wwPDB X-ray Structure Validation Summary Report (i)

Aug 21, 2023 – 04:22 pm BST

PDB ID	:	7Z6P
Title	:	Diels-Alderase AbyU mutant - Y76F
Authors	:	Tiwari, K.; Burton, N.M.; Yang, S.; Race, P.R.
Deposited on	:	2022-03-14
Resolution	:	1.56 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:

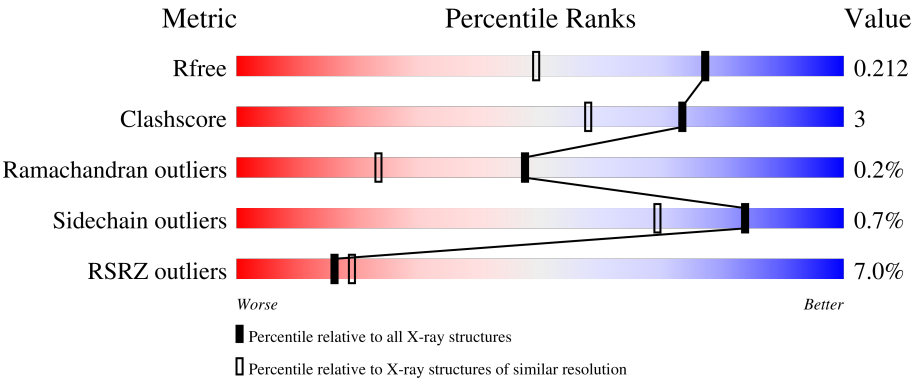
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 1.56 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	1483 (1.56-1.56)
Clashscore	141614	1529 (1.56-1.56)
Ramachandran outliers	138981	1498 (1.56-1.56)
Sidechain outliers	138945	1495 (1.56-1.56)
RSRZ outliers	127900	1465 (1.56-1.56)

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	AAA	141	6% 87% 6% • 6%
1	BBB	141	6% 89% 6% 6%
1	CCC	141	% 89% 6% 6%
1	DDD	141	13% 82% 11% • 6%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 8874 atoms, of which 4162 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.

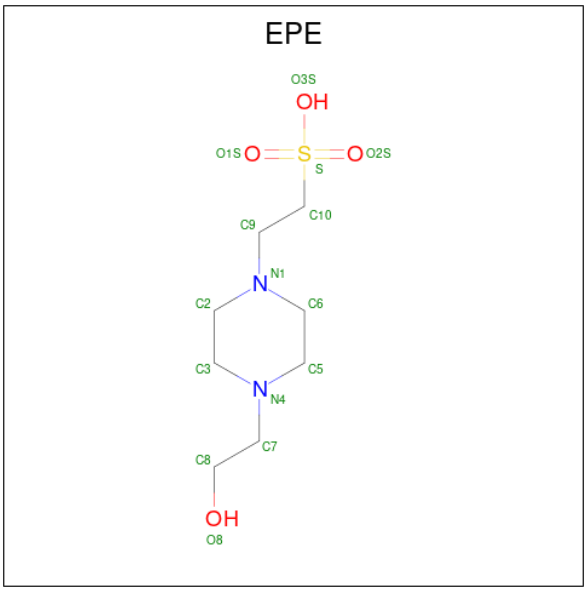
- Molecule 1 is a protein called YD repeat-containing protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	AAA	133	Total	C	H	N	O	S	38	2	0
			2064	665	1022	175	201	1			
1	BBB	133	Total	C	H	N	O	S	36	1	0
			2054	662	1019	174	198	1			
1	CCC	133	Total	C	H	N	O	S	36	3	0
			2072	668	1026	174	202	2			
1	DDD	133	Total	C	H	N	O	S	36	0	0
			2040	657	1010	174	198	1			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	76	PHE	TYR	engineered mutation	UNP F4F7G1
BBB	76	PHE	TYR	engineered mutation	UNP F4F7G1
CCC	76	PHE	TYR	engineered mutation	UNP F4F7G1
DDD	76	PHE	TYR	engineered mutation	UNP F4F7G1

- Molecule 2 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (three-letter code: EPE) (formula: $C_8H_{18}N_2O_4S$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
2	AAA	1	Total	C	H	N	O	S	2	1
			64	16	34	4	8	2		
2	BBB	1	Total	C	H	N	O	S	1	0
			32	8	17	2	4	1		
2	CCC	1	Total	C	H	N	O	S	1	0
			32	8	17	2	4	1		
2	DDD	1	Total	C	H	N	O	S	1	0
			32	8	17	2	4	1		

- Molecule 3 is water.

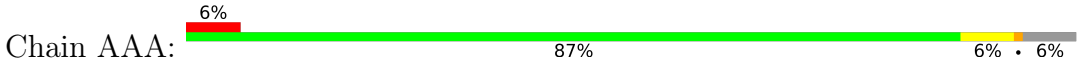
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	AAA	135	Total	O	0	0
			135	135		
3	BBB	108	Total	O	0	0
			108	108		
3	CCC	138	Total	O	0	0
			138	138		
3	DDD	103	Total	O	0	0
			103	103		

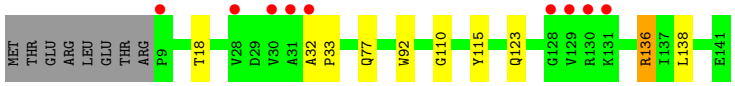
3 Residue-property plots (i)

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: YD repeat-containing protein

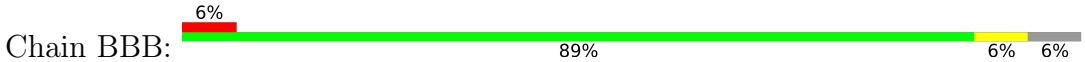
Chain AAA:

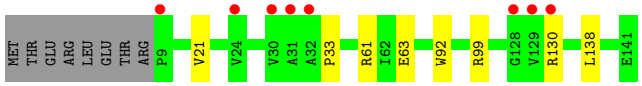




• Molecule 1: YD repeat-containing protein

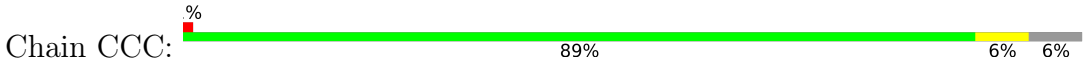
Chain BBB:

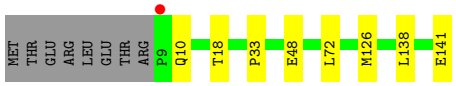




• Molecule 1: YD repeat-containing protein

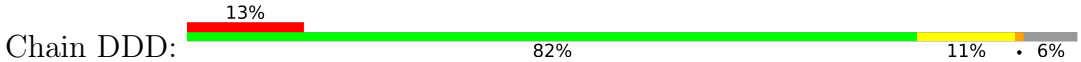
Chain CCC:

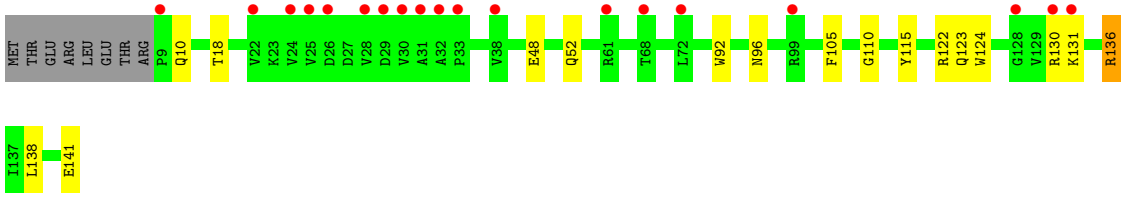




• Molecule 1: YD repeat-containing protein

Chain DDD:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	49.71Å 53.82Å 60.48Å 90.11° 96.98° 94.66°	Depositor
Resolution (Å)	40.58 – 1.56 40.54 – 1.56	Depositor EDS
% Data completeness (in resolution range)	95.5 (40.58-1.56) 95.5 (40.54-1.56)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.84 (at 1.56Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.176 , 0.209 0.188 , 0.212	Depositor DCC
R_{free} test set	4233 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	28.3	Xtriage
Anisotropy	0.564	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.44 , 54.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.98	EDS
Total number of atoms	8874	wwPDB-VP
Average B, all atoms (Å^2)	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.08% 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: EPE

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	AAA	0.72	0/1070	0.88	1/1453 (0.1%)
1	BBB	0.71	0/1060	0.87	0/1440
1	CCC	0.73	0/1077	0.89	0/1462
1	DDD	0.74	1/1052 (0.1%)	0.90	1/1429 (0.1%)
All	All	0.72	1/4259 (0.0%)	0.89	2/5784 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	DDD	48	GLU	CD-OE1	7.97	1.34	1.25

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	DDD	136	ARG	NE-CZ-NH1	5.87	123.24	120.30
1	AAA	136	ARG	NE-CZ-NH1	5.34	122.97	120.30

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	AAA	1042	1022	1022	6	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	BBB	1035	1019	1019	6	0
1	CCC	1046	1026	1027	6	0
1	DDD	1030	1010	1008	11	0
2	AAA	30	34	36	0	0
2	BBB	15	17	18	1	0
2	CCC	15	17	18	1	0
2	DDD	15	17	18	2	0
3	AAA	135	0	0	2	0
3	BBB	108	0	0	3	0
3	CCC	138	0	0	1	0
3	DDD	103	0	0	3	0
All	All	4712	4162	4166	28	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 3.

The worst 5 of 28 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:BBB:21:VAL:HG22	3:BBB:321:HOH:O	1.72	0.88
2:BBB:201:EPE:H91	3:BBB:321:HOH:O	1.93	0.69
1:DDD:130:ARG:HG2	1:DDD:131:LYS:HG3	1.75	0.68
1:DDD:92:TRP:O	3:DDD:302:HOH:O	2.15	0.65
1:BBB:33:PRO:HG2	1:BBB:99:ARG:HD3	1.82	0.61

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	AAA	133/141 (94%)	129 (97%)	4 (3%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	BBB	132/141 (94%)	128 (97%)	3 (2%)	1 (1%)	19 4
1	CCC	134/141 (95%)	130 (97%)	4 (3%)	0	100 100
1	DDD	131/141 (93%)	128 (98%)	3 (2%)	0	100 100
All	All	530/564 (94%)	515 (97%)	14 (3%)	1 (0%)	47 23

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	BBB	130	ARG

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	AAA	109/115 (95%)	108 (99%)	1 (1%)	78 61
1	BBB	108/115 (94%)	108 (100%)	0	100 100
1	CCC	110/115 (96%)	109 (99%)	1 (1%)	78 61
1	DDD	107/115 (93%)	106 (99%)	1 (1%)	78 61
All	All	434/460 (94%)	431 (99%)	3 (1%)	84 69

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	18	THR
1	CCC	18	THR
1	DDD	18	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

5 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	EPE	AAA	201[A]	-	15,15,15	2.09	1 (6%)	18,20,20	1.59	2 (11%)
2	EPE	CCC	201	-	15,15,15	1.90	1 (6%)	18,20,20	1.44	2 (11%)
2	EPE	DDD	201	-	15,15,15	1.75	1 (6%)	18,20,20	1.74	3 (16%)
2	EPE	AAA	201[B]	-	15,15,15	1.84	1 (6%)	18,20,20	1.56	5 (27%)
2	EPE	BBB	201	-	15,15,15	1.89	1 (6%)	18,20,20	1.93	2 (11%)

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	EPE	AAA	201[A]	-	-	4/9/19/19	0/1/1/1
2	EPE	CCC	201	-	-	4/9/19/19	0/1/1/1
2	EPE	DDD	201	-	-	6/9/19/19	0/1/1/1
2	EPE	AAA	201[B]	-	-	3/9/19/19	0/1/1/1
2	EPE	BBB	201	-	-	3/9/19/19	0/1/1/1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	AAA	201[A]	EPE	C10-S	-7.65	1.66	1.77
2	BBB	201	EPE	C10-S	-7.04	1.67	1.77
2	CCC	201	EPE	C10-S	-6.98	1.67	1.77
2	AAA	201[B]	EPE	C10-S	-6.85	1.67	1.77
2	DDD	201	EPE	C10-S	-6.37	1.68	1.77

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	BBB	201	EPE	O3S-S-C10	6.58	116.41	105.77
2	AAA	201[A]	EPE	O2S-S-C10	4.50	112.33	106.92
2	DDD	201	EPE	O3S-S-C10	3.99	112.23	105.77
2	DDD	201	EPE	C6-C5-N4	3.58	117.98	110.64
2	AAA	201[B]	EPE	O3S-S-C10	3.50	111.44	105.77

There are no chirality outliers.

5 of 20 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	AAA	201[A]	EPE	S-C10-C9-N1
2	AAA	201[A]	EPE	C9-C10-S-O1S
2	CCC	201	EPE	S-C10-C9-N1
2	DDD	201	EPE	S-C10-C9-N1
2	DDD	201	EPE	N4-C7-C8-O8

There are no ring outliers.

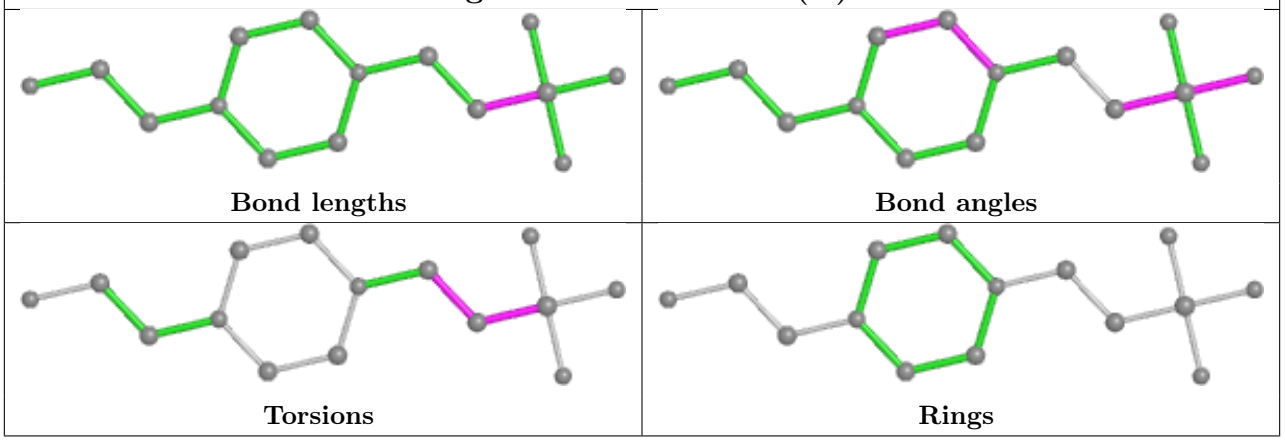
3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	CCC	201	EPE	1	0
2	DDD	201	EPE	2	0
2	BBB	201	EPE	1	0

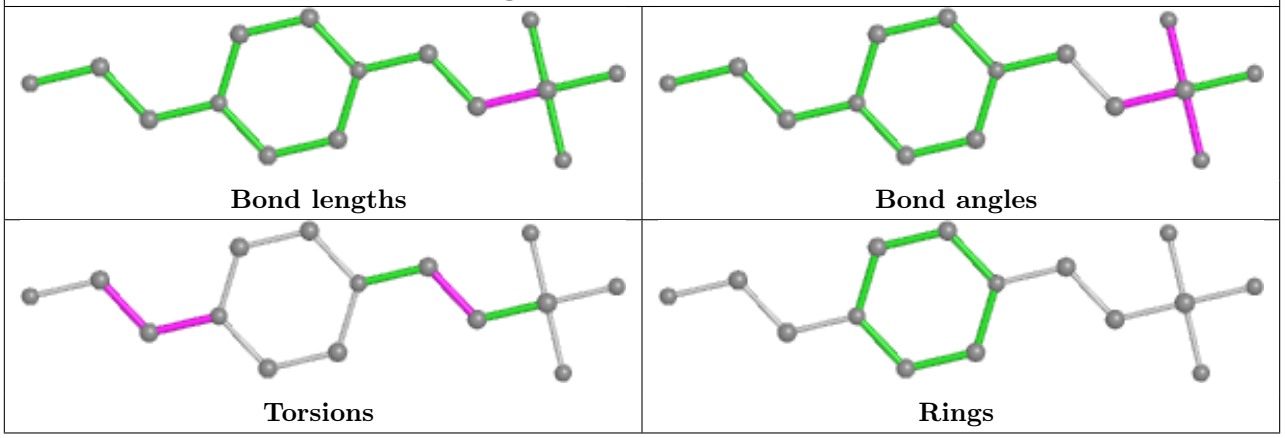
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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

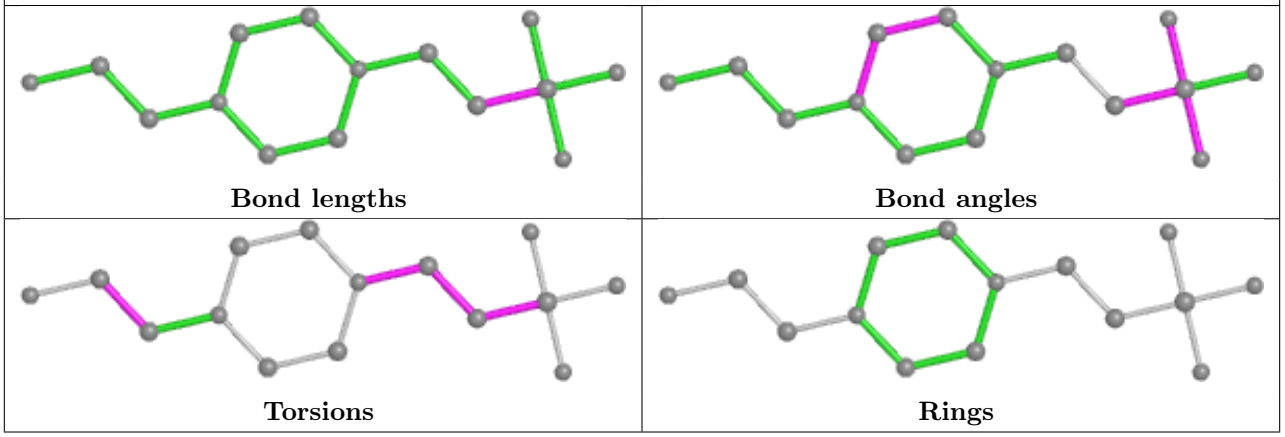
Ligand EPE AAA 201 (A)

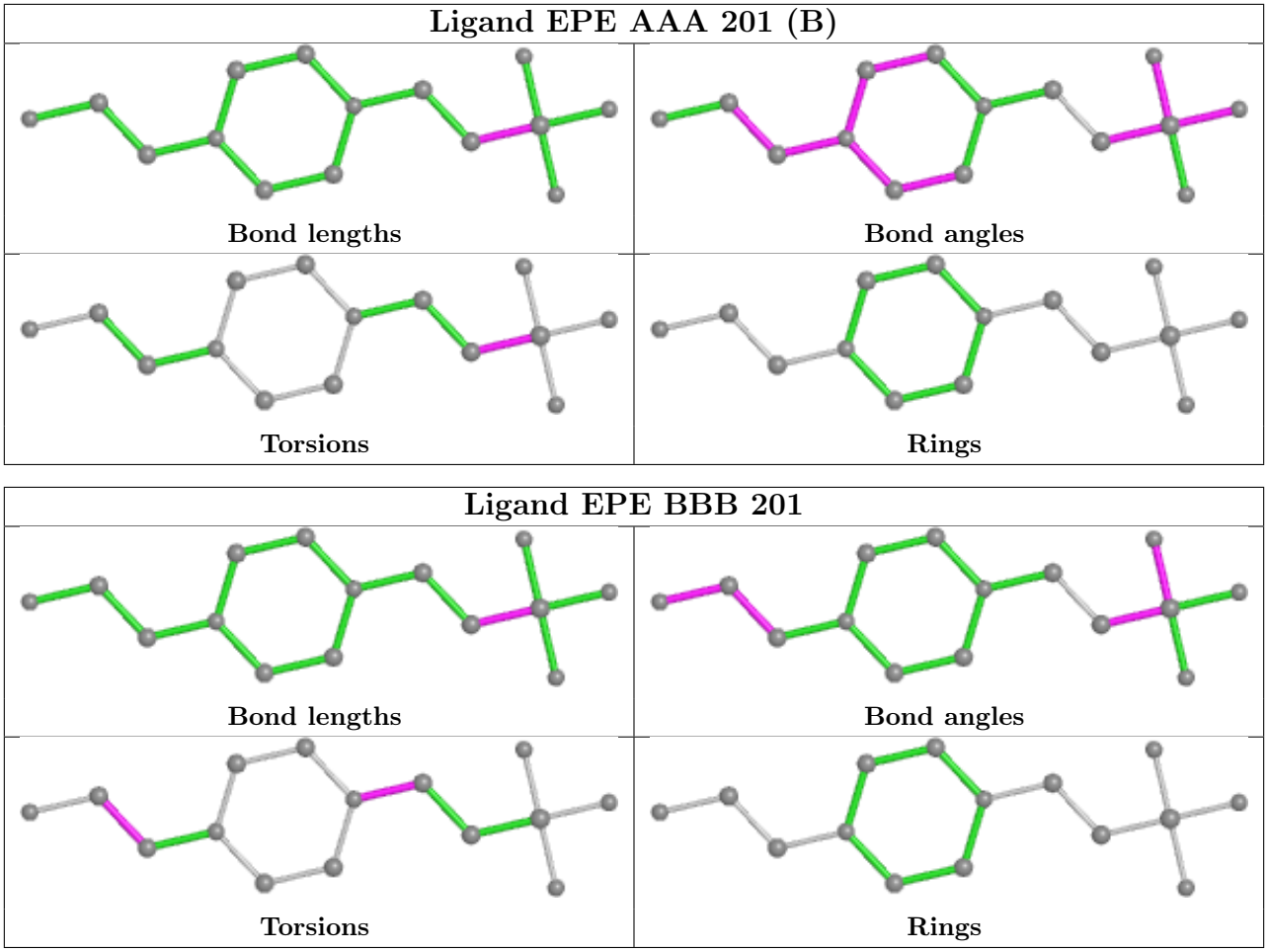


Ligand EPE CCC 201



Ligand EPE DDD 201





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	AAA	133/141 (94%)	0.41	9 (6%) 17 20	22, 34, 70, 104	0
1	BBB	133/141 (94%)	0.28	8 (6%) 21 25	22, 35, 66, 94	0
1	CCC	133/141 (94%)	0.15	1 (0%) 86 89	21, 34, 52, 81	0
1	DDD	133/141 (94%)	0.77	19 (14%) 2 2	22, 37, 76, 114	0
All	All	532/564 (94%)	0.40	37 (6%) 16 19	21, 35, 70, 114	0

The worst 5 of 37 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	DDD	31	ALA	13.4
1	DDD	30	VAL	12.7
1	DDD	32	ALA	8.5
1	AAA	31	ALA	6.7
1	AAA	130	ARG	6.5

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

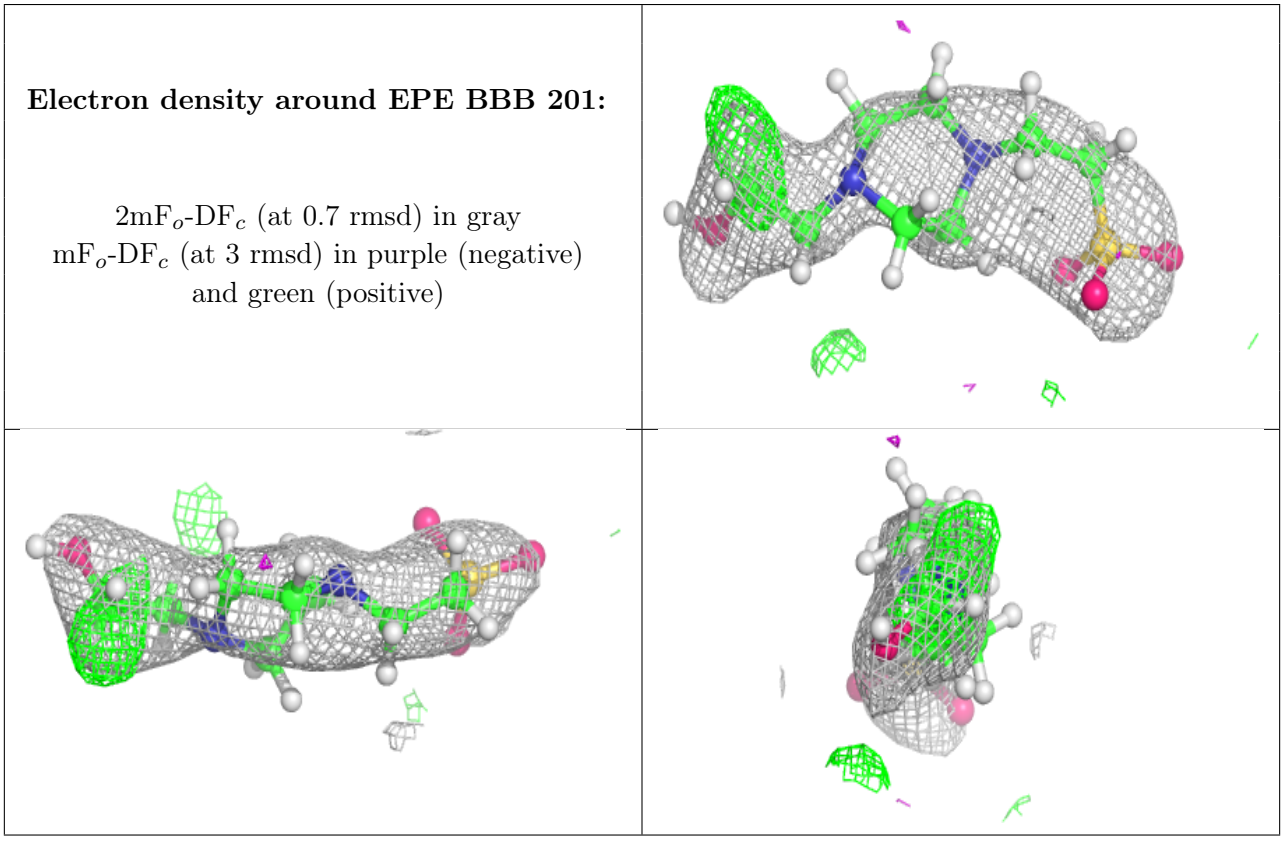
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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	EPE	BBB	201	15/15	0.83	0.20	47,71,98,109	1
2	EPE	CCC	201	15/15	0.84	0.21	72,83,124,136	1
2	EPE	DDD	201	15/15	0.89	0.22	53,65,78,93	1
2	EPE	AAA	201[B]	15/15	0.95	0.24	34,38,44,46	32
2	EPE	AAA	201[A]	15/15	0.95	0.24	58,70,71,77	32

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.

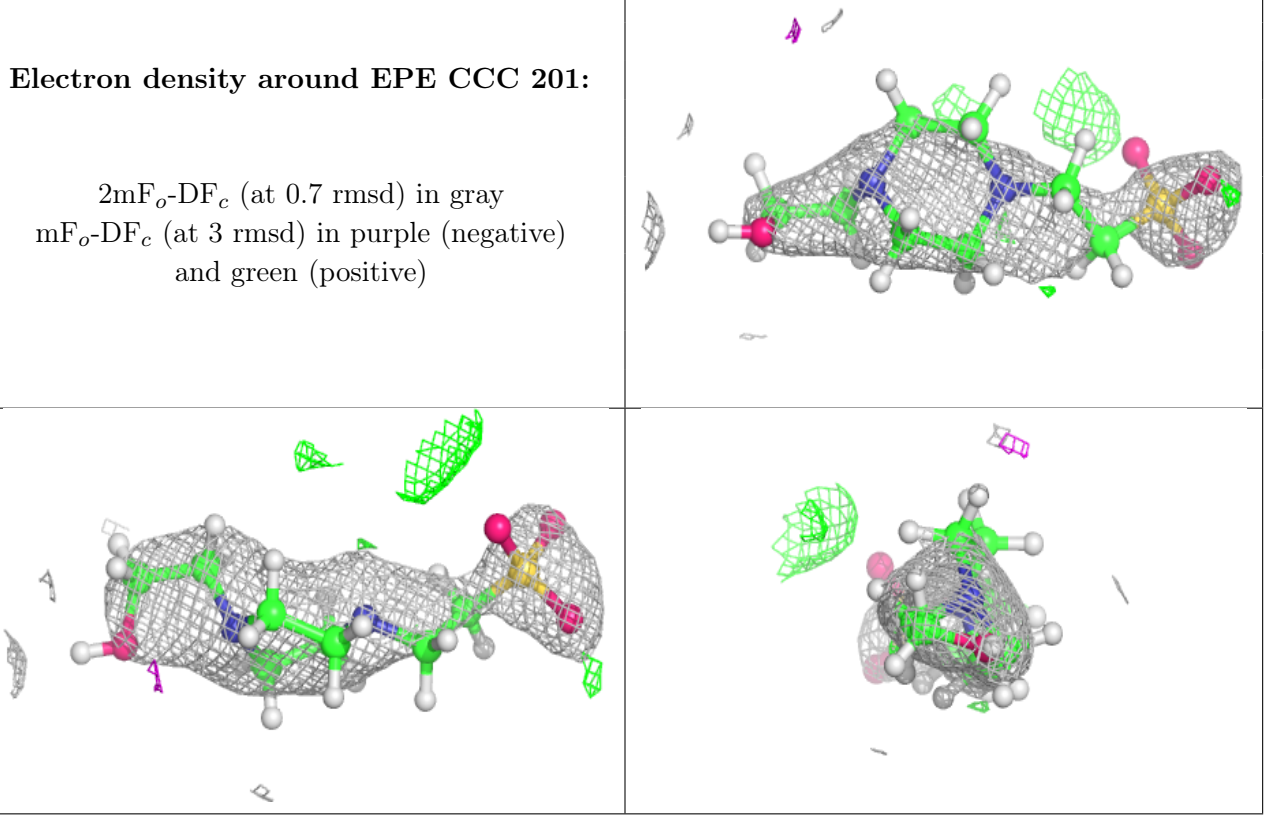
Electron density around EPE BBB 201:

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



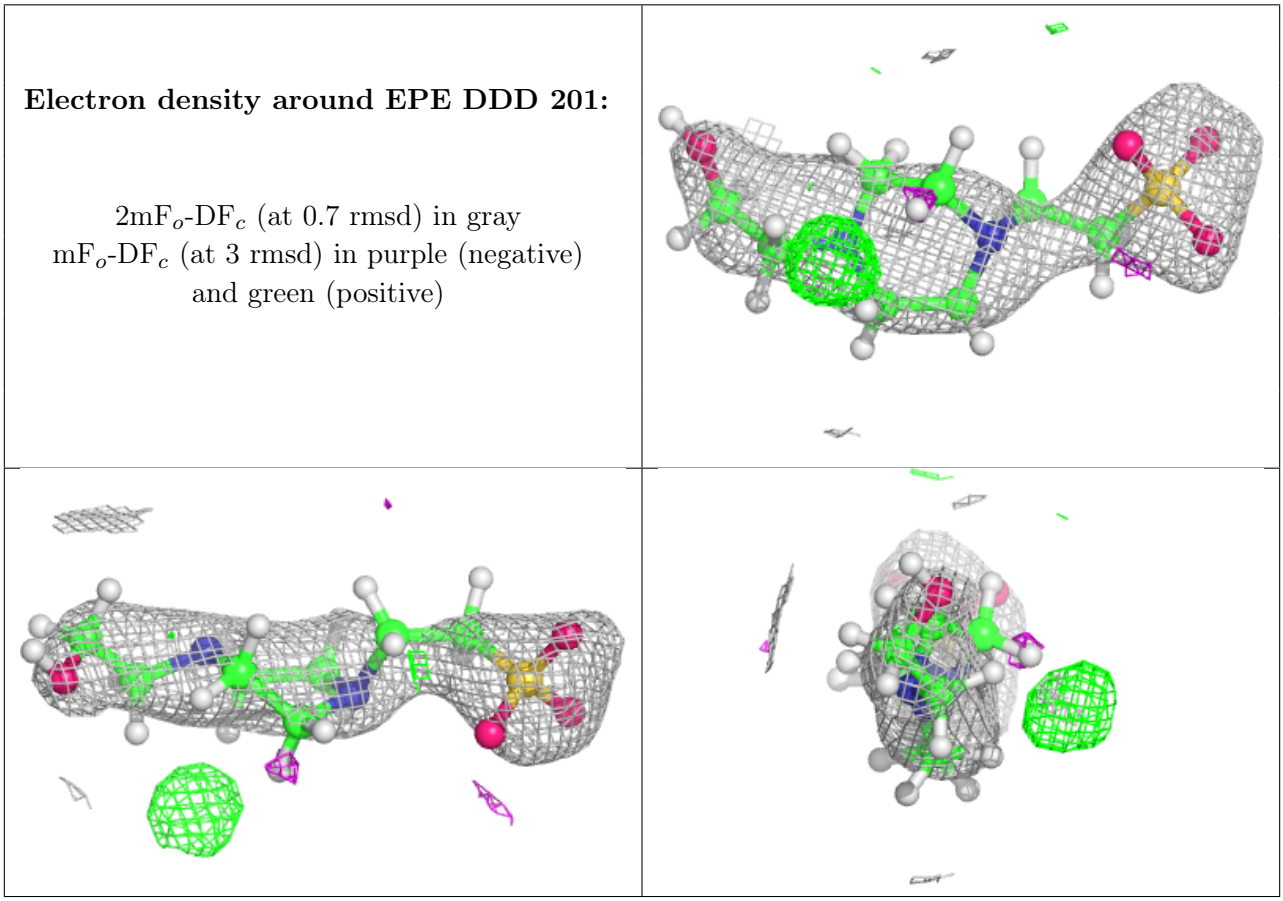
Electron density around EPE CCC 201:

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



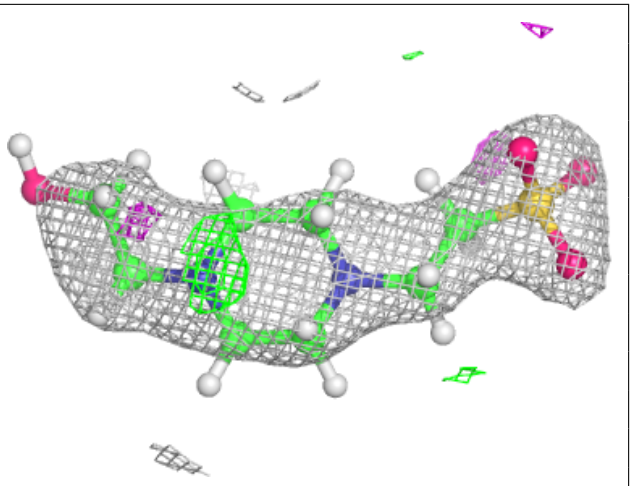
Electron density around EPE DDD 201:

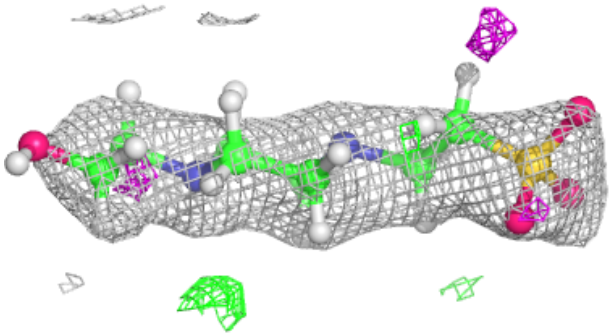
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

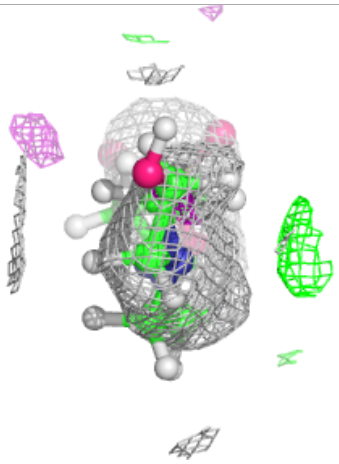


Electron density around EPE AAA 201 (B):

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

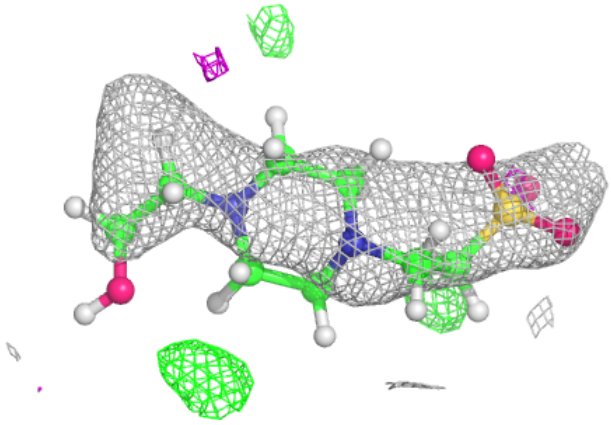


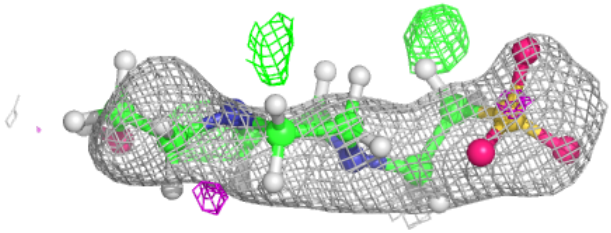


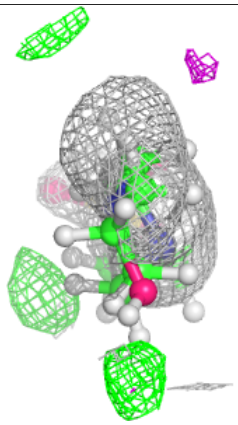


Electron density around EPE AAA 201 (A):

$2mF_o$-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.