0
3	D	34	0	0	0	0
3	E	35	0	0	0	0
3	F	39	0	0	1	0
3	G	44	0	0	0	0
3	H	42	0	0	0	0
All	All	8996	0	8301	63	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 4.

All (63) 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:B:38:ARG:HH11	1:D:129:LEU:HD11	1.27	0.97
1:E:93:TRP:CH2	1:F:107:PRO:HG2	1.98	0.97
1:C:55:ARG:HH21	1:C:55:ARG:HG3	1.46	0.81
1:C:129:LEU:HD11	1:E:38:ARG:HD3	1.61	0.81
1:C:26:VAL:HG12	1:C:137:GLU:HG2	1.70	0.74
1:D:125:THR:O	1:D:129:LEU:HD13	1.94	0.68
1:D:67:GLN:HG3	1:D:80:VAL:HG22	1.76	0.67
1:H:67:GLN:HG3	1:H:80:VAL:HG22	1.76	0.66
1:B:38:ARG:NH1	1:D:129:LEU:HD11	2.08	0.65
1:A:67:GLN:HG3	1:A:80:VAL:HG22	1.80	0.62
1:D:26:VAL:HG22	1:D:52:LYS:HG2	1.80	0.62

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:F:126:ASN:HA	1:F:129:LEU:HD22	1.82	0.62
1:E:67:GLN:HG3	1:E:80:VAL:HG22	1.81	0.61
1:B:67:GLN:HG3	1:B:80:VAL:HG22	1.82	0.61
1:C:129:LEU:O	1:E:34:PHE:O	2.19	0.61
1:C:67:GLN:HG3	1:C:80:VAL:HG22	1.81	0.60
1:G:32:ARG:HG3	1:G:133:THR:HG22	1.82	0.59
1:C:55:ARG:HH22	1:E:54:TYR:HA	1.67	0.59
1:G:26:VAL:CG2	1:G:52:LYS:HG2	2.33	0.58
1:C:125:THR:O	1:C:129:LEU:HD13	2.03	0.57
1:C:30:VAL:HG12	1:C:135:VAL:HG22	1.87	0.57
1:H:30:VAL:HG12	1:H:135:VAL:HG22	1.86	0.57
1:F:81:VAL:HG11	1:F:86:TYR:HA	1.87	0.57
1:G:26:VAL:HG22	1:G:52:LYS:HG2	1.87	0.56
1:D:26:VAL:CG2	1:D:52:LYS:HG2	2.34	0.56
1:B:32:ARG:HG3	1:B:133:THR:HG22	1.87	0.56
1:G:110:ARG:HH12	1:G:112:VAL:HG23	1.71	0.55
1:F:30:VAL:HG12	1:F:135:VAL:HG22	1.87	0.55
1:C:26:VAL:CG1	1:C:137:GLU:HG2	2.36	0.54
1:F:26:VAL:HG22	1:F:52:LYS:HG2	1.90	0.54
1:E:32:ARG:HG3	1:E:133:THR:HG22	1.89	0.53
1:F:38:ARG:HG2	1:F:40:GLU:HG2	1.89	0.53
1:D:30:VAL:HG12	1:D:135:VAL:HG22	1.91	0.52
1:F:26:VAL:CG2	1:F:52:LYS:HG2	2.40	0.52
1:G:30:VAL:HG12	1:G:135:VAL:HG22	1.92	0.52
1:C:55:ARG:HG3	1:C:55:ARG:NH2	2.21	0.51
1:B:30:VAL:HG12	1:B:135:VAL:HG22	1.94	0.50
1:A:110:ARG:HG3	1:F:75:GLY:HA3	1.94	0.48
1:A:115:TYR:CE1	1:F:40:GLU:HB2	2.49	0.47
1:D:110:ARG:H	1:D:110:ARG:HG3	1.55	0.47
1:A:141:GLU:OE1	1:H:24:PRO:HG3	2.15	0.47
1:F:100:ILE:HB	1:F:114:LEU:HB2	1.97	0.47
1:E:30:VAL:HG12	1:E:135:VAL:HG22	1.96	0.46
1:E:93:TRP:HH2	1:F:107:PRO:HG2	1.70	0.46
1:C:139:TYR:CD1	1:E:85:PRO:HD3	2.51	0.46
1:D:122:GLN:HB2	1:D:134:VAL:HG22	1.99	0.45
1:F:28:GLY:HA3	1:F:136:SER:O	2.17	0.44
1:F:67:GLN:OE1	1:F:78:LEU:HD11	2.18	0.44
1:G:75:GLY:O	1:G:77:PRO:HD3	2.18	0.44
1:A:54:TYR:CE1	1:H:139:TYR:HB2	2.53	0.43
1:D:26:VAL:HG22	1:D:52:LYS:CG	2.49	0.42
1:H:43:HIS:HD2	1:H:91:THR:OG1	2.02	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:19:VAL:HG23	1:A:20:THR:H	1.84	0.42
1:G:67:GLN:HG3	1:G:80:VAL:HG22	2.01	0.42
1:G:26:VAL:HG22	1:G:52:LYS:CG	2.49	0.42
1:G:104:PHE:HD2	1:G:109:GLU:HB3	1.85	0.41
1:C:129:LEU:HG	1:E:34:PHE:CD2	2.56	0.41
1:G:81:VAL:HG11	1:G:86:TYR:HA	2.01	0.41
1:D:55:ARG:HD2	1:D:55:ARG:HA	1.93	0.41
1:F:47:TRP:CZ2	1:F:90:GLU:HB2	2.56	0.40
1:G:104:PHE:CD2	1:G:109:GLU:HB3	2.57	0.40
1:B:38:ARG:HD3	1:D:129:LEU:HD12	2.03	0.40
1:F:69:LYS:HD2	3:F:4132:HOH:O	2.22	0.40

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	124/155 (80%)	124 (100%)	0	0	100	100
1	B	124/155 (80%)	121 (98%)	3 (2%)	0	100	100
1	C	127/155 (82%)	126 (99%)	0	1 (1%)	19	37
1	D	123/155 (79%)	122 (99%)	1 (1%)	0	100	100
1	E	125/155 (81%)	124 (99%)	1 (1%)	0	100	100
1	F	124/155 (80%)	119 (96%)	5 (4%)	0	100	100
1	G	124/155 (80%)	121 (98%)	3 (2%)	0	100	100
1	H	124/155 (80%)	123 (99%)	0	1 (1%)	19	37
All	All	995/1240 (80%)	980 (98%)	13 (1%)	2 (0%)	47	69

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	H	56	ASN
1	C	147	PRO

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	115/139 (83%)	113 (98%)	2 (2%)	60	79
1	B	116/139 (84%)	111 (96%)	5 (4%)	29	52
1	C	116/139 (84%)	112 (97%)	4 (3%)	37	60
1	D	113/139 (81%)	110 (97%)	3 (3%)	44	68
1	E	116/139 (84%)	109 (94%)	7 (6%)	19	37
1	F	115/139 (83%)	111 (96%)	4 (4%)	36	59
1	G	115/139 (83%)	111 (96%)	4 (4%)	36	59
1	H	113/139 (81%)	108 (96%)	5 (4%)	28	51
All	All	919/1112 (83%)	885 (96%)	34 (4%)	34	57

All (34) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	132	LYS
1	A	140	ASP
1	B	21	ILE
1	B	32	ARG
1	B	110	ARG
1	B	140	ASP
1	B	145	GLN
1	C	26	VAL
1	C	124	ASP
1	C	129	LEU
1	C	146	ASP
1	D	110	ARG
1	D	124	ASP
1	D	129	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	E	19	VAL
1	E	20	THR
1	E	32	ARG
1	E	55	ARG
1	E	59	MET
1	E	124	ASP
1	E	140	ASP
1	F	67	GLN
1	F	124	ASP
1	F	132	LYS
1	F	144	PHE
1	G	55	ARG
1	G	112	VAL
1	G	140	ASP
1	G	144	PHE
1	H	43	HIS
1	H	70	LEU
1	H	72	GLU
1	H	110	ARG
1	H	140	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	108	ASN
1	E	145	GLN
1	H	43	HIS

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 ⓘ

8 ligands are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#$|Z| > 2$	Counts	RMSZ	#$|Z| > 2$
2	SJI	E	4000	-	30,33,33	0.64	0	35,48,48	0.91	2 (5%)
2	SJI	C	4000	-	30,33,33	0.77	1 (3%)	35,48,48	0.90	1 (2%)
2	SJI	H	4000	-	30,33,33	0.74	1 (3%)	35,48,48	0.93	1 (2%)
2	SJI	F	4000	-	30,33,33	0.78	1 (3%)	35,48,48	0.97	1 (2%)
2	SJI	A	4000	-	30,33,33	0.62	0	35,48,48	0.92	1 (2%)
2	SJI	B	4000	-	30,33,33	0.55	0	35,48,48	0.76	1 (2%)
2	SJI	G	4000	-	30,33,33	0.69	0	35,48,48	0.84	1 (2%)
2	SJI	D	4000	-	30,33,33	0.68	0	35,48,48	0.85	1 (2%)

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	SJI	E	4000	-	-	0/12/39/39	0/4/4/4
2	SJI	C	4000	-	-	0/12/39/39	0/4/4/4
2	SJI	H	4000	-	-	0/12/39/39	0/4/4/4
2	SJI	F	4000	-	-	0/12/39/39	0/4/4/4
2	SJI	A	4000	-	-	0/12/39/39	0/4/4/4
2	SJI	B	4000	-	-	1/12/39/39	0/4/4/4
2	SJI	G	4000	-	-	0/12/39/39	0/4/4/4
2	SJI	D	4000	-	-	0/12/39/39	0/4/4/4

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	4000	SJI	C24-C25	-2.74	1.49	1.54
2	F	4000	SJI	C24-C25	-2.67	1.50	1.54

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	H	4000	SJI	C24-C25	-2.57	1.50	1.54

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	4000	SJI	C27-C26-C25	2.81	116.73	112.96
2	F	4000	SJI	C20-N19-C24	-2.66	107.12	111.92
2	B	4000	SJI	C29-C30-C25	2.64	116.49	112.96
2	H	4000	SJI	C29-C30-C25	2.28	116.01	112.96
2	G	4000	SJI	C29-C30-C25	2.26	115.99	112.96
2	E	4000	SJI	C29-C30-C25	2.19	115.89	112.96
2	D	4000	SJI	C29-C30-C25	2.14	115.82	112.96
2	C	4000	SJI	C29-C30-C25	2.10	115.78	112.96
2	E	4000	SJI	C27-C26-C25	2.04	115.70	112.96

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	4000	SJI	C25-C24-N19-C18

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 SJI E 4000

Bond lengths

Bond angles

Torsions

Rings

Ligand SJI C 4000

Bond lengths

Bond angles

Torsions

Rings

Ligand SJI H 4000

Bond lengths

Bond angles

Torsions

Rings

Ligand SJI F 4000

Bond lengths

Bond angles

Torsions

Rings

Ligand SJI A 4000

Bond lengths

Bond angles

Torsions

Rings

Ligand SJI B 4000

Bond lengths

Bond angles

Torsions

Rings

Ligand SJI G 4000

Bond lengths

Bond angles

Torsions

Rings

Ligand SJI D 4000

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(Å²)	Q<0.9
1	A	126/155 (81%)	0.27	6 (4%)	30	26	36, 63, 106, 125	0
1	B	126/155 (81%)	0.07	2 (1%)	72	69	33, 56, 91, 117	0
1	C	129/155 (83%)	0.46	15 (11%)	4	3	35, 59, 145, 157	0
1	D	125/155 (80%)	0.64	17 (13%)	3	2	41, 71, 128, 144	0
1	E	127/155 (81%)	0.49	10 (7%)	12	10	42, 74, 117, 144	0
1	F	126/155 (81%)	0.56	15 (11%)	4	3	43, 78, 138, 159	0
1	G	126/155 (81%)	0.43	13 (10%)	6	5	44, 62, 116, 134	0
1	H	126/155 (81%)	0.58	15 (11%)	4	3	52, 82, 132, 144	0
All	All	1011/1240 (81%)	0.44	93 (9%)	9	7	33, 70, 126, 159	0

All (93) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	128	MET	9.9
1	C	130	GLY	7.6
1	G	128	MET	6.7
1	E	105	ILE	6.6
1	D	129	LEU	6.5
1	F	128	MET	6.0
1	D	128	MET	5.9
1	C	129	LEU	5.9
1	F	130	GLY	5.6
1	H	36	LYS	5.5
1	F	21	ILE	5.4
1	C	34	PHE	5.2
1	F	129	LEU	5.1
1	A	19	VAL	4.9
1	E	62	TYR	4.9
1	E	19	VAL	4.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	21	ILE	4.5
1	D	127	ALA	4.3
1	D	125	THR	4.2
1	D	130	GLY	4.2
1	E	144	PHE	3.9
1	H	37	LYS	3.9
1	E	145	GLN	3.8
1	D	126	ASN	3.8
1	F	124	ASP	3.8
1	H	105	ILE	3.7
1	G	127	ALA	3.7
1	C	127	ALA	3.7
1	G	106	ASP	3.6
1	C	39	GLU	3.5
1	D	123	SER	3.4
1	C	35	GLY	3.3
1	F	125	THR	3.2
1	F	143	ILE	3.2
1	G	125	THR	3.2
1	F	19	VAL	3.2
1	G	130	GLY	3.1
1	D	34	PHE	3.1
1	H	35	GLY	3.0
1	D	108	ASN	3.0
1	B	36	LYS	3.0
1	G	21	ILE	2.9
1	D	107	PRO	2.9
1	F	20	THR	2.9
1	F	144	PHE	2.9
1	H	21	ILE	2.9
1	H	146	ASP	2.8
1	D	37	LYS	2.8
1	E	21	ILE	2.8
1	C	125	THR	2.8
1	G	129	LEU	2.8
1	A	143	ILE	2.7
1	H	107	PRO	2.7
1	C	131	LYS	2.7
1	H	93	TRP	2.7
1	D	105	ILE	2.6
1	H	114	LEU	2.6
1	C	124	ASP	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	131	LYS	2.6
1	E	20	THR	2.5
1	F	33	TYR	2.5
1	D	124	ASP	2.5
1	E	112	VAL	2.5
1	F	34	PHE	2.5
1	H	145	GLN	2.5
1	D	35	GLY	2.5
1	G	36	LYS	2.5
1	H	104	PHE	2.5
1	F	40	GLU	2.4
1	C	38	ARG	2.4
1	G	33	TYR	2.4
1	E	109	GLU	2.4
1	C	147	PRO	2.4
1	C	33	TYR	2.4
1	G	108	ASN	2.4
1	H	42	GLY	2.4
1	H	34	PHE	2.3
1	A	144	PHE	2.3
1	F	35	GLY	2.3
1	F	36	LYS	2.3
1	H	108	ASN	2.3
1	A	34	PHE	2.3
1	A	108	ASN	2.3
1	E	143	ILE	2.3
1	B	93	TRP	2.2
1	C	148	THR	2.2
1	G	37	LYS	2.2
1	D	42	GLY	2.2
1	H	143	ILE	2.2
1	C	40	GLU	2.2
1	D	36	LYS	2.1
1	G	143	ILE	2.0
1	G	107	PRO	2.0

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, 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	SJI	H	4000	30/30	0.90	0.34	102,107,110,113	0
2	SJI	F	4000	30/30	0.92	0.20	69,78,84,86	0
2	SJI	D	4000	30/30	0.94	0.19	58,67,85,90	0
2	SJI	G	4000	30/30	0.94	0.21	49,70,79,85	0
2	SJI	E	4000	30/30	0.94	0.25	78,84,96,97	0
2	SJI	B	4000	30/30	0.96	0.20	46,58,69,73	0
2	SJI	C	4000	30/30	0.98	0.16	52,58,77,80	0
2	SJI	A	4000	30/30	0.98	0.17	37,58,63,64	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.
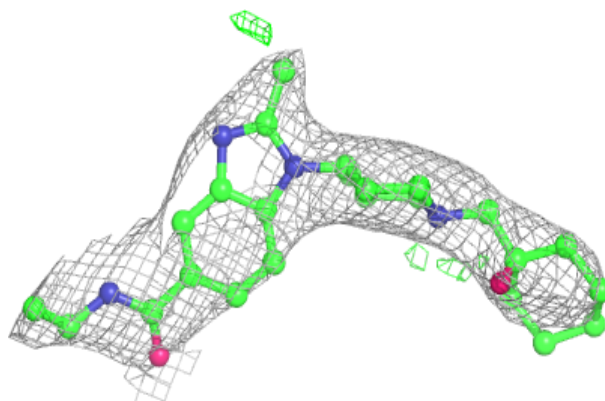
Electron density around SJI H 4000:

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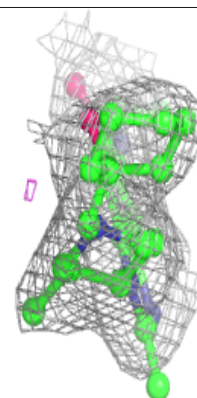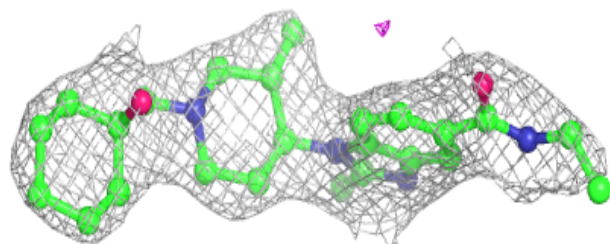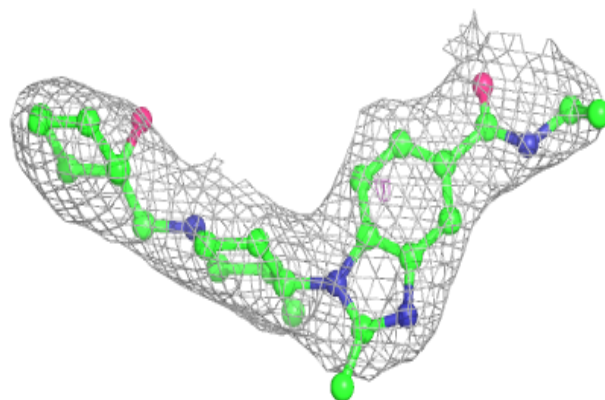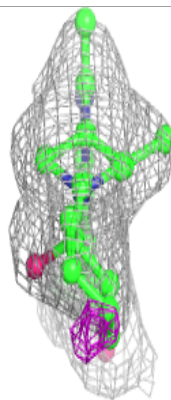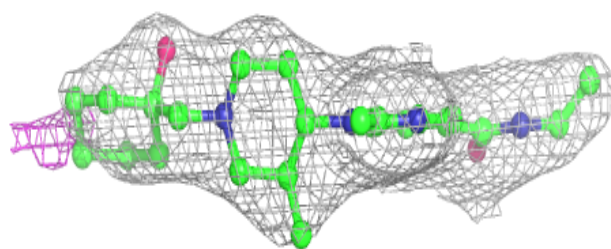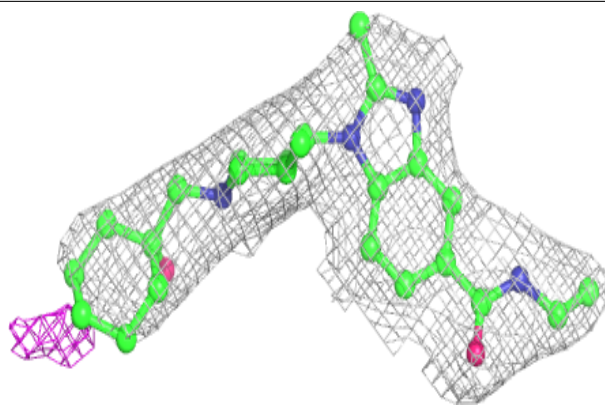mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
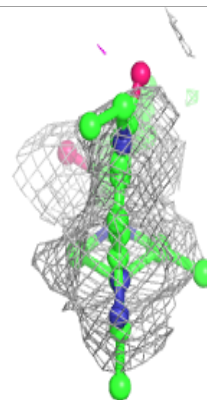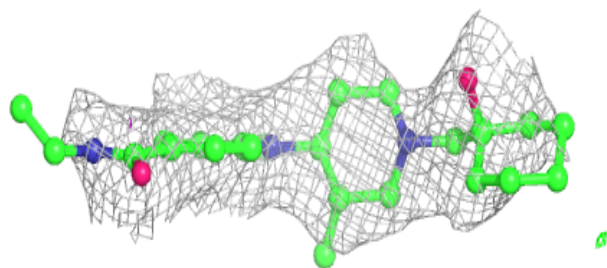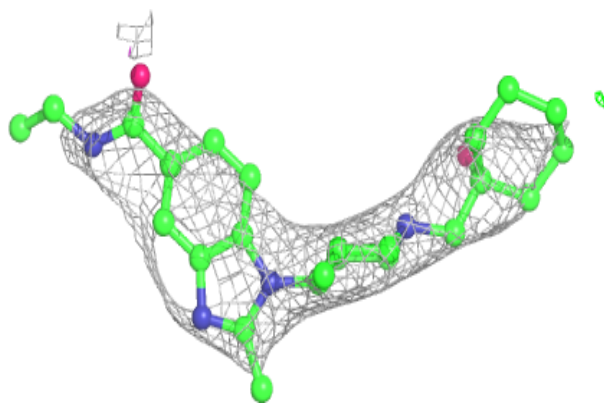

Electron density around SJI F 4000:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
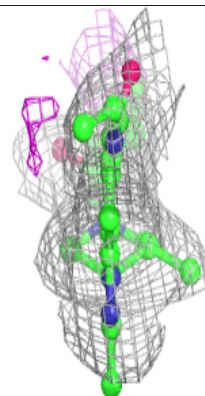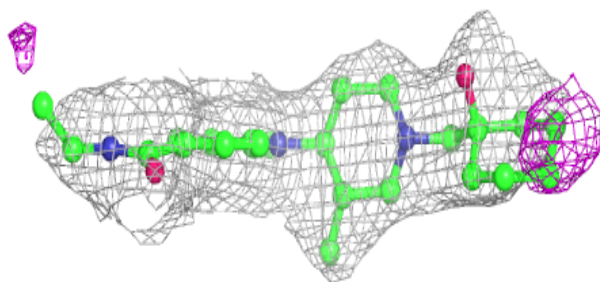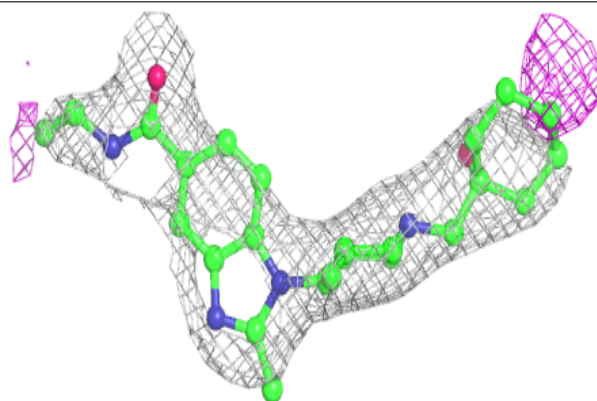
Electron density around SJI D 4000:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)
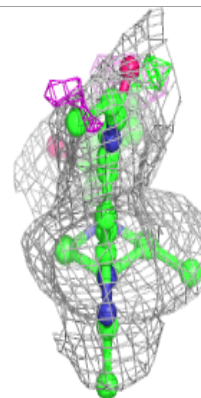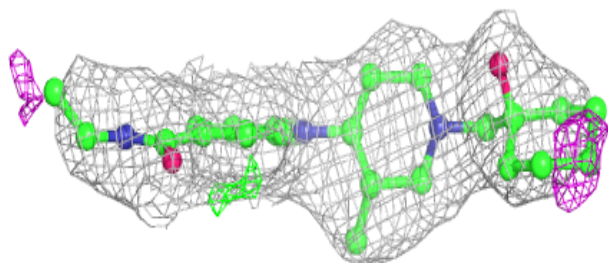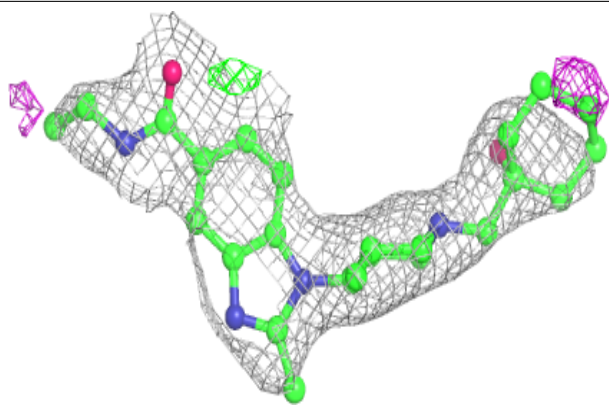
Electron density around SJI G 4000:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)
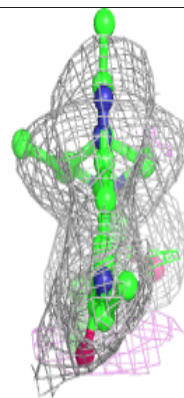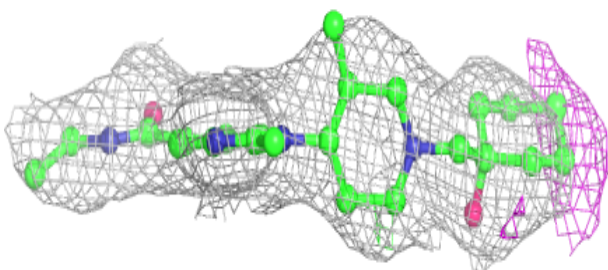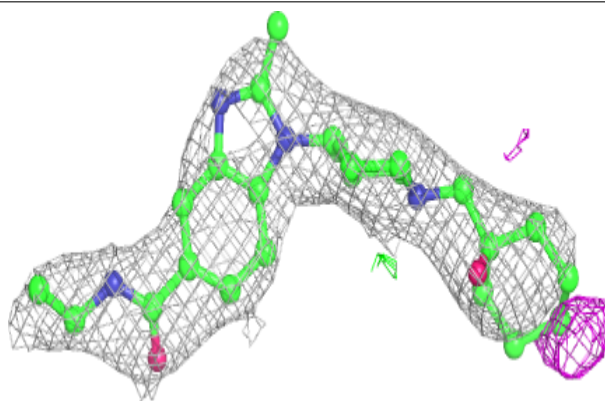
Electron density around SJI E 4000:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
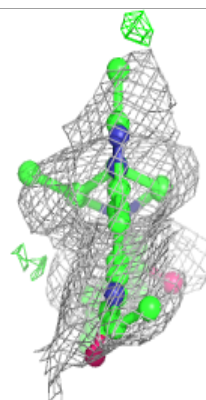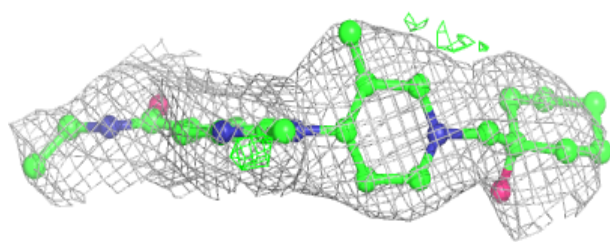
Electron density around SJI B 4000:

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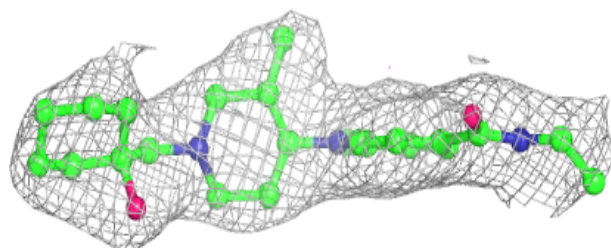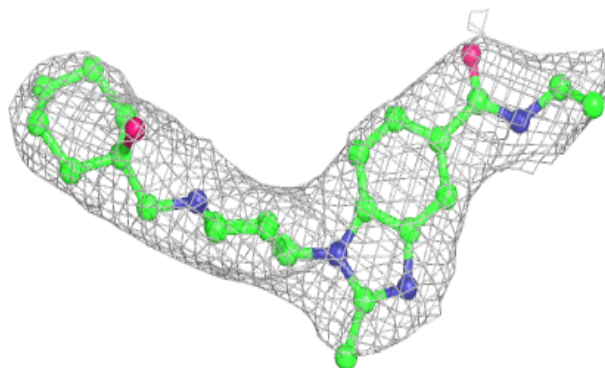mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

Electron density around SJI C 4000:

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

Electron density around SJI A 4000:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)
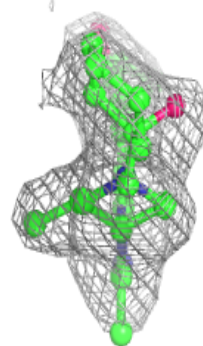
6.5 Other polymers (i)

There are no such residues in this entry.