Full wwPDB X-ray Structure Validation Report ⓘ

Aug 7, 2020 – 06:33 PM BST

PDB ID	:	6XY9
Title	:	Crystal structure of haloalkane dehalogenase DbeA-M1 loop variant from Bradyrhizobium elkanii
Authors	:	Prudnikova, T.; Rezacova, P.; Kuta Smatanova, I.; Chaloupkova, R.; Damborsky, J.
Deposited on	:	2020-01-29
Resolution	:	2.20 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the ⓘ symbol.

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

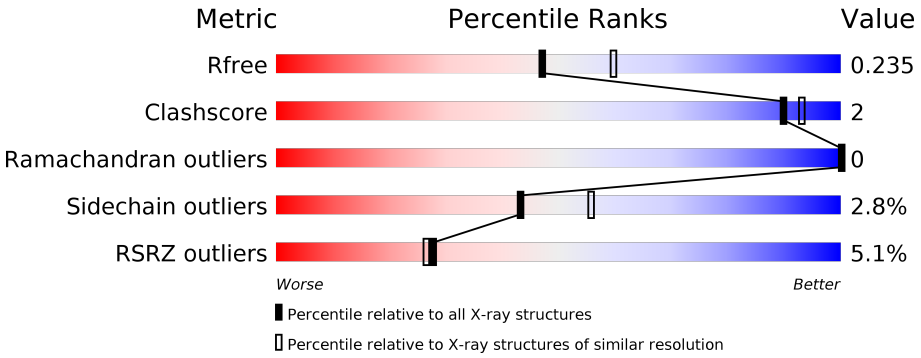
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.13.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.13.1

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 2.20 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	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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 on 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	309	5% 92% 6% ·
1	B	309	6% 90% 7% ·

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 5046 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 Haloalkane dehalogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	303	Total	C	N	O	S	0	1	0
			2341	1489	417	424	11			
1	B	301	Total	C	N	O	S	0	5	0
			2341	1491	414	426	10			

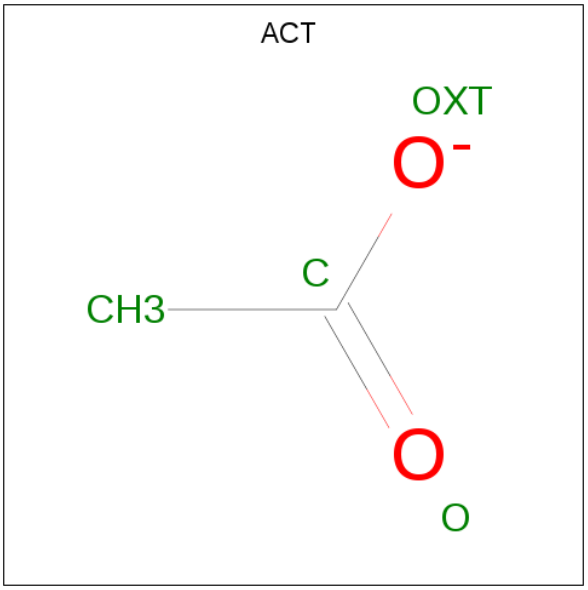
There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	143	VAL	-	insertion	UNP E2RV62
A	144	ALA	-	insertion	UNP E2RV62
A	145	GLU	-	insertion	UNP E2RV62
A	146	GLU	-	insertion	UNP E2RV62
A	147	GLN	-	insertion	UNP E2RV62
A	148	ASP	-	insertion	UNP E2RV62
A	149	HIS	-	insertion	UNP E2RV62
A	150	ALA	-	insertion	UNP E2RV62
A	151	GLU	-	insertion	UNP E2RV62
A	312	LEU	-	expression tag	UNP E2RV62
A	313	GLU	-	expression tag	UNP E2RV62
B	143	VAL	-	insertion	UNP E2RV62
B	144	ALA	-	insertion	UNP E2RV62
B	145	GLU	-	insertion	UNP E2RV62
B	146	GLU	-	insertion	UNP E2RV62
B	147	GLN	-	insertion	UNP E2RV62
B	148	ASP	-	insertion	UNP E2RV62
B	149	HIS	-	insertion	UNP E2RV62
B	150	ALA	-	insertion	UNP E2RV62
B	151	GLU	-	insertion	UNP E2RV62
B	312	LEU	-	expression tag	UNP E2RV62
B	313	GLU	-	expression tag	UNP E2RV62

- Molecule 2 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	2	Total Cl 2 2	0	0
2	A	2	Total Cl 2 2	0	0

- Molecule 3 is ACETATE ION (three-letter code: ACT) (formula: $C_2H_3O_2$) (labeled as "Ligand of Interest" by author).



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

- Molecule 4 is water.

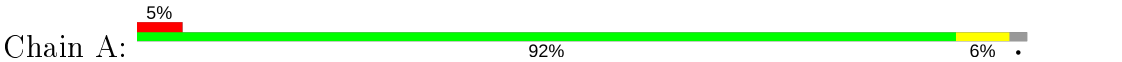
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	167	Total O 167 167	0	0
4	B	185	Total O 185 185	0	0

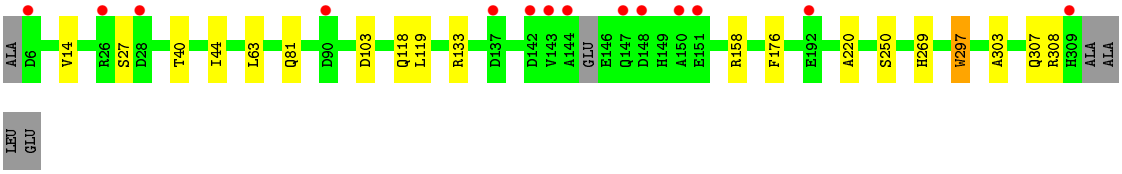
3 Residue-property plots (i)

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: Haloalkane dehalogenase

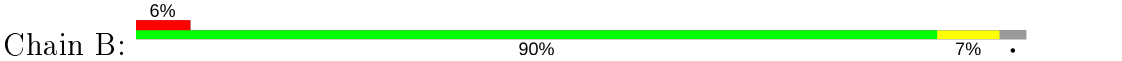
Chain A:

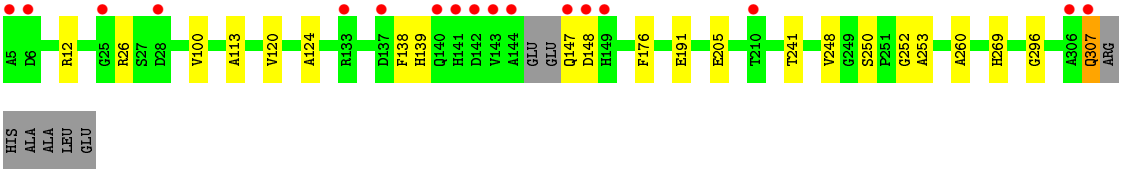




- Molecule 1: Haloalkane dehalogenase

Chain B:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	133.75Å 75.13Å 77.60Å 90.00° 91.90° 90.00°	Depositor
Resolution (Å)	19.60 – 2.20 19.39 – 2.20	Depositor EDS
% Data completeness (in resolution range)	98.9 (19.60-2.20) 99.1 (19.39-2.20)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.31 (at 2.21Å)	Xtriage
Refinement program	REFMAC 5.8.0123	Depositor
R, R_{free}	0.175 , 0.235 0.176 , 0.235	Depositor DCC
R_{free} test set	1938 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	28.2	Xtriage
Anisotropy	0.062	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 42.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.006 for -1/2*h-3/2*k,-1/2*h+1/2*k,-l 0.007 for -1/2*h+3/2*k,1/2*h+1/2*k,-l 0.016 for 1/2*h-3/2*k,-1/2*h-1/2*k,-l 0.014 for 1/2*h+3/2*k,1/2*h-1/2*k,-l 0.017 for -h,-k,l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	5046	wwPDB-VP
Average B, all atoms (Å^2)	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.12% 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: CL, ACT

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	0.91	1/2406 (0.0%)	0.92	1/3271 (0.0%)
1	B	6.45	3/2411 (0.1%)	1.29	4/3279 (0.1%)
All	All	4.61	4/4817 (0.1%)	1.12	5/6550 (0.1%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	307[A]	GLN	CD-OE1	221.53	6.11	1.24
1	B	307[B]	GLN	CD-OE1	221.53	6.11	1.24
1	B	205	GLU	CG-CD	5.33	1.59	1.51
1	A	103	ASP	N-CA	5.32	1.56	1.46

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	307[A]	GLN	CG-CD-OE1	-33.62	54.35	121.60
1	B	307[B]	GLN	CG-CD-OE1	-33.62	54.35	121.60
1	B	307[A]	GLN	OE1-CD-NE2	-14.42	88.74	121.90
1	B	307[B]	GLN	OE1-CD-NE2	-14.42	88.74	121.90
1	A	297	TRP	CA-CB-CG	5.05	123.29	113.70

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	2341	0	2288	7	0
1	B	2341	0	2293	14	0
2	A	2	0	0	0	0
2	B	2	0	0	0	0
3	A	4	0	3	0	0
3	B	4	0	3	0	0
4	A	167	0	0	0	1
4	B	185	0	0	2	2
All	All	5046	0	4587	17	2

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 2.

All (17) 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:248[A]:VAL:HG11	1:B:260:ALA:CB	2.04	0.87
1:B:139:HIS:HA	1:B:253:ALA:O	1.99	0.62
1:B:248[A]:VAL:HG11	1:B:260:ALA:HB2	1.83	0.59
1:B:12:ARG:HD3	4:B:601:HOH:O	2.02	0.58
1:B:248[A]:VAL:HG11	1:B:260:ALA:HB1	1.88	0.50
1:A:269:HIS:NE2	1:B:307[B]:GLN:HG3	2.26	0.50
1:B:191:GLU:HA	1:B:191:GLU:OE1	2.13	0.48
1:B:113:ALA:HB2	1:B:120:VAL:HG21	1.96	0.48
1:A:303:ALA:HB1	1:B:269:HIS:HB3	1.97	0.47
1:B:138:PHE:O	1:B:139:HIS:HB2	2.15	0.47
1:A:133:ARG:O	1:A:220:ALA:HA	2.16	0.45
1:A:307:GLN:OE1	1:B:241[A]:THR:HG23	2.18	0.44
1:B:252:GLY:N	4:B:603:HOH:O	2.37	0.44
1:A:40:THR:HG21	1:A:44:ILE:HD12	1.99	0.43
1:A:63:LEU:HB3	1:A:81:GLN:OE1	2.20	0.42
1:B:100:VAL:HA	1:B:124:ALA:O	2.21	0.41
1:A:297:TRP:HA	1:B:296:GLY:O	2.21	0.40

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:A:662:HOH:O	4:B:686:HOH:O[4_545]	2.15	0.05

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:B:678:HOH:O	4:B:760:HOH:O[2_656]	2.18	0.02

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	300/309 (97%)	289 (96%)	11 (4%)	0	100	100
1	B	301/309 (97%)	289 (96%)	12 (4%)	0	100	100
All	All	601/618 (97%)	578 (96%)	23 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	235/237 (99%)	227 (97%)	8 (3%)	37	47
1	B	236/237 (100%)	231 (98%)	5 (2%)	53	67
All	All	471/474 (99%)	458 (97%)	13 (3%)	43	56

All (13) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	14	VAL
1	A	27	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	118	GLN
1	A	119	LEU
1	A	158	ARG
1	A	176	PHE
1	A	250	SER
1	A	308	ARG
1	B	26	ARG
1	B	147	GLN
1	B	148	ASP
1	B	176	PHE
1	B	250	SER

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

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

Of 6 ligands modelled in this entry, 4 are monoatomic - leaving 2 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	$\#\|Z\|>2$
3	ACT	A	503	-	1,3,3	4.11	1 (100%)	0,3,3	0.00	-
3	ACT	B	503	-	1,3,3	5.64	1 (100%)	0,3,3	0.00	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	503	ACT	CH3-C	5.64	1.55	1.48
3	A	503	ACT	CH3-C	4.11	1.54	1.48

There are no bond angle outliers.

There are no chirality outliers.

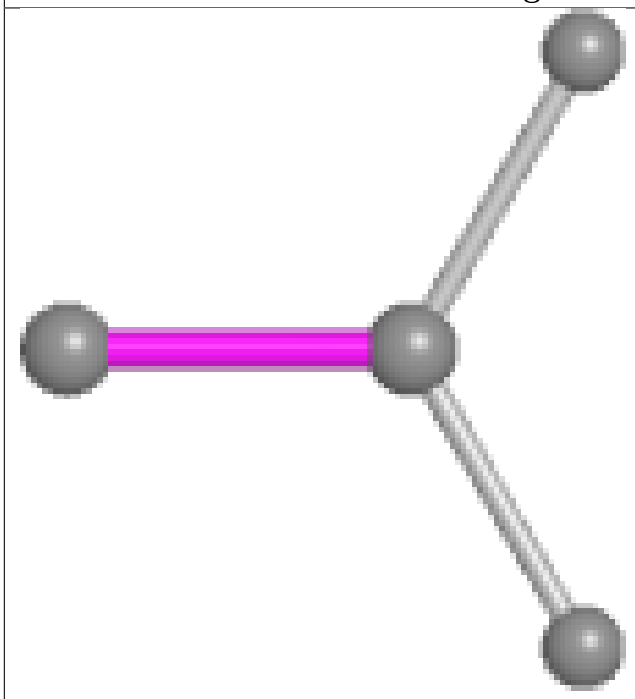
There are no torsion outliers.

There are no ring outliers.

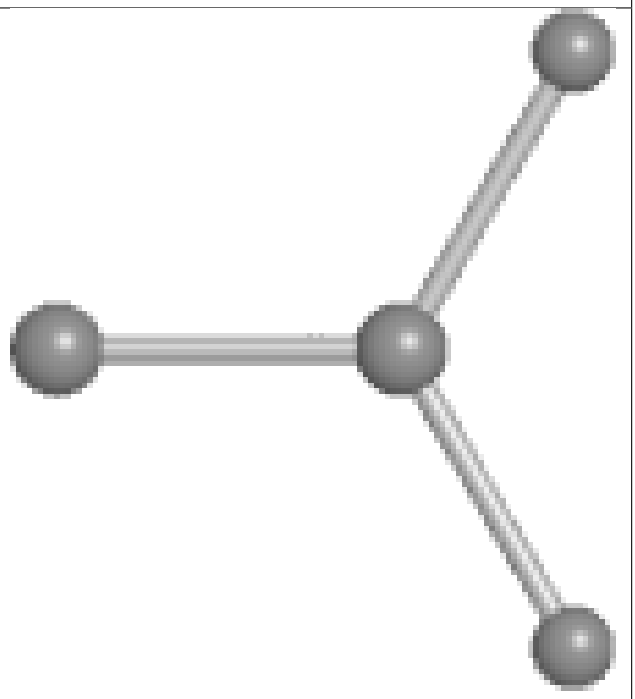
No monomer is involved in short contacts.

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.

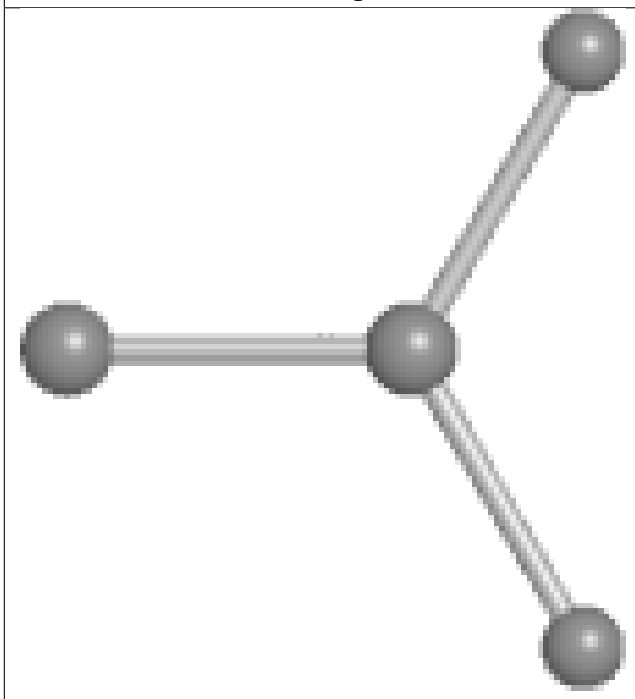
Ligand ACT A 503



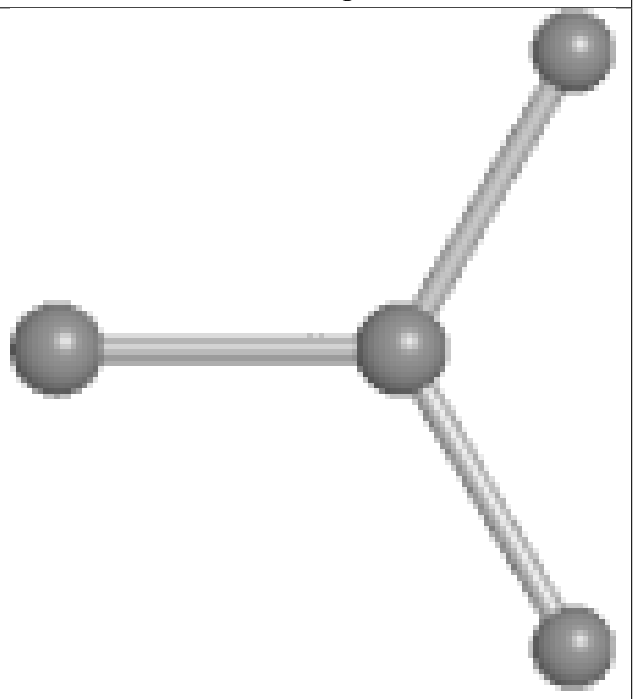
Bond lengths



Bond angles

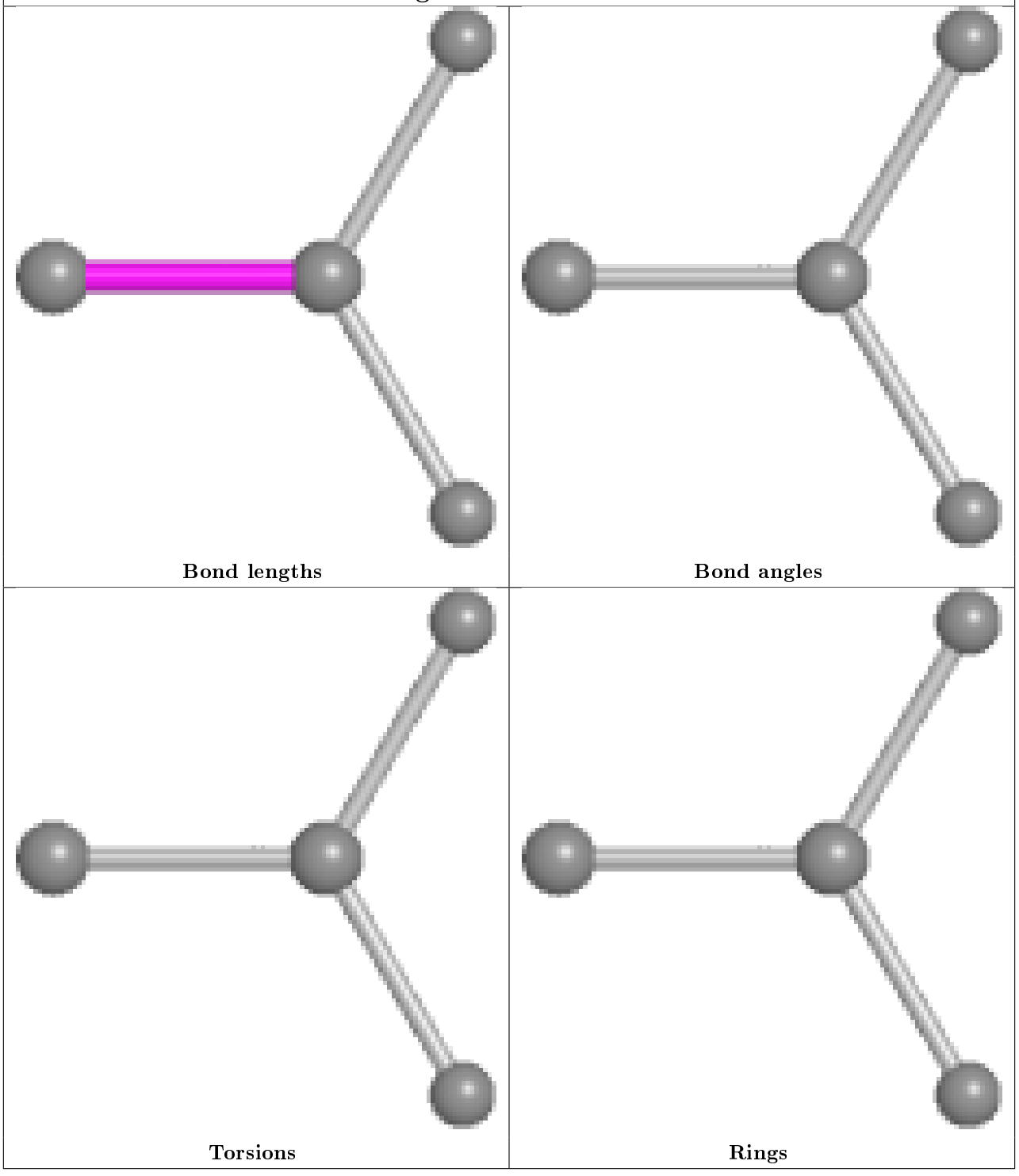


Torsions



Rings

Ligand ACT B 503



5.7 Other polymers (i)

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, 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	303/309 (98%)	-0.13	14 (4%)	32	31	19, 29, 60, 107	2 (0%)
1	B	301/309 (97%)	-0.22	17 (5%)	24	23	17, 25, 51, 113	2 (0%)
All	All	604/618 (97%)	-0.18	31 (5%)	28	26	17, 27, 60, 113	4 (0%)

All (31) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	143	VAL	7.2
1	B	5	ALA	6.5
1	A	144	ALA	5.6
1	B	307[A]	GLN	5.3
1	B	144	ALA	5.0
1	B	142	ASP	4.9
1	A	143	VAL	4.8
1	A	148	ASP	4.2
1	A	26	ARG	4.1
1	B	148	ASP	4.1
1	A	309	HIS	4.0
1	B	6	ASP	3.5
1	B	147	GLN	3.5
1	B	140	GLN	3.4
1	A	28	ASP	3.2
1	A	147	GLN	3.0
1	B	306	ALA	3.0
1	A	6	ASP	2.9
1	A	142	ASP	2.7
1	B	28	ASP	2.6
1	A	151	GLU	2.4
1	A	150	ALA	2.4
1	A	137	ASP	2.2
1	B	141	HIS	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	133	ARG	2.1
1	A	192	GLU	2.1
1	A	90	ASP	2.1
1	B	25	GLY	2.1
1	B	137	ASP	2.0
1	B	210	THR	2.0
1	B	149	HIS	2.0

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

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

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

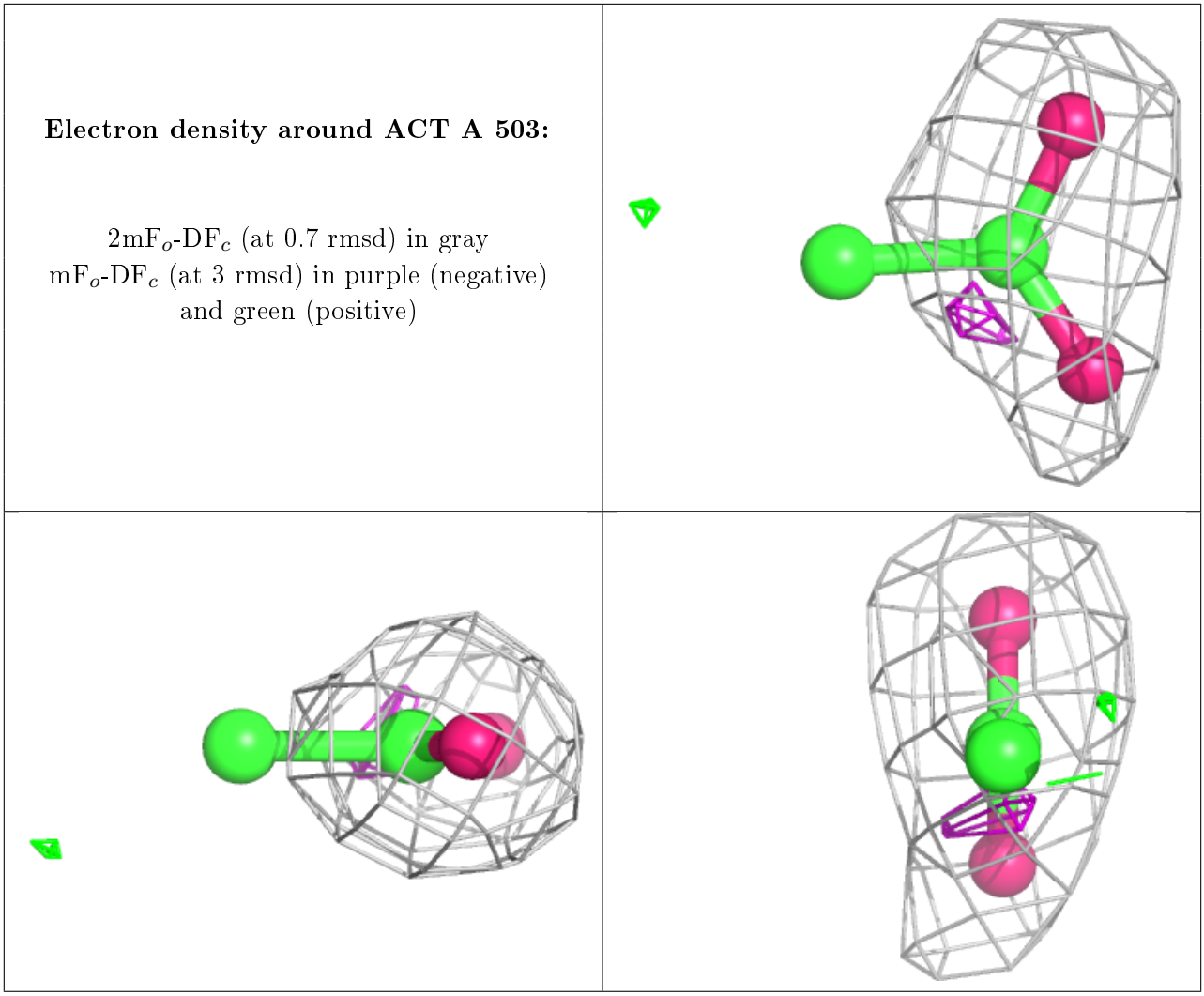
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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
3	ACT	A	503	4/4	0.94	0.25	25,36,38,39	0
3	ACT	B	503	4/4	0.95	0.18	31,32,35,35	0
2	CL	B	501	1/1	0.99	0.08	21,21,21,21	0
2	CL	A	501	1/1	0.99	0.15	25,25,25,25	0
2	CL	B	502	1/1	0.99	0.04	27,27,27,27	0
2	CL	A	502	1/1	1.00	0.04	34,34,34,34	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.

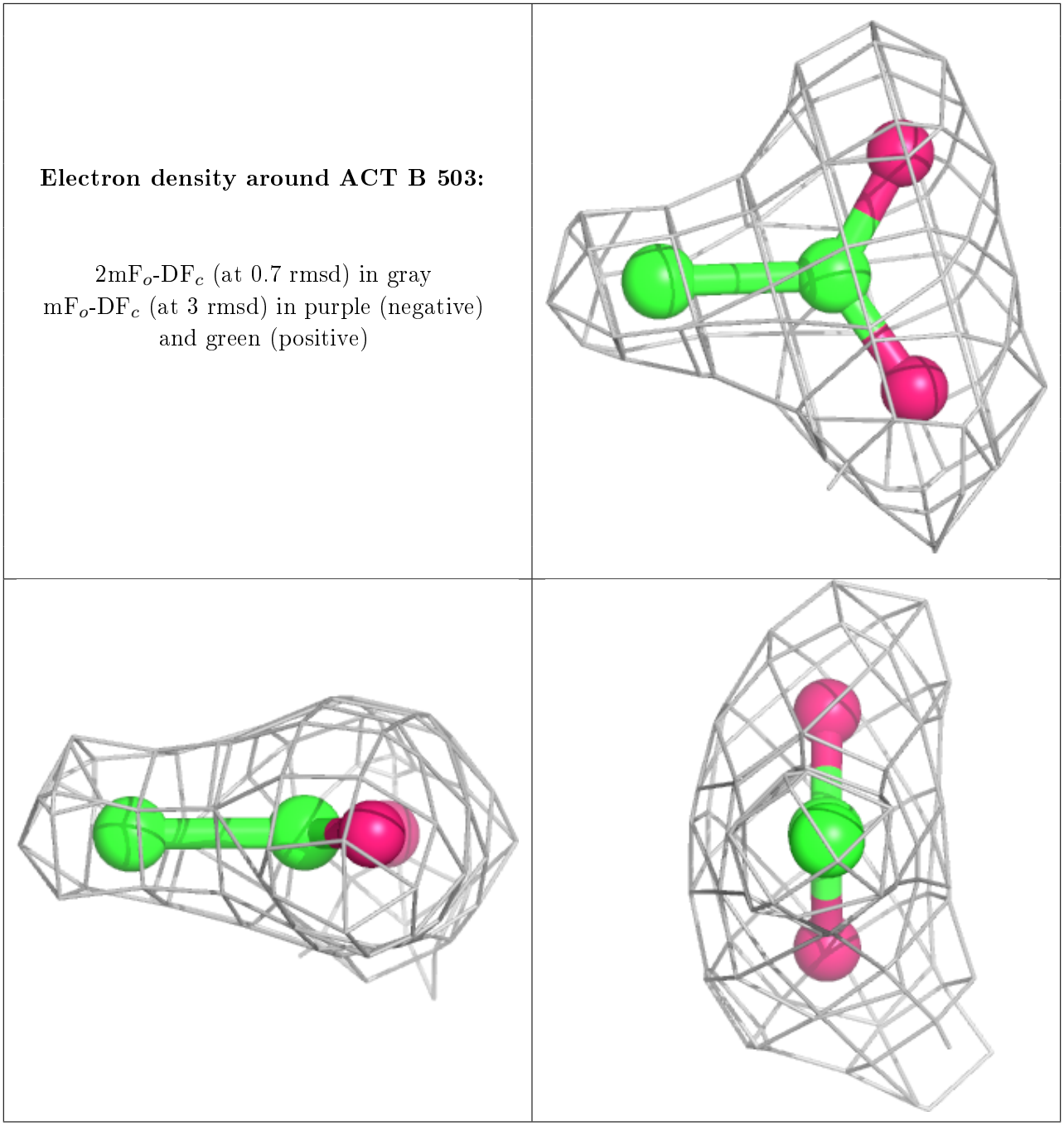
Electron density around ACT A 503:

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



Electron density around ACT B 503:

$2mF_o$-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.