



Full wwPDB X-ray Structure Validation Report ⓘ

Aug 17, 2026 – 04:17 am BST

PDB ID : 9THV / pdb_00009thv
Title : Bacteriodes thetaiotamicron sulphatase BT1636_S77C in complex with Chromate Ions
Authors : Tomlinson, C.W.E.; Cartmell, A.
Deposited on : 2025-12-04
Resolution : 1.70 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	:	4-5-2 with Phenix2.0
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Buster-report	:	wwPDB partial adaption of 1.1.7 (2018)
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.015 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.50

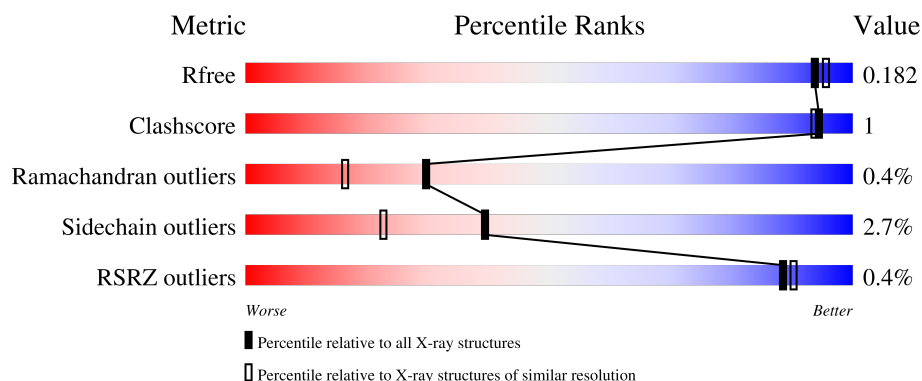
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.70 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}	180053	5551 (1.70-1.70)
Clashscore	190562	5924 (1.70-1.70)
Ramachandran outliers	187476	5846 (1.70-1.70)
Sidechain outliers	187428	5846 (1.70-1.70)
RSRZ outliers	180081	5554 (1.70-1.70)

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 $\leq 5\%$. 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	510	 88% 6% 6%

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 7950 atoms, of which 3751 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 Arylsulfatase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	481	7622	2483	3737	653	723	26	114	6	0

There are 28 discrepancies between the modelled and reference sequences:

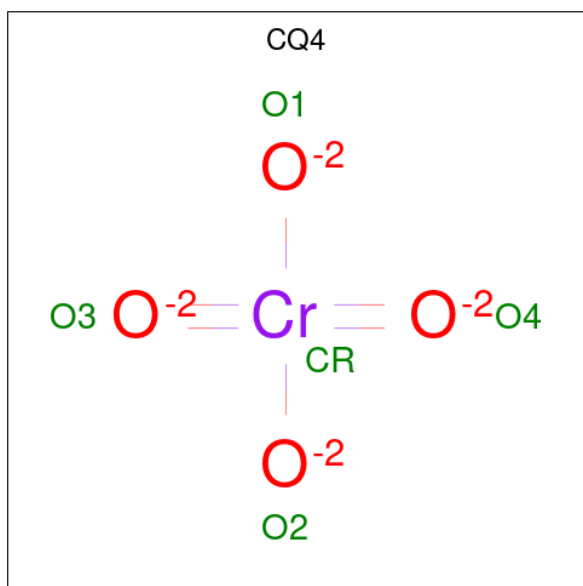
Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	initiating methionine	UNP Q8A789
A	1	GLY	-	expression tag	UNP Q8A789
A	2	SER	-	expression tag	UNP Q8A789
A	3	SER	-	expression tag	UNP Q8A789
A	4	HIS	-	expression tag	UNP Q8A789
A	5	HIS	-	expression tag	UNP Q8A789
A	6	HIS	-	expression tag	UNP Q8A789
A	7	HIS	-	expression tag	UNP Q8A789
A	8	HIS	-	expression tag	UNP Q8A789
A	9	HIS	-	expression tag	UNP Q8A789
A	10	SER	-	expression tag	UNP Q8A789
A	11	SER	-	expression tag	UNP Q8A789
A	12	GLY	-	expression tag	UNP Q8A789
A	13	LEU	-	expression tag	UNP Q8A789
A	14	VAL	-	expression tag	UNP Q8A789
A	15	PRO	-	expression tag	UNP Q8A789
A	16	ARG	-	expression tag	UNP Q8A789
A	17	GLY	-	expression tag	UNP Q8A789
A	18	SER	-	expression tag	UNP Q8A789
A	19	HIS	-	expression tag	UNP Q8A789
A	20	MET	-	expression tag	UNP Q8A789
A	21	ALA	-	expression tag	UNP Q8A789
A	22	SER	-	expression tag	UNP Q8A789
A	23	GLN	-	expression tag	UNP Q8A789
A	24	LYS	-	expression tag	UNP Q8A789
A	25	ASN	-	expression tag	UNP Q8A789
A	26	ASN	-	expression tag	UNP Q8A789

Continued on next page...

Continued from previous page...

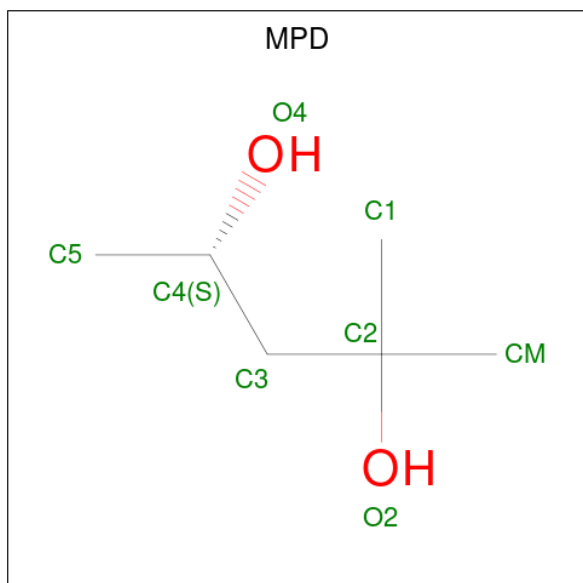
Chain	Residue	Modelled	Actual	Comment	Reference
A	27	THR	-	expression tag	UNP Q8A789

- Molecule 2 is Chromate (CCD ID: CQ4) (formula: CrO_4) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	Cr	O	0	0
			5	1	4		

- Molecule 3 is (4S)-2-METHYL-2,4-PENTANEDIOL (CCD ID: MPD) (formula: $\text{C}_6\text{H}_{14}\text{O}_2$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	H	O	2	0
			22	6	14	2		

- Molecule 4 is CALCIUM ION (CCD ID: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	Ca	0	0
			1	1		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	300	Total	O	0	0
			300	300		

- Molecule 1: Arylsulfatase

Amino Acid	Frequency (%)	Group / Annotation
MET	~10	Grey Group
GLY	~10	
SER	~10	
HIS	~10	
THR	~10	
LYS	~10	
ARG	~10	
VAL	~10	
LEU	~10	
PRO	~10	
K28	~10	Green Bar
D38	~10	Yellow Group
A78	~10	
P79	~10	Orange Group
T107	~10	
K114	~10	Light Blue Group
I130	~10	
L131	~10	Dark Blue Group
M144	~10	
F145	~10	Purple Group
R162	~10	
Y167	~10	Pink Group
Q173	~10	
K189	~10	Brown Group
P218	~10	
M230	~10	Red Group
G238	~10	
K239	~10	

4 Data and refinement statistics

Property	Value
Space group	P 21 21 21
Cell constants a, b, c, α , β , γ	74.65Å 87.59Å 103.32Å 90.00° 90.00° 90.00°
Resolution (Å)	66.81 – 1.70 66.81 – 1.70
% Data completeness (in resolution range)	100.0 (66.81-1.70) 100.0 (66.81-1.70)
R_{merge}	0.11
R_{sym}	(Not available)
$\langle I/\sigma(I) \rangle$ ¹	1.00 (at 1.70Å)
Refinement program	REFMAC 5.8.0431 (refmacat 0.4.105), REFMAC 5.8.0431 (refmacat 0.4.105)
R, R_{free}	0.153 , 0.194 0.141 , 0.182
R_{free} test set	3869 reflections (5.15%)
Wilson B-factor (Å ²)	27.3
Anisotropy	0.402
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.43 , 48.7
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$
Estimated twinning fraction	No twinning to report.
F_o, F_c correlation	0.98
Total number of atoms	7950
Average B, all atoms (Å ²)	32.0

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.72% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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: CQ4, CA, MPD

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	0.61	0/4019	1.03	8/5436 (0.1%)

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
1	A	0	1

There are no bond length outliers.

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	273	ASP	CA-CB-CG	7.76	120.36	112.60
1	A	469	ASP	CA-CB-CG	5.98	118.58	112.60
1	A	482	GLU	CB-CG-CD	5.82	122.50	112.60
1	A	173	GLN	CB-CA-C	-5.81	99.88	110.63
1	A	145	PHE	CA-CB-CG	-5.46	108.34	113.80
1	A	399	ASN	CA-CB-CG	-5.44	107.16	112.60
1	A	107	THR	CA-CB-OG1	-5.29	101.66	109.60
1	A	38	ASP	CA-CB-CG	5.20	117.80	112.60

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	238	GLY	Peptide

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	3885	3737	3699	9	0
2	A	5	0	0	1	0
3	A	8	14	14	1	0
4	A	1	0	0	0	0
5	A	300	0	0	2	0
All	All	4199	3751	3713	10	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 1.

All (10) 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:A:427:LYS:NZ	5:A:701:HOH:O	2.11	0.71
1:A:230[A]:MET:HE2	1:A:230[A]:MET:HA	1.83	0.60
3:A:602:MPD:HM1	3:A:602:MPD:O4	2.09	0.53
1:A:267:LYS:O	5:A:702:HOH:O	2.20	0.46
1:A:130:ILE:HD11	1:A:162:ARG:HB3	2.00	0.44
1:A:252:HIS:NE2	2:A:601:CQ4:O1	2.35	0.44
1:A:508:LYS:HD2	1:A:508:LYS:C	2.42	0.44
1:A:420:ALA:N	1:A:421:PRO:CD	2.83	0.42
1:A:78:ALA:N	1:A:79:PRO:HD2	2.35	0.41
1:A:144:MET:O	1:A:167:TYR:HA	2.21	0.41

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was

analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	485/510 (95%)	467 (96%)	16 (3%)	2 (0%)	30	16

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	352	LYS
1	A	239	LYS

5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	414/433 (96%)	403 (97%)	11 (3%)	39	23

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	114	LYS
1	A	131	LEU
1	A	189	LYS
1	A	218	PRO
1	A	239	LYS
1	A	251	PRO
1	A	329	ASN
1	A	362	ILE
1	A	407	LYS
1	A	427	LYS
1	A	428	LYS

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

Mol	Chain	Res	Type
1	A	320	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	479	GLN
1	A	500	HIS

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates ⓘ

There are no oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 3 ligands modelled in this entry, 1 is 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	MPD	A	602	-	7,7,7	0.42	0	9,10,10	0.79	0
2	CQ4	A	601	-	2,4,4	1.34	0	-		

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	MPD	A	602	-	-	3/5/5/5	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) torsion outliers are listed below:

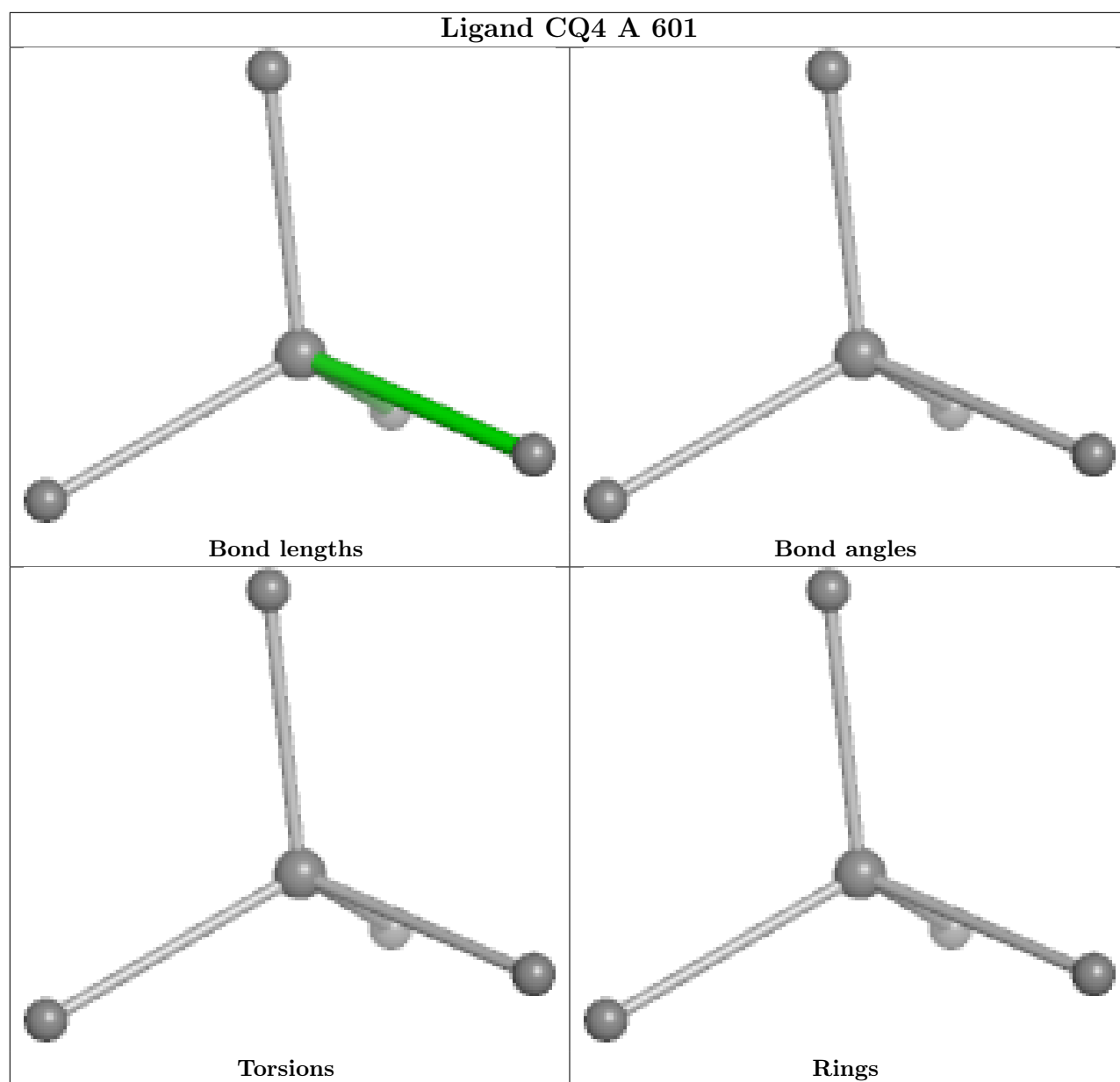
Mol	Chain	Res	Type	Atoms
3	A	602	MPD	O2-C2-C3-C4
3	A	602	MPD	CM-C2-C3-C4
3	A	602	MPD	C2-C3-C4-O4

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	602	MPD	1	0
2	A	601	CQ4	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 than 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.



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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	481/510 (94%)	-0.80	2 (0%) 88 90	14, 29, 49, 83	3 (0%)

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	508	LYS	2.9
1	A	239	LYS	2.5

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 oligosaccharides in this entry.

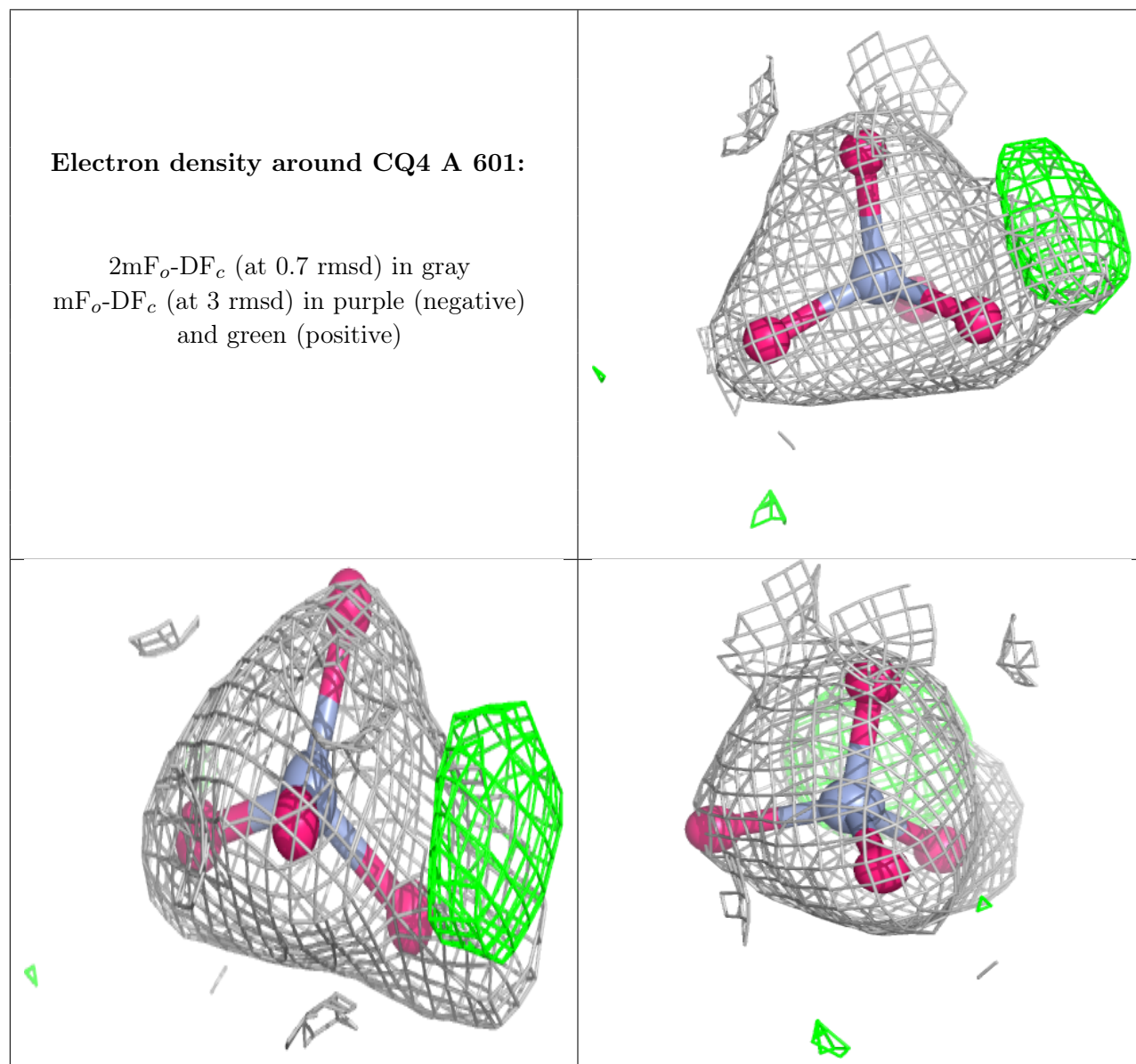
6.4 Ligands [i](#)

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, 95th percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
3	MPD	A	602	8/8	0.92	0.11	37,46,49,50	2
2	CQ4	A	601	5/5	0.98	0.08	39,41,43,46	5
4	CA	A	603	1/1	1.00	0.03	24,24,24,24	1

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.



6.5 Other polymers ⓘ

There are no such residues in this entry.