Full wwPDB X-ray Structure Validation Report ⓘ

Jan 20, 2024 – 03:30 pm GMT

PDB ID	:	7NSW
Title	:	The structure of the SBP TarP_Csal in complex with coumarate
Authors	:	Bisson, C.; Salmon, R.C.; West, L.; Rafferty, J.B.; Hitchcock, A.; Thomas, G.H.; Kelly, D.J.
Deposited on	:	2021-03-08
Resolution	:	1.67 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

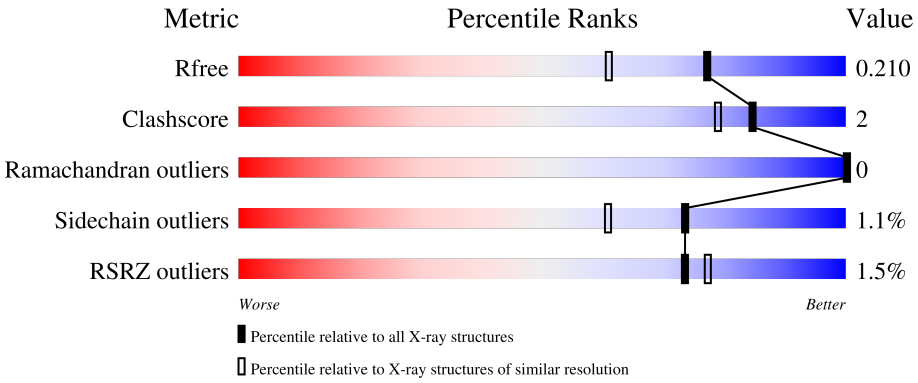
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
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

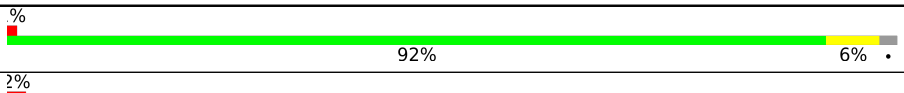
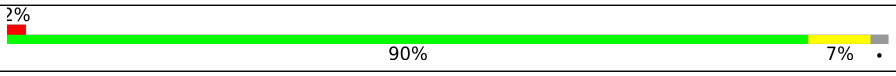
The reported resolution of this entry is 1.67 Å.

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	6780 (1.70-1.66)
Clashscore	141614	7310 (1.70-1.66)
Ramachandran outliers	138981	7173 (1.70-1.66)
Sidechain outliers	138945	7172 (1.70-1.66)
RSRZ outliers	127900	6661 (1.70-1.66)

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	335	
1	BBB	335	

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 5368 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.

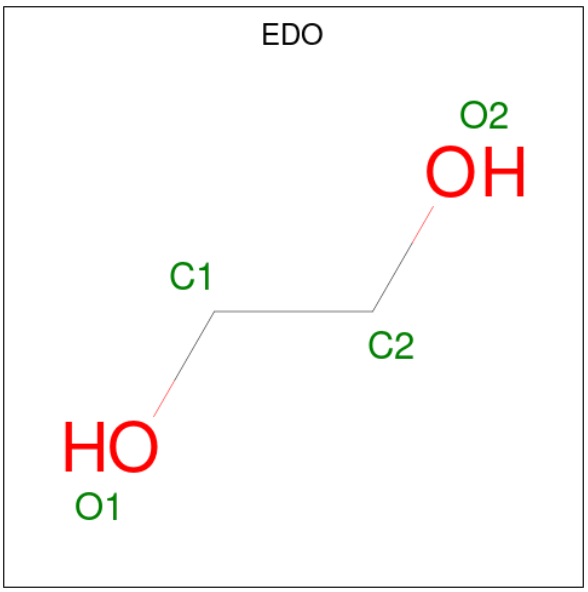
- Molecule 1 is a protein called TRAP dicarboxylate transporter-DctP subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	AAA	327	Total	C	N	O	S	0	0	0
			2546	1604	434	497	11			
1	BBB	328	Total	C	N	O	S	0	0	0
			2555	1609	436	499	11			

There are 24 discrepancies between the modelled and reference sequences:

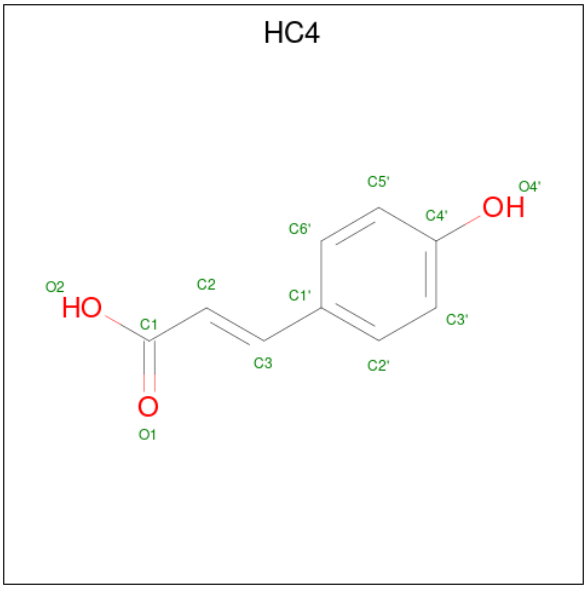
Chain	Residue	Modelled	Actual	Comment	Reference
AAA	0	MET	-	initiating methionine	UNP Q1R0W5
AAA	324	ALA	-	expression tag	UNP Q1R0W5
AAA	325	ALA	-	expression tag	UNP Q1R0W5
AAA	326	ALA	-	expression tag	UNP Q1R0W5
AAA	327	LEU	-	expression tag	UNP Q1R0W5
AAA	328	GLU	-	expression tag	UNP Q1R0W5
AAA	329	HIS	-	expression tag	UNP Q1R0W5
AAA	330	HIS	-	expression tag	UNP Q1R0W5
AAA	331	HIS	-	expression tag	UNP Q1R0W5
AAA	332	HIS	-	expression tag	UNP Q1R0W5
AAA	333	HIS	-	expression tag	UNP Q1R0W5
AAA	334	HIS	-	expression tag	UNP Q1R0W5
BBB	0	MET	-	initiating methionine	UNP Q1R0W5
BBB	324	ALA	-	expression tag	UNP Q1R0W5
BBB	325	ALA	-	expression tag	UNP Q1R0W5
BBB	326	ALA	-	expression tag	UNP Q1R0W5
BBB	327	LEU	-	expression tag	UNP Q1R0W5
BBB	328	GLU	-	expression tag	UNP Q1R0W5
BBB	329	HIS	-	expression tag	UNP Q1R0W5
BBB	330	HIS	-	expression tag	UNP Q1R0W5
BBB	331	HIS	-	expression tag	UNP Q1R0W5
BBB	332	HIS	-	expression tag	UNP Q1R0W5
BBB	333	HIS	-	expression tag	UNP Q1R0W5
BBB	334	HIS	-	expression tag	UNP Q1R0W5

- Molecule 2 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	AAA	1	Total C O 4 2 2	0	0

- Molecule 3 is 4'-HYDROXYCINNAMIC ACID (three-letter code: HC4) (formula: $C_9H_8O_3$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	AAA	1	Total C O 12 9 3	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	BBB	1	Total C O 12 9 3	0	0

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	AAA	2	Total Mg 2 2	0	0
4	BBB	1	Total Mg 1 1	0	0

- Molecule 5 is water.

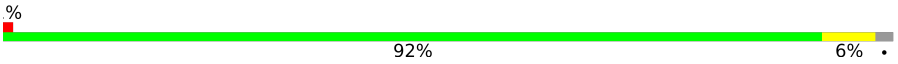
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	AAA	130	Total O 130 130	0	0
5	BBB	106	Total O 106 106	0	0

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.

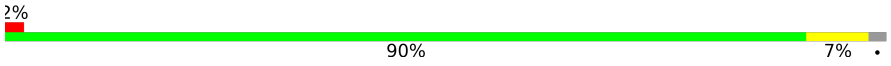
• Molecule 1: TRAP dicarboxylate transporter-DctP subunit

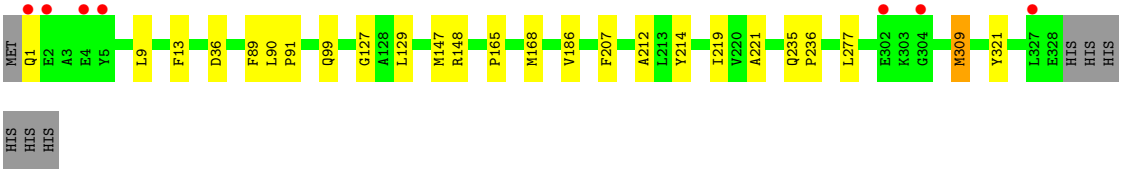
Chain AAA:



• Molecule 1: TRAP dicarboxylate transporter-DctP subunit

Chain BBB:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	82.17Å 119.94Å 62.04Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	34.25 – 1.67 34.25 – 1.67	Depositor EDS
% Data completeness (in resolution range)	99.5 (34.25-1.67) 99.5 (34.25-1.67)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.97 (at 1.67Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.197 , 0.251 0.209 , 0.210	Depositor DCC
R_{free} test set	3544 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å^2)	17.7	Xtriage
Anisotropy	0.668	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 42.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	5368	wwPDB-VP
Average B, all atoms (Å^2)	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 40.67 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 2.6419e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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: EDO, HC4, MG

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	0.77	1/2603 (0.0%)	0.87	0/3538
1	BBB	0.74	0/2612	0.85	0/3550
All	All	0.76	1/5215 (0.0%)	0.86	0/7088

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	AAA	203	GLU	CD-OE2	-5.13	1.20	1.25

There are no bond angle outliers.

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	2546	0	2465	9	0
1	BBB	2555	0	2476	14	0
2	AAA	4	0	6	0	0
3	AAA	12	0	7	0	0
3	BBB	12	0	6	0	0
4	AAA	2	0	0	0	0
4	BBB	1	0	0	0	0
5	AAA	130	0	0	1	1

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	BBB	106	0	0	0	1
All	All	5368	0	4960	23	1

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 (23) 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:99:GLN:NE2	1:BBB:321:TYR:CE1	2.67	0.62
1:AAA:89:PHE:HA	1:AAA:212:ALA:O	1.99	0.61
1:BBB:148:ARG:HG3	1:BBB:168:MET:HG2	1.81	0.61
1:BBB:129:LEU:HD23	1:BBB:186:VAL:HG12	1.83	0.60
1:AAA:148:ARG:HG3	1:AAA:168:MET:HG2	1.87	0.57
1:BBB:127:GLY:HA3	1:BBB:214:TYR:CE1	2.40	0.56
1:BBB:90:LEU:N	1:BBB:91:PRO:CD	2.71	0.54
1:AAA:150:PRO:HD2	1:AAA:154:VAL:HG11	1.91	0.51
1:BBB:89:PHE:HA	1:BBB:212:ALA:O	2.11	0.49
1:BBB:1:GLN:HA	1:BBB:36:ASP:O	2.13	0.49
1:BBB:147:MET:O	1:BBB:165:PRO:HA	2.14	0.48
1:AAA:267:SER:HB3	5:AAA:523:HOH:O	2.15	0.46
1:BBB:99:GLN:NE2	1:BBB:321:TYR:CD1	2.85	0.45
1:AAA:125:SER:HB2	1:AAA:126:PRO:CD	2.47	0.45
1:AAA:90:LEU:N	1:AAA:91:PRO:CD	2.81	0.44
1:AAA:125:SER:HB2	1:AAA:126:PRO:HD2	2.01	0.43
1:BBB:219:ILE:HG12	1:BBB:221:ALA:HB2	2.01	0.42
1:BBB:9:LEU:C	1:BBB:9:LEU:HD23	2.39	0.42
1:BBB:235:GLN:HB2	1:BBB:236:PRO:HD3	2.01	0.42
1:AAA:127:GLY:HA3	1:AAA:214:TYR:CE1	2.56	0.41
1:AAA:316:ARG:O	1:AAA:319:GLU:HG2	2.21	0.41
1:BBB:207:PHE:CE2	1:BBB:277:LEU:HD12	2.56	0.40
1:BBB:309:MET:C	1:BBB:309:MET:SD	3.00	0.40

All (1) 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)
5:AAA:577:HOH:O	5:BBB:604:HOH:O[3_646]	1.88	0.32

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	325/335 (97%)	321 (99%)	4 (1%)	0	100	100
1	BBB	326/335 (97%)	319 (98%)	7 (2%)	0	100	100
All	All	651/670 (97%)	640 (98%)	11 (2%)	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	AAA	262/270 (97%)	258 (98%)	4 (2%)	65	48
1	BBB	263/270 (97%)	261 (99%)	2 (1%)	81	72
All	All	525/540 (97%)	519 (99%)	6 (1%)	73	61

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	13	PHE
1	AAA	114	GLU
1	AAA	152	ARG
1	AAA	156	ARG
1	BBB	13	PHE
1	BBB	309	MET

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)

Of 6 ligands modelled in this entry, 3 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	$\#\|Z\| > 2$
3	HC4	BBB	401	-	12,12,12	0.79	1 (8%)	15,15,15	0.48	0
2	EDO	AAA	401	-	3,3,3	0.13	0	2,2,2	0.16	0
3	HC4	AAA	402	-	12,12,12	0.87	1 (8%)	15,15,15	1.05	1 (6%)

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
3	HC4	BBB	401	-	-	0/5/5/5	0/1/1/1
2	EDO	AAA	401	-	-	0/1/1/1	-
3	HC4	AAA	402	-	-	0/5/5/5	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	BBB	401	HC4	O2-C1	-2.15	1.24	1.30
3	AAA	402	HC4	O2-C1	-2.11	1.24	1.30

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	AAA	402	HC4	C2'-C3'-C4'	2.34	122.45	119.88

There are no chirality outliers.

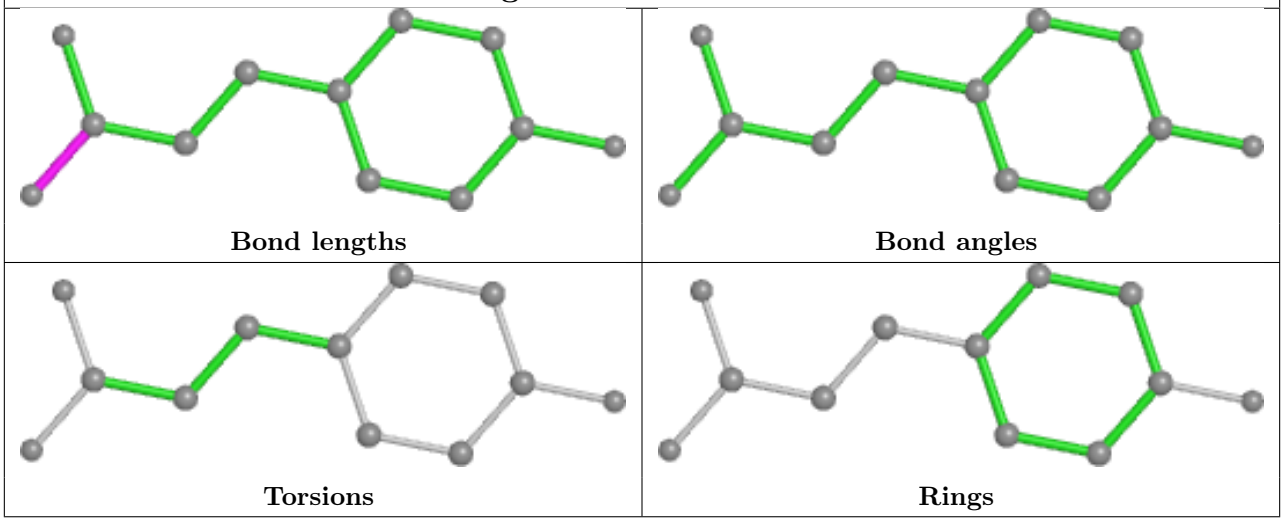
There are no torsion outliers.

There are no ring outliers.

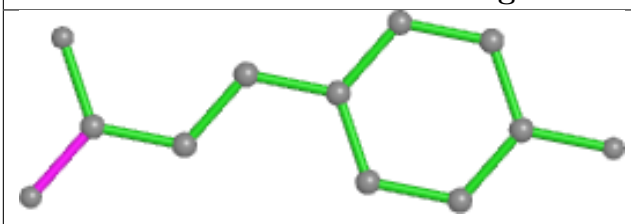
No monomer is involved in short contacts.

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.

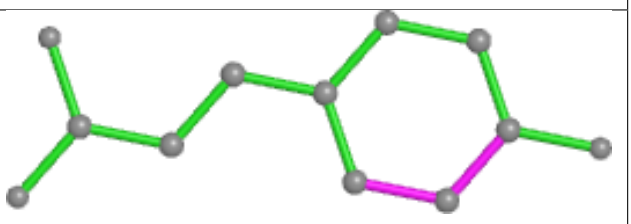
Ligand HC4 BBB 401



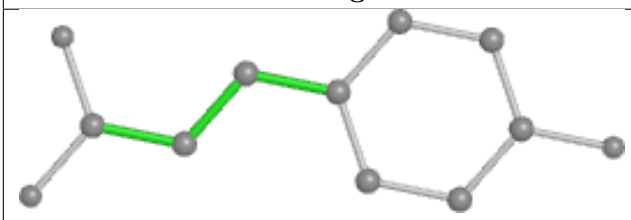
Ligand HC4 AAA 402



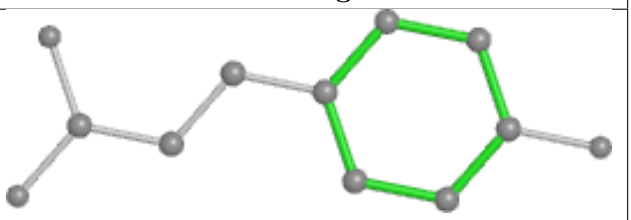
Bond lengths



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, 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	327/335 (97%)	-0.22	3 (0%)	84	87	12, 23, 44, 101	0
1	BBB	328/335 (97%)	-0.06	7 (2%)	63	67	14, 25, 52, 89	0
All	All	655/670 (97%)	-0.14	10 (1%)	73	77	12, 24, 49, 101	0

All (10) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	2	GLU	3.5
1	BBB	1	GLN	3.4
1	BBB	327	LEU	3.3
1	BBB	304	GLY	2.9
1	BBB	5	TYR	2.6
1	BBB	302	GLU	2.4
1	AAA	3	ALA	2.1
1	BBB	4	GLU	2.1
1	BBB	2	GLU	2.0
1	AAA	5	TYR	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 (i)

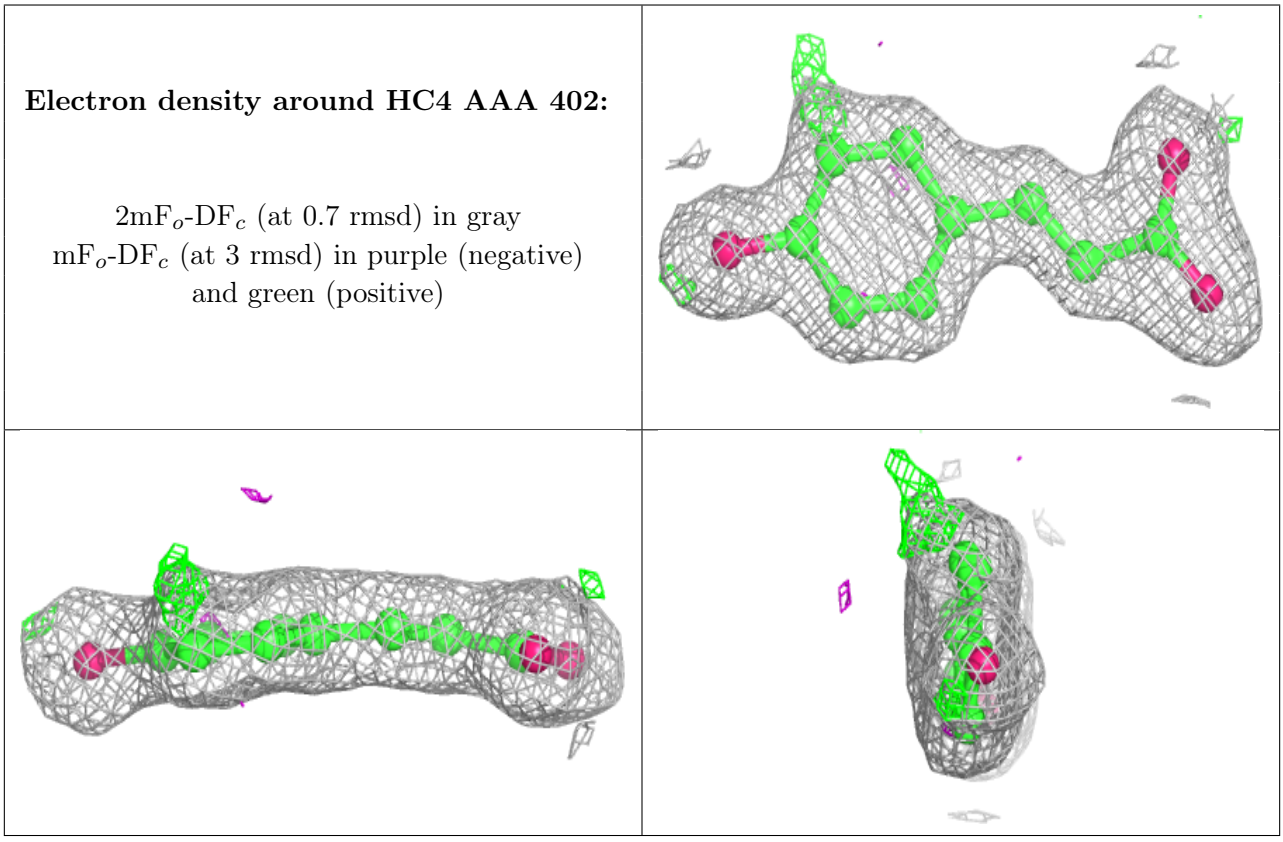
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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	EDO	AAA	401	4/4	0.86	0.14	32,37,57,66	0
4	MG	AAA	404	1/1	0.95	0.11	30,30,30,30	0
3	HC4	AAA	402	12/12	0.96	0.09	9,16,19,21	0
3	HC4	BBB	401	12/12	0.97	0.07	13,19,25,31	0
4	MG	AAA	403	1/1	0.98	0.06	25,25,25,25	0
4	MG	BBB	402	1/1	0.98	0.04	29,29,29,29	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.

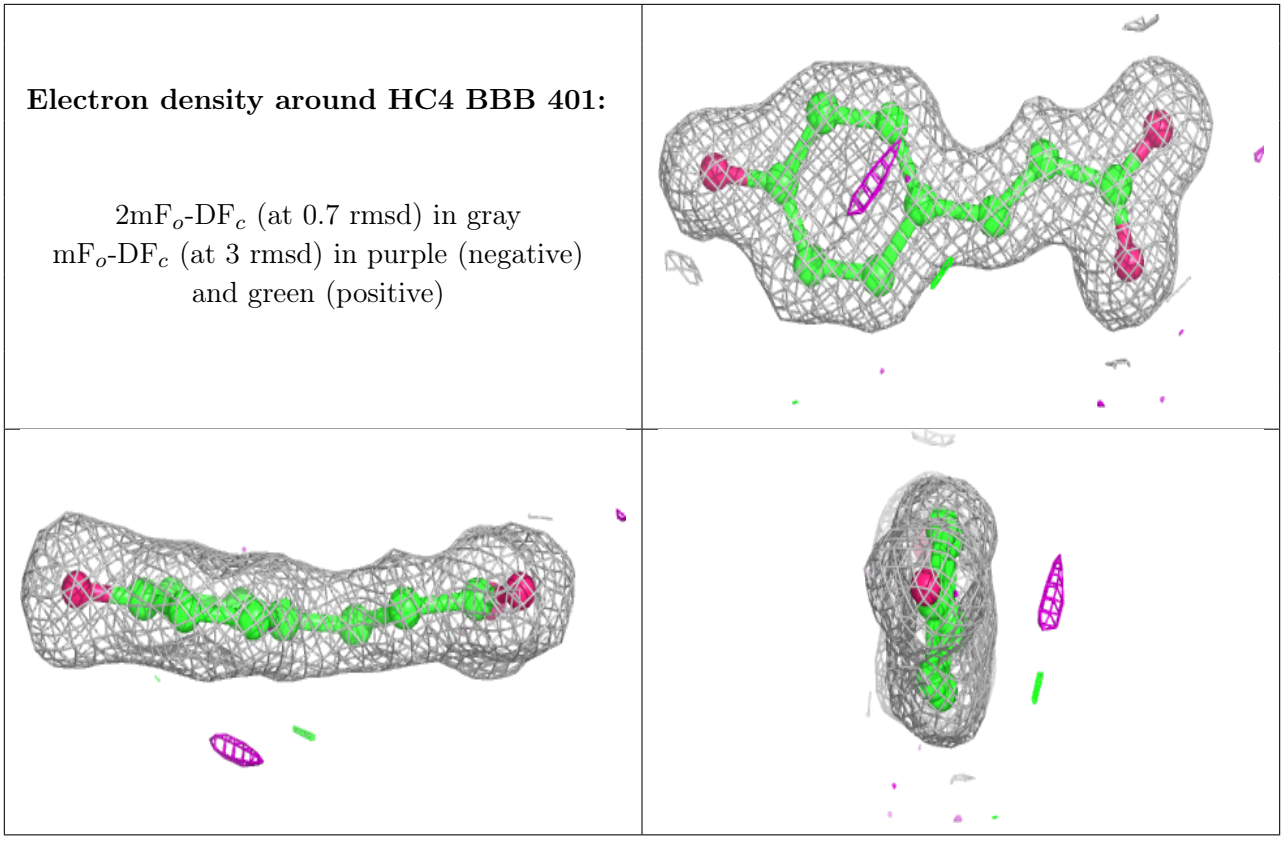
Electron density around HC4 AAA 402:

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



Electron density around HC4 BBB 401:

$2mF_o$-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.