wwPDB X-ray Structure Validation Summary Report (i)

Jan 18, 2024 – 12:09 PM JST

PDB ID	:	8JNR
Title	:	Crystal structure of human ALKBH3 bound to 3mC containing ssDNA through distal crosslink
Authors	:	Zhang, L.
Deposited on	:	2023-06-06
Resolution	:	3.66 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 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:

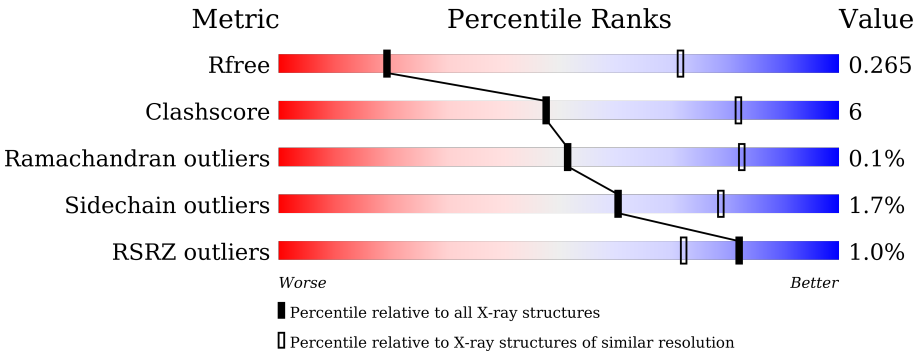
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (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 3.66 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	1557 (3.82-3.50)
Clashscore	141614	1037 (3.80-3.52)
Ramachandran outliers	138981	1004 (3.80-3.52)
Sidechain outliers	138945	1002 (3.80-3.52)
RSRZ outliers	127900	1441 (3.82-3.50)

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	238	% 78% 10% 11%
1	C	238	3% 69% 8% 23%
1	D	238	4% 82% 6% 12%
1	E	238	2% 72% 10% 18%
2	B	3	67% 33%
2	F	3	33% 33% 33%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	G	216	88% 11%
3	H	216	82% 14%
3	I	216	83% 12%
3	J	216	% 82% 14%
4	K	217	88% 11%
4	L	217	89% 10%
4	M	217	89% 9%
4	N	217	91% 6%

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 19651 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 Alpha-ketoglutarate-dependent dioxygenase alkB homolog 3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	211	Total	C	N	O	S	0	1	0
			1750	1109	310	324	7			
1	C	184	Total	C	N	O	S	0	0	0
			1520	964	274	276	6			
1	D	210	Total	C	N	O	S	0	0	0
			1735	1099	307	323	6			
1	E	195	Total	C	N	O	S	0	0	0
			1613	1023	288	296	6			

There are 96 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	49	MET	-	initiating methionine	UNP Q96Q83
A	50	GLY	-	expression tag	UNP Q96Q83
A	51	SER	-	expression tag	UNP Q96Q83
A	52	SER	-	expression tag	UNP Q96Q83
A	53	HIS	-	expression tag	UNP Q96Q83
A	54	HIS	-	expression tag	UNP Q96Q83
A	55	HIS	-	expression tag	UNP Q96Q83
A	56	HIS	-	expression tag	UNP Q96Q83
A	57	HIS	-	expression tag	UNP Q96Q83
A	58	HIS	-	expression tag	UNP Q96Q83
A	59	SER	-	expression tag	UNP Q96Q83
A	60	SER	-	expression tag	UNP Q96Q83
A	61	GLY	-	expression tag	UNP Q96Q83
A	62	LEU	-	expression tag	UNP Q96Q83
A	63	VAL	-	expression tag	UNP Q96Q83
A	64	PRO	-	expression tag	UNP Q96Q83
A	65	ARG	-	expression tag	UNP Q96Q83
A	66	GLY	-	expression tag	UNP Q96Q83
A	67	SER	-	expression tag	UNP Q96Q83
A	68	HIS	-	expression tag	UNP Q96Q83
A	69	MET	-	expression tag	UNP Q96Q83

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	110	SER	CYS	engineered mutation	UNP Q96Q83
A	189	CYS	ASP	engineered mutation	UNP Q96Q83
A	201	SER	CYS	engineered mutation	UNP Q96Q83
C	49	MET	-	initiating methionine	UNP Q96Q83
C	50	GLY	-	expression tag	UNP Q96Q83
C	51	SER	-	expression tag	UNP Q96Q83
C	52	SER	-	expression tag	UNP Q96Q83
C	53	HIS	-	expression tag	UNP Q96Q83
C	54	HIS	-	expression tag	UNP Q96Q83
C	55	HIS	-	expression tag	UNP Q96Q83
C	56	HIS	-	expression tag	UNP Q96Q83
C	57	HIS	-	expression tag	UNP Q96Q83
C	58	HIS	-	expression tag	UNP Q96Q83
C	59	SER	-	expression tag	UNP Q96Q83
C	60	SER	-	expression tag	UNP Q96Q83
C	61	GLY	-	expression tag	UNP Q96Q83
C	62	LEU	-	expression tag	UNP Q96Q83
C	63	VAL	-	expression tag	UNP Q96Q83
C	64	PRO	-	expression tag	UNP Q96Q83
C	65	ARG	-	expression tag	UNP Q96Q83
C	66	GLY	-	expression tag	UNP Q96Q83
C	67	SER	-	expression tag	UNP Q96Q83
C	68	HIS	-	expression tag	UNP Q96Q83
C	69	MET	-	expression tag	UNP Q96Q83
C	110	SER	CYS	engineered mutation	UNP Q96Q83
C	189	CYS	ASP	engineered mutation	UNP Q96Q83
C	201	SER	CYS	engineered mutation	UNP Q96Q83
D	49	MET	-	initiating methionine	UNP Q96Q83
D	50	GLY	-	expression tag	UNP Q96Q83
D	51	SER	-	expression tag	UNP Q96Q83
D	52	SER	-	expression tag	UNP Q96Q83
D	53	HIS	-	expression tag	UNP Q96Q83
D	54	HIS	-	expression tag	UNP Q96Q83
D	55	HIS	-	expression tag	UNP Q96Q83
D	56	HIS	-	expression tag	UNP Q96Q83
D	57	HIS	-	expression tag	UNP Q96Q83
D	58	HIS	-	expression tag	UNP Q96Q83
D	59	SER	-	expression tag	UNP Q96Q83
D	60	SER	-	expression tag	UNP Q96Q83
D	61	GLY	-	expression tag	UNP Q96Q83
D	62	LEU	-	expression tag	UNP Q96Q83
D	63	VAL	-	expression tag	UNP Q96Q83

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	64	PRO	-	expression tag	UNP Q96Q83
D	65	ARG	-	expression tag	UNP Q96Q83
D	66	GLY	-	expression tag	UNP Q96Q83
D	67	SER	-	expression tag	UNP Q96Q83
D	68	HIS	-	expression tag	UNP Q96Q83
D	69	MET	-	expression tag	UNP Q96Q83
D	110	SER	CYS	engineered mutation	UNP Q96Q83
D	189	CYS	ASP	engineered mutation	UNP Q96Q83
D	201	SER	CYS	engineered mutation	UNP Q96Q83
E	49	MET	-	initiating methionine	UNP Q96Q83
E	50	GLY	-	expression tag	UNP Q96Q83
E	51	SER	-	expression tag	UNP Q96Q83
E	52	SER	-	expression tag	UNP Q96Q83
E	53	HIS	-	expression tag	UNP Q96Q83
E	54	HIS	-	expression tag	UNP Q96Q83
E	55	HIS	-	expression tag	UNP Q96Q83
E	56	HIS	-	expression tag	UNP Q96Q83
E	57	HIS	-	expression tag	UNP Q96Q83
E	58	HIS	-	expression tag	UNP Q96Q83
E	59	SER	-	expression tag	UNP Q96Q83
E	60	SER	-	expression tag	UNP Q96Q83
E	61	GLY	-	expression tag	UNP Q96Q83
E	62	LEU	-	expression tag	UNP Q96Q83
E	63	VAL	-	expression tag	UNP Q96Q83
E	64	PRO	-	expression tag	UNP Q96Q83
E	65	ARG	-	expression tag	UNP Q96Q83
E	66	GLY	-	expression tag	UNP Q96Q83
E	67	SER	-	expression tag	UNP Q96Q83
E	68	HIS	-	expression tag	UNP Q96Q83
E	69	MET	-	expression tag	UNP Q96Q83
E	110	SER	CYS	engineered mutation	UNP Q96Q83
E	189	CYS	ASP	engineered mutation	UNP Q96Q83
E	201	SER	CYS	engineered mutation	UNP Q96Q83

- Molecule 2 is a DNA chain called DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	3	Total	C	N	O	P	0	0	0
			60	30	7	20	3			
2	F	2	Total	C	N	O	P	0	0	0
			40	20	5	13	2			

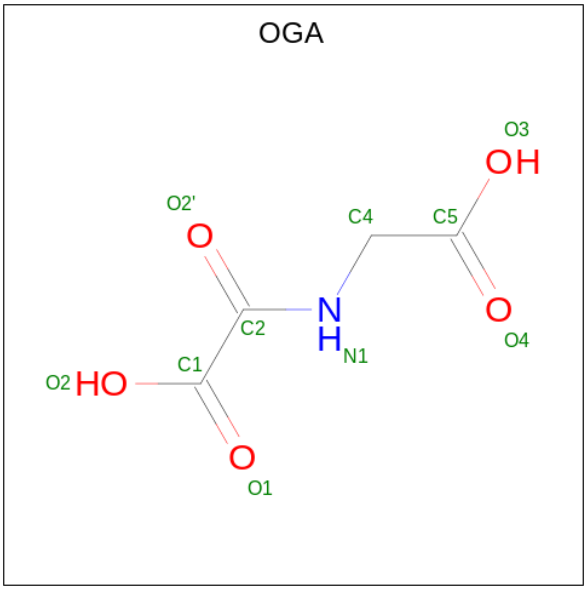
- Molecule 3 is a protein called Synthetic antibody heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	G	214	Total	C	N	O	S	0	0	0
			1601	1014	264	317	6			
3	H	210	Total	C	N	O	S	0	0	0
			1578	1002	260	310	6			
3	I	210	Total	C	N	O	S	0	0	0
			1572	997	259	310	6			
3	J	209	Total	C	N	O	S	0	0	0
			1566	994	258	308	6			

- Molecule 4 is a protein called Synthetic antibody light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	K	214	Total	C	N	O	S	0	0	0
			1643	1032	274	332	5			
4	L	214	Total	C	N	O	S	0	0	0
			1643	1032	274	332	5			
4	M	214	Total	C	N	O	S	0	0	0
			1643	1032	274	332	5			
4	N	214	Total	C	N	O	S	0	0	0
			1643	1032	274	332	5			

- Molecule 5 is N-OXALYLGLYCINE (three-letter code: OGA) (formula: $C_4H_5NO_5$).

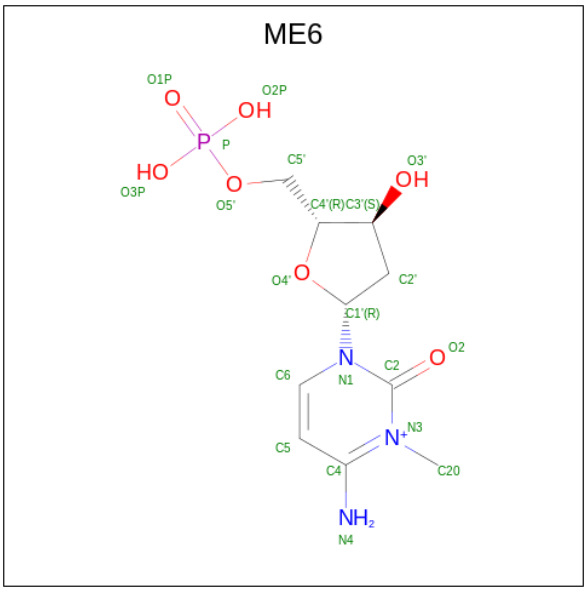


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	N	O	0	0
			10	4	1	5		
5	D	1	Total	C	N	O	0	0
			10	4	1	5		

- Molecule 6 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total Mn 1 1	0	0
6	C	1	Total Mn 1 1	0	0
6	D	1	Total Mn 1 1	0	0
6	E	1	Total Mn 1 1	0	0

- Molecule 7 is [(2R,3S,5R)-5-(4-azanyl-3-methyl-2-oxo-pyrimidin-3-ium-1-yl)-3-hydroxy-oxolan-2-yl]methyl dihydrogen phosphate (three-letter code: ME6) (formula: $C_{10}H_{17}N_3O_7P$) (labeled as "Ligand of Interest" by depositor).



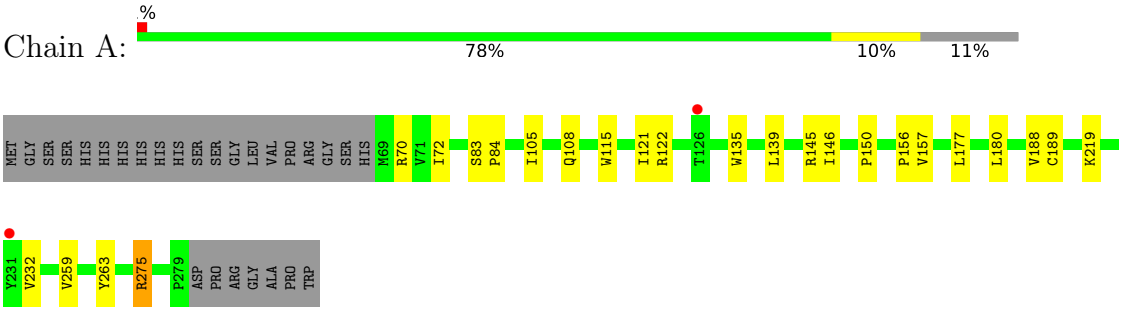
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	D	1	Total C N O P 20 10 3 6 1	0	0

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.

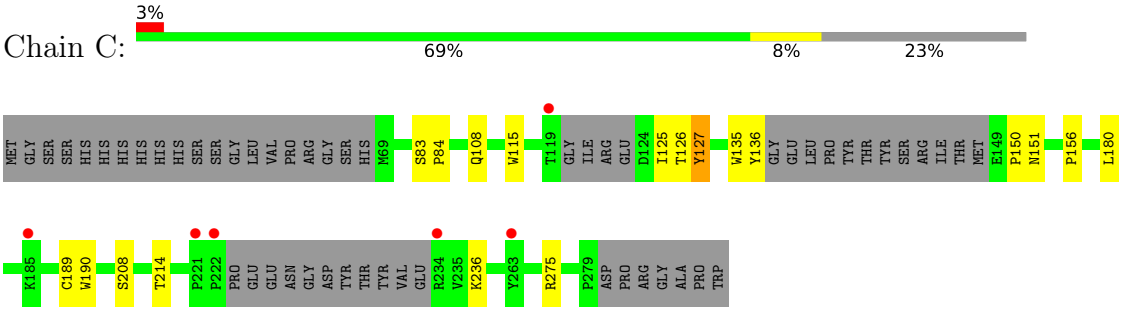
• Molecule 1: Alpha-ketoglutarate-dependent dioxygenase alkB homolog 3

Chain A:



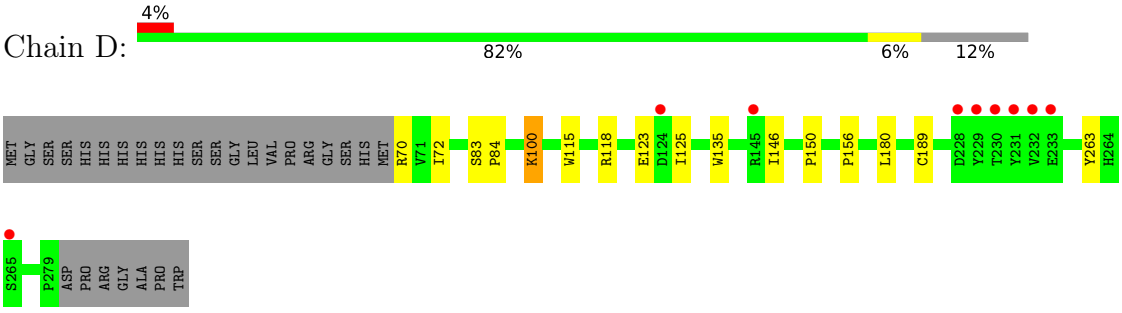
• Molecule 1: Alpha-ketoglutarate-dependent dioxygenase alkB homolog 3

Chain C:



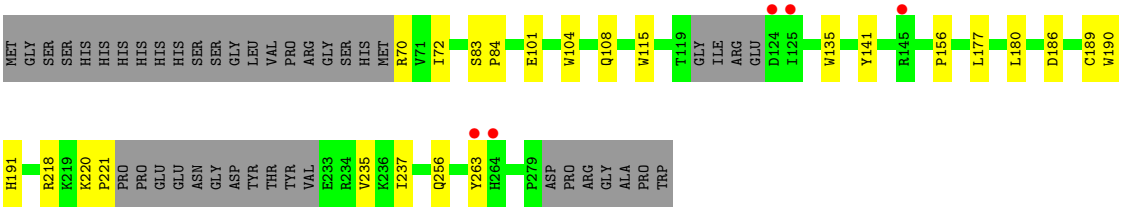
• Molecule 1: Alpha-ketoglutarate-dependent dioxygenase alkB homolog 3

Chain D:



• Molecule 1: Alpha-ketoglutarate-dependent dioxygenase alkB homolog 3

Chain E:

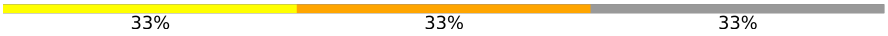


• Molecule 2: DNA

Chain B:

• Molecule 2: DNA

Chain F:



• Molecule 3: Synthetic antibody heavy chain

Chain G:

• Molecule 3: Synthetic antibody heavy chain

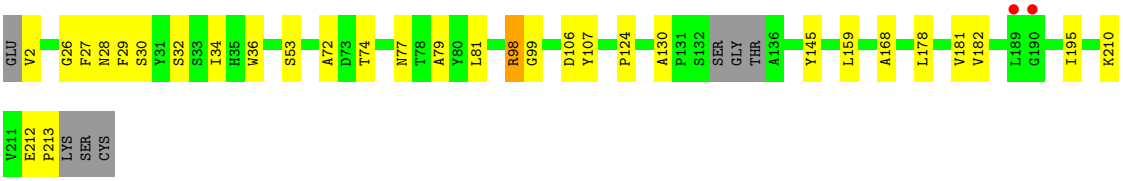
Chain H:

• Molecule 3: Synthetic antibody heavy chain

Chain I:

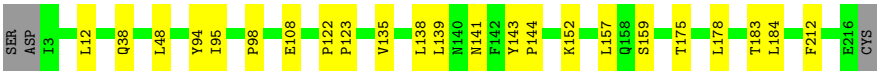
• Molecule 3: Synthetic antibody heavy chain

Chain J:



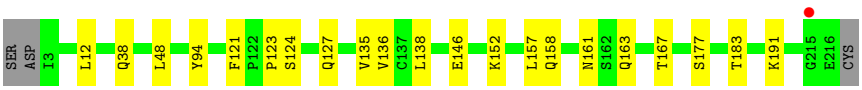
• Molecule 4: Synthetic antibody light chain

Chain K:



• Molecule 4: Synthetic antibody light chain

Chain L:



• Molecule 4: Synthetic antibody light chain

Chain M:

• Molecule 4: Synthetic antibody light chain

Chain N:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	108.61Å 142.17Å 130.11Å 90.00° 113.07° 90.00°	Depositor
Resolution (Å)	34.38 – 3.66 34.38 – 3.66	Depositor EDS
% Data completeness (in resolution range)	88.5 (34.38-3.66) 88.5 (34.38-3.66)	Depositor EDS
R_{merge}	0.24	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.23 (at 3.66Å)	Xtriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, R_{free}	0.227 , 0.273 0.230 , 0.265	Depositor DCC
R_{free} test set	796 reflections (2.24%)	wwPDB-VP
Wilson B-factor (Å^2)	80.7	Xtriage
Anisotropy	0.224	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.26 , 18.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	0.019 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.88	EDS
Total number of atoms	19651	wwPDB-VP
Average B, all atoms (Å^2)	84.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.36% 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: MN, OGA, ME6

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	0.39	0/1805	0.60	0/2459
1	C	0.35	0/1562	0.58	0/2123
1	D	0.37	0/1786	0.58	0/2434
1	E	0.36	0/1658	0.57	0/2255
2	B	0.45	0/42	0.82	0/60
2	F	0.47	0/21	0.67	0/30
3	G	0.39	0/1643	0.58	1/2242 (0.0%)
3	H	0.43	0/1619	0.60	1/2208 (0.0%)
3	I	0.46	0/1613	0.64	1/2201 (0.0%)
3	J	0.38	0/1607	0.56	0/2193
4	K	0.37	0/1680	0.54	0/2283
4	L	1.13	2/1680 (0.1%)	0.55	1/2283 (0.0%)
4	M	0.37	0/1680	0.54	0/2283
4	N	0.36	0/1680	0.54	0/2283
All	All	0.49	2/20076 (0.0%)	0.57	4/27337 (0.0%)

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
3	G	0	1
4	M	0	1
4	N	0	1
All	All	0	3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	L	146	GLU	CD-OE1	32.17	1.61	1.25

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	L	146	GLU	CD-OE2	28.46	1.56	1.25

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	G	98	ARG	CB-CA-C	-7.00	96.40	110.40
3	H	98	ARG	CB-CA-C	-6.39	97.62	110.40
4	L	146	GLU	OE1-CD-OE2	-5.87	116.26	123.30
3	I	99	GLY	N-CA-C	5.10	125.85	113.10

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	G	98	ARG	Sidechain
4	M	25	ARG	Sidechain
4	N	25	ARG	Sidechain

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	1750	0	1697	26	0
1	C	1520	0	1485	9	0
1	D	1735	0	1681	12	0
1	E	1613	0	1573	20	0
2	B	60	0	38	6	0
2	F	40	0	27	2	0
3	G	1601	0	1548	23	0
3	H	1578	0	1527	24	0
3	I	1572	0	1515	48	0
3	J	1566	0	1510	32	0
4	K	1643	0	1603	24	0
4	L	1643	0	1603	17	0
4	M	1643	0	1603	15	0
4	N	1643	0	1603	12	0
5	A	10	0	3	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	D	10	0	3	0	0
6	A	1	0	0	0	0
6	C	1	0	0	0	0
6	D	1	0	0	0	0
6	E	1	0	0	0	0
7	D	20	0	15	3	0
All	All	19651	0	19034	236	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 6.

The worst 5 of 236 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)
3:I:98:ARG:HB2	3:I:107:TYR:HB2	1.33	1.10
1:A:108:GLN:HE22	1:A:157:VAL:HG21	1.12	1.07
3:I:98:ARG:HB3	3:I:107:TYR:H	1.13	1.07
3:I:98:ARG:HG2	3:I:107:TYR:CD1	1.91	1.06
2:F:5:ME6:H2'A	2:F:6:DT:OP2	1.53	1.01

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	210/238 (88%)	204 (97%)	6 (3%)	0	100	100
1	C	176/238 (74%)	174 (99%)	2 (1%)	0	100	100
1	D	208/238 (87%)	202 (97%)	6 (3%)	0	100	100
1	E	189/238 (79%)	189 (100%)	0	0	100	100
3	G	212/216 (98%)	200 (94%)	11 (5%)	1 (0%)	29	66

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
3	H	206/216 (95%)	196 (95%)	10 (5%)	0	100 100
3	I	206/216 (95%)	193 (94%)	12 (6%)	1 (0%)	29 66
3	J	205/216 (95%)	198 (97%)	7 (3%)	0	100 100
4	K	212/217 (98%)	203 (96%)	9 (4%)	0	100 100
4	L	212/217 (98%)	203 (96%)	9 (4%)	0	100 100
4	M	212/217 (98%)	203 (96%)	9 (4%)	0	100 100
4	N	212/217 (98%)	203 (96%)	9 (4%)	0	100 100
All	All	2460/2684 (92%)	2368 (96%)	90 (4%)	2 (0%)	51 82

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	G	98	ARG
3	I	30	SER

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	195/216 (90%)	190 (97%)	5 (3%)	46 68
1	C	170/216 (79%)	165 (97%)	5 (3%)	42 66
1	D	193/216 (89%)	192 (100%)	1 (0%)	88 94
1	E	180/216 (83%)	178 (99%)	2 (1%)	73 85
3	G	177/179 (99%)	173 (98%)	4 (2%)	50 71
3	H	174/179 (97%)	171 (98%)	3 (2%)	60 79
3	I	174/179 (97%)	165 (95%)	9 (5%)	23 55
3	J	173/179 (97%)	170 (98%)	3 (2%)	60 79
4	K	189/192 (98%)	187 (99%)	2 (1%)	73 85
4	L	189/192 (98%)	189 (100%)	0	100 100
4	M	189/192 (98%)	187 (99%)	2 (1%)	73 85

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
4	N	189/192 (98%)	187 (99%)	2 (1%)	73 85
All	All	2192/2348 (93%)	2154 (98%)	38 (2%)	60 79

5 of 38 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	I	193	THR
4	M	108	GLU
3	J	27	PHE
4	K	108	GLU
4	N	135	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
4	M	150	GLN
4	N	201	HIS
3	J	164	HIS
4	K	141	ASN
4	K	150	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

2 non-standard protein/DNA/RNA residues 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	ME6	B	5	2	17,21,22	4.17	11 (64%)	21,30,33	1.39	3 (14%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
2	ME6	F	5	2	17,21,22	3.87	10 (58%)	21,30,33	1.06	1 (4%)

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	ME6	B	5	2	-	2/7/21/22	0/2/2/2
2	ME6	F	5	2	-	3/7/21/22	0/2/2/2

The worst 5 of 21 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	5	ME6	O4'-C4'	8.03	1.62	1.45
2	F	5	ME6	O4'-C4'	7.02	1.60	1.45
2	B	5	ME6	C3'-C4'	-6.72	1.34	1.53
2	B	5	ME6	C4-N4	6.45	1.48	1.33
2	B	5	ME6	C6-C5	6.27	1.49	1.35

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	5	ME6	O4'-C1'-N1	3.26	113.69	107.86
2	B	5	ME6	O4'-C1'-C2'	-3.16	100.28	106.25
2	F	5	ME6	C2'-C3'-C4'	2.67	108.32	102.76
2	B	5	ME6	C6-C5-C4	2.44	121.43	117.96

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	5	ME6	C4'-C5'-O5'-P
2	F	5	ME6	O4'-C4'-C5'-O5'
2	F	5	ME6	C3'-C4'-C5'-O5'
2	F	5	ME6	C4'-C5'-O5'-P
2	B	5	ME6	C3'-C4'-C5'-O5'

There are no ring outliers.

2 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	5	ME6	3	0
2	F	5	ME6	2	0

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 7 ligands modelled in this entry, 4 are monoatomic - leaving 3 for Mogul analysis.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
5	OGA	A	1001	6	9,9,9	2.61	2 (22%)	10,11,11	1.36	1 (10%)
5	OGA	D	1001	6	9,9,9	2.62	2 (22%)	10,11,11	1.45	2 (20%)
7	ME6	D	1003	-	17,21,22	3.93	11 (64%)	21,30,33	1.34	5 (23%)

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
5	OGA	A	1001	6	-	5/8/9/9	-
5	OGA	D	1001	6	-	0/8/9/9	-
7	ME6	D	1003	-	-	1/7/21/22	0/2/2/2

The worst 5 of 15 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	D	1003	ME6	O4'-C4'	7.95	1.62	1.45
5	D	1001	OGA	C2-N1	6.84	1.45	1.33
5	A	1001	OGA	C2-N1	6.74	1.45	1.33

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	D	1003	ME6	C3'-C4'	-6.09	1.36	1.53
7	D	1003	ME6	C4-N4	6.07	1.47	1.33

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
7	D	1003	ME6	O4'-C1'-N1	2.76	112.79	107.86
7	D	1003	ME6	C5-C4-N4	-2.54	116.51	120.85
7	D	1003	ME6	C1'-N1-C2	2.47	122.26	116.89
5	D	1001	OGA	O2-C1-C2	2.47	120.42	113.15
5	A	1001	OGA	O3-C5-C4	2.22	120.47	112.74

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

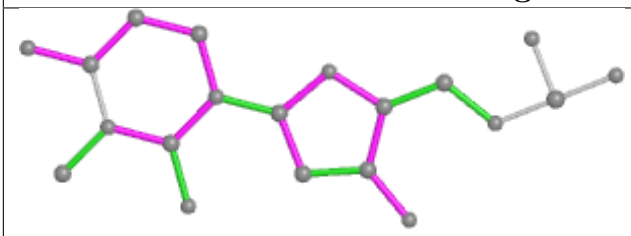
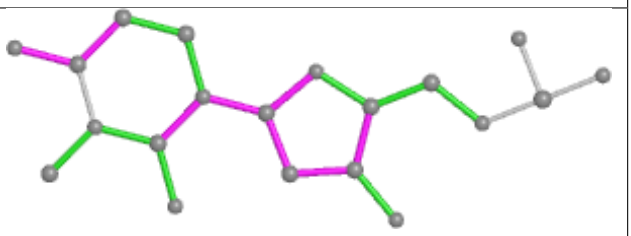
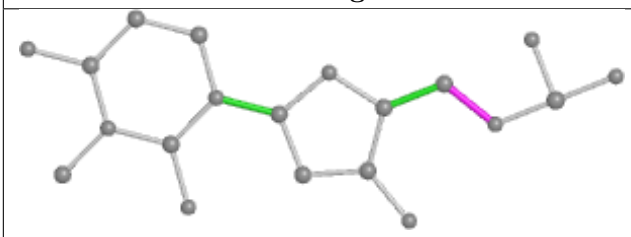
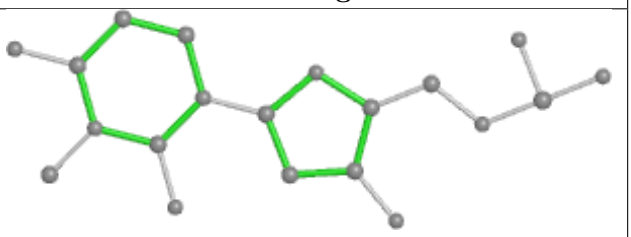
Mol	Chain	Res	Type	Atoms
5	A	1001	OGA	O2-C1-C2-N1
5	A	1001	OGA	N1-C4-C5-O4
5	A	1001	OGA	O1-C1-C2-O2'
5	A	1001	OGA	N1-C4-C5-O3
5	A	1001	OGA	O1-C1-C2-N1

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	D	1003	ME6	3	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 ME6 D 1003	
 Bond lengths	 Bond angles
 Torsions	 Rings

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

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	211/238 (88%)	-0.16	2 (0%) 84 74	50, 79, 123, 135	0
1	C	184/238 (77%)	-0.03	6 (3%) 46 33	73, 106, 128, 153	0
1	D	210/238 (88%)	-0.07	9 (4%) 35 24	50, 74, 121, 152	0
1	E	195/238 (81%)	-0.08	5 (2%) 56 42	56, 92, 123, 157	0
2	B	2/3 (66%)	0.19	0 100 100	111, 111, 111, 133	0
2	F	1/3 (33%)	1.13	0 100 100	106, 106, 106, 106	0
3	G	214/216 (99%)	-0.39	0 100 100	42, 70, 107, 129	1 (0%)
3	H	210/216 (97%)	-0.33	1 (0%) 91 85	48, 78, 116, 139	0
3	I	210/216 (97%)	-0.33	0 100 100	53, 79, 109, 128	0
3	J	209/216 (96%)	-0.18	2 (0%) 82 72	72, 90, 122, 153	0
4	K	214/217 (98%)	-0.39	0 100 100	47, 74, 94, 105	0
4	L	214/217 (98%)	-0.34	1 (0%) 91 85	44, 82, 113, 141	0
4	M	214/217 (98%)	-0.41	0 100 100	53, 76, 94, 104	0
4	N	214/217 (98%)	-0.30	0 100 100	66, 87, 109, 116	0
All	All	2502/2690 (93%)	-0.25	26 (1%) 82 72	42, 82, 118, 157	1 (0%)

The worst 5 of 26 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	222	PRO	3.8
1	E	125	ILE	3.4
1	D	231	TYR	3.4
3	J	189	LEU	3.3
1	D	229	TYR	3.1

6.2 Non-standard residues in protein, DNA, RNA chains (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
2	ME6	F	5	20/21	0.90	0.21	79,108,122,123	0
2	ME6	B	5	20/21	0.93	0.20	79,93,118,124	0

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

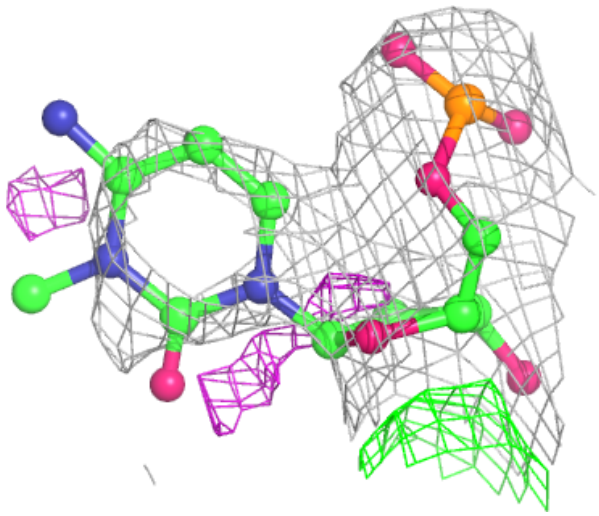
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

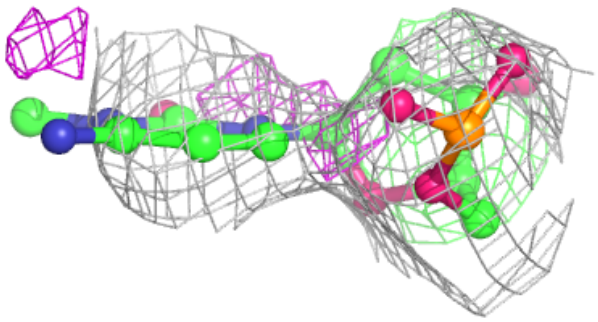
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
7	ME6	D	1003	20/21	0.80	0.33	95,110,123,124	0
6	MN	C	1001	1/1	0.95	0.31	60,60,60,60	0
5	OGA	A	1001	10/10	0.96	0.32	58,73,82,84	0
5	OGA	D	1001	10/10	0.96	0.28	60,68,80,87	0
6	MN	E	1001	1/1	0.97	0.29	60,60,60,60	0
6	MN	A	1002	1/1	0.98	0.28	48,48,48,48	0
6	MN	D	1002	1/1	0.98	0.24	46,46,46,46	0

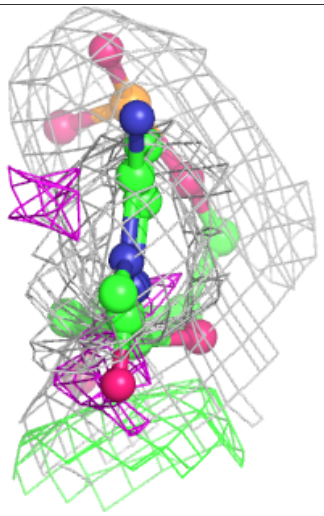
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around ME6 D 1003:

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







6.5 Other polymers

There are no such residues in this entry.