



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 4, 2023 – 01:11 pm BST

PDB ID : 7QGO  
Title : Human CD73 (ecto 5'-nucleotidase) in complex with MRS4602 (a 3-methyl-CMPCP derivative, compound 21 in paper) in the closed state (crystal form III)  
Authors : Strater, N.  
Deposited on : 2021-12-09  
Resolution : 2.21 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](mailto: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)  
Xtrriage (Phenix) : 1.13  
EDS : 2.35.1  
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) : 2.35.1

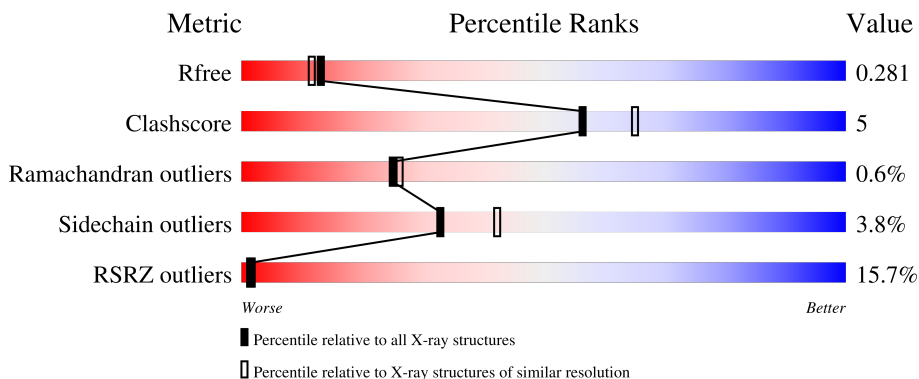
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

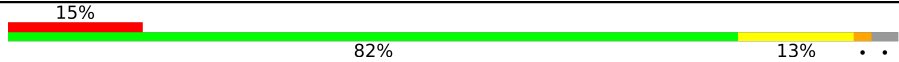
The reported resolution of this entry is 2.21 Å.

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 of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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	541	

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 4302 atoms, of which 26 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 5'-nucleotidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	523	4087	2599	696	773	19	0	1	0

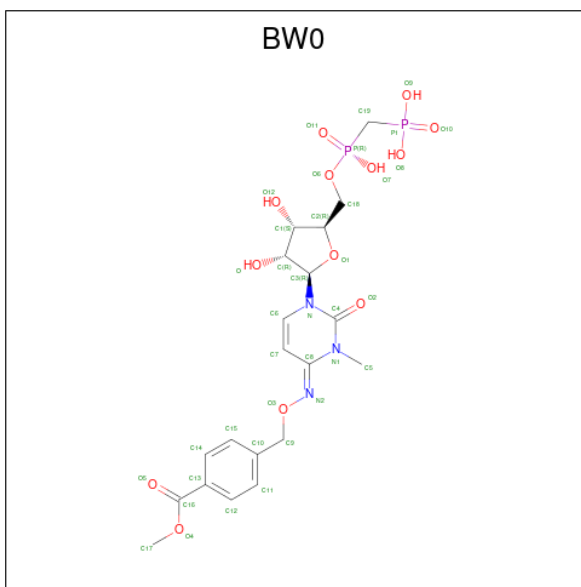
There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	53	ASP	ASN	engineered mutation	UNP P21589
A	311	ASP	ASN	engineered mutation	UNP P21589
A	333	ASP	ASN	engineered mutation	UNP P21589
A	403	ASP	ASN	engineered mutation	UNP P21589
A	550	GLY	-	expression tag	UNP P21589
A	551	GLY	-	expression tag	UNP P21589
A	552	GLY	-	expression tag	UNP P21589
A	553	GLY	-	expression tag	UNP P21589
A	554	ALA	-	expression tag	UNP P21589
A	555	GLY	-	expression tag	UNP P21589
A	556	GLY	-	expression tag	UNP P21589
A	557	GLY	-	expression tag	UNP P21589
A	558	GLY	-	expression tag	UNP P21589
A	559	GLY	-	expression tag	UNP P21589
A	560	THR	-	expression tag	UNP P21589
A	561	LYS	-	expression tag	UNP P21589
A	562	HIS	-	expression tag	UNP P21589
A	563	HIS	-	expression tag	UNP P21589
A	564	HIS	-	expression tag	UNP P21589
A	565	HIS	-	expression tag	UNP P21589
A	566	HIS	-	expression tag	UNP P21589
A	567	HIS	-	expression tag	UNP P21589

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	2	Total	Zn	0	0
			2	2		

- Molecule 3 is [(2 {R},3 {S},4 {R},5 {R})-5-[(4 {E})-4-[(4-methoxycarbonylphenyl)methoxyimino]-3-methyl-2-oxidanylidene-pyrimidin-1-yl]-3,4-bis(oxidanyl)oxolan-2-yl]methoxy-oxidanyl-phosphoryl]methylphosphonic acid (three-letter code: BW0) (formula: C<sub>20</sub>H<sub>27</sub>N<sub>3</sub>O<sub>13</sub>P<sub>2</sub>) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	
			Total	C	H	N	O			P
3	A	1	64	20	26	3	13	2	26	0

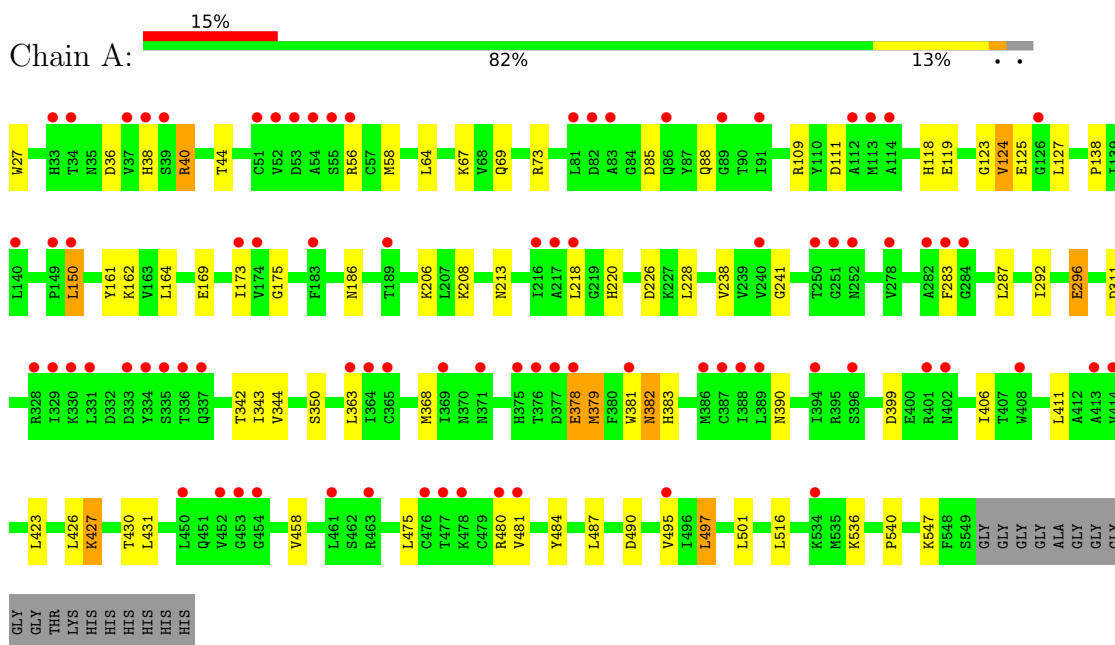
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	149	Total	O	0	0
			149	149		

### 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: 5'-nucleotidase



## 4 Data and refinement statistics i

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	55.04Å 96.83Å 232.69Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	27.84 – 2.21 27.84 – 2.21	Depositor EDS
% Data completeness (in resolution range)	99.7 (27.84-2.21) 99.7 (27.84-2.21)	Depositor EDS
$R_{merge}$	0.15	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.02 (at 2.20Å)	Xtrriage
Refinement program	BUSTER 2.10.4 (20-OCT-2021)	Depositor
R, $R_{free}$	0.239 , 0.292 0.231 , 0.281	Depositor DCC
$R_{free}$ test set	1664 reflections (5.24%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	50.1	Xtrriage
Anisotropy	0.601	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 64.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	0.001 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.030 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	4302	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	80.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.24% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: BW0, ZN

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	0.42	0/4176	0.63	0/5655

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 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	4087	0	4062	41	0
2	A	2	0	0	0	0
3	A	38	26	0	0	0
4	A	149	0	0	1	0
All	All	4276	26	4062	41	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 5.

All (41) 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:208:LYS:NZ	1:A:213:ASN:HA	2.02	0.74
1:A:40:ARG:HG2	1:A:44:THR:HG21	1.75	0.67
1:A:40:ARG:HD3	1:A:283:PHE:HB3	1.76	0.67
1:A:378:GLU:O	1:A:379:MET:HB2	1.96	0.66
1:A:382:ASN:ND2	1:A:383:HIS:HD2	1.94	0.64
1:A:296:GLU:H	1:A:296:GLU:CD	2.04	0.61
1:A:208:LYS:HZ1	1:A:213:ASN:HA	1.65	0.60
1:A:390:ASN:HB3	1:A:501:LEU:HD21	1.84	0.60
1:A:111:ASP:O	1:A:138:PRO:HD2	2.08	0.54
1:A:368:MET:HE1	1:A:497:LEU:HA	1.90	0.53
1:A:64:LEU:HD11	1:A:287:LEU:HD13	1.91	0.52
1:A:363:LEU:HB2	1:A:540:PRO:HG3	1.91	0.52
1:A:161:TYR:CE2	1:A:206:LYS:NZ	2.78	0.51
1:A:69[B]:GLN:HE22	1:A:109:ARG:HH21	1.59	0.51
1:A:226:ASP:OD2	1:A:241:GLY:HA3	2.12	0.50
1:A:343:ILE:HG13	1:A:344:VAL:HG13	1.94	0.50
1:A:381:TRP:HB3	1:A:423:LEU:HD22	1.93	0.50
1:A:119:GLU:HB3	1:A:127:LEU:HD13	1.94	0.49
1:A:426:LEU:HD22	1:A:430:THR:HG21	1.96	0.47
1:A:138:PRO:HG3	1:A:162:LYS:HG2	1.97	0.46
1:A:342:THR:HB	1:A:406:ILE:HD11	1.97	0.46
1:A:123:GLY:HA2	1:A:186:ASN:O	2.17	0.45
1:A:38:HIS:CE1	1:A:85:ASP:HB3	2.52	0.45
1:A:124:VAL:HG11	1:A:150:LEU:HD11	1.99	0.44
1:A:124:VAL:HG23	4:A:759:HOH:O	2.18	0.44
1:A:475:LEU:HD13	1:A:484:TYR:CE1	2.53	0.44
1:A:175:GLY:HA2	1:A:218:LEU:O	2.18	0.43
1:A:238:VAL:HG21	1:A:292:ILE:CD1	2.49	0.43
1:A:36:ASP:OD1	1:A:220:HIS:HE1	2.02	0.43
1:A:427:LYS:HD2	1:A:516:LEU:HD21	2.00	0.42
1:A:431:LEU:HD11	1:A:495:VAL:HG21	2.01	0.42
1:A:164:LEU:HD22	1:A:173:ILE:HD12	2.01	0.42
1:A:536:LYS:HE2	1:A:536:LYS:HA	2.02	0.42
1:A:58:MET:HG2	1:A:311:ASP:HA	2.01	0.42
1:A:238:VAL:HG21	1:A:292:ILE:HD13	2.01	0.42
1:A:27:TRP:CE2	1:A:169:GLU:HG3	2.56	0.41
1:A:458:VAL:HG22	1:A:547:LYS:HB2	2.02	0.41
1:A:411:LEU:HD23	1:A:411:LEU:HA	1.98	0.41
1:A:27:TRP:NE1	1:A:169:GLU:HG3	2.36	0.41
1:A:85:ASP:OD2	1:A:118:HIS:HD2	2.05	0.40

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	522/541 (96%)	494 (95%)	25 (5%)	3 (1%)	25 26

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	378	GLU
1	A	379	MET
1	A	88	GLN

### 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	451/459 (98%)	434 (96%)	17 (4%)	33 42

All (17) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	40	ARG
1	A	56	ARG
1	A	67	LYS
1	A	124	VAL
1	A	125	GLU
1	A	150	LEU
1	A	228	LEU
1	A	296	GLU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	350	SER
1	A	382	ASN
1	A	399	ASP
1	A	427	LYS
1	A	480	ARG
1	A	481	VAL
1	A	487	LEU
1	A	490	ASP
1	A	497	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	70	GLN
1	A	153	GLN
1	A	372	ASN
1	A	382	ASN
1	A	383	HIS
1	A	503	ASN

### 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 3 ligands modelled in this entry, 2 are monoatomic - leaving 1 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	BW0	A	603	2	37,40,40	0.25	0	47,59,59	0.77	2 (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
3	BW0	A	603	2	-	0/26/44/44	0/3/3/3

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	603	BW0	O11-P-C19	3.90	119.39	109.07
3	A	603	BW0	O3-N2-C8	2.52	112.02	108.83

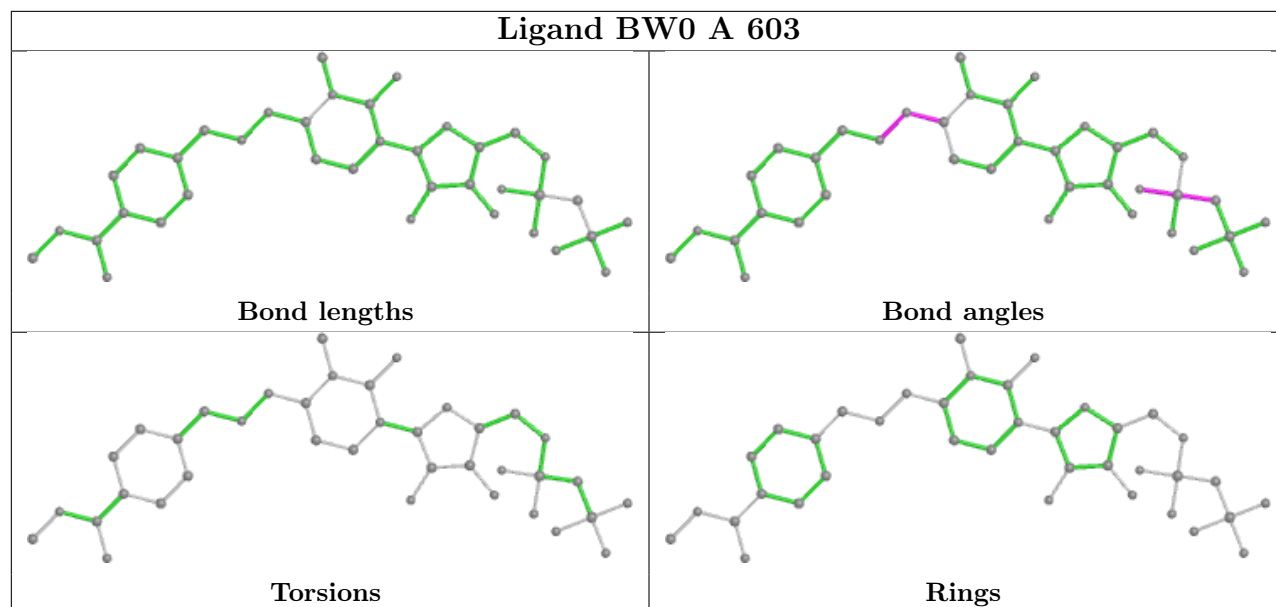
There are no chirality outliers.

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 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

### 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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	523/541 (96%)	0.76	82 (15%) <b>2</b> <b>1</b>	49, 77, 114, 152	5 (0%)

All (82) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	376	THR	8.5
1	A	377	ASP	6.4
1	A	478	LYS	5.6
1	A	52	VAL	5.3
1	A	401	ARG	5.1
1	A	481	VAL	5.1
1	A	218	LEU	5.1
1	A	388	ILE	5.1
1	A	330	LYS	5.0
1	A	387	CYS	5.0
1	A	334	TYR	4.7
1	A	331	LEU	4.7
1	A	37	VAL	4.7
1	A	495	VAL	4.7
1	A	329	ILE	4.6
1	A	83	ALA	4.6
1	A	81	LEU	4.5
1	A	364	ILE	4.4
1	A	378	GLU	4.4
1	A	477	THR	4.2
1	A	452	VAL	4.2
1	A	34	THR	4.0
1	A	82	ASP	4.0
1	A	365	CYS	4.0
1	A	89	GLY	4.0
1	A	140	LEU	3.9
1	A	336	THR	3.9

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	A	240	VAL	3.9
1	A	450	LEU	3.7
1	A	453	GLY	3.7
1	A	183	PHE	3.6
1	A	173	ILE	3.4
1	A	39	SER	3.3
1	A	389	LEU	3.3
1	A	534	LYS	3.2
1	A	414	VAL	3.2
1	A	216	ILE	3.2
1	A	375	HIS	3.1
1	A	112	ALA	3.1
1	A	396	SER	3.1
1	A	381	TRP	3.0
1	A	54	ALA	3.0
1	A	149	PRO	3.0
1	A	56	ARG	2.9
1	A	282	ALA	2.9
1	A	53	ASP	2.9
1	A	337	GLN	2.8
1	A	386	MET	2.8
1	A	461	LEU	2.8
1	A	363	LEU	2.8
1	A	408	TRP	2.7
1	A	217	ALA	2.6
1	A	371	ASN	2.5
1	A	174	VAL	2.5
1	A	480	ARG	2.5
1	A	335	SER	2.5
1	A	402	ASN	2.5
1	A	150	LEU	2.4
1	A	86	GLN	2.4
1	A	113	MET	2.4
1	A	251	GLY	2.3
1	A	189	THR	2.3
1	A	278	VAL	2.3
1	A	463	ARG	2.3
1	A	394	ILE	2.3
1	A	114	ALA	2.3
1	A	55	SER	2.3
1	A	126	GLY	2.3
1	A	33	HIS	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	283	PHE	2.2
1	A	91	ILE	2.2
1	A	454	GLY	2.2
1	A	413	ALA	2.2
1	A	369	ILE	2.1
1	A	51	CYS	2.1
1	A	328	ARG	2.1
1	A	252	ASN	2.1
1	A	250	THR	2.1
1	A	284	GLY	2.1
1	A	476	CYS	2.1
1	A	38	HIS	2.0
1	A	333	ASP	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](#)

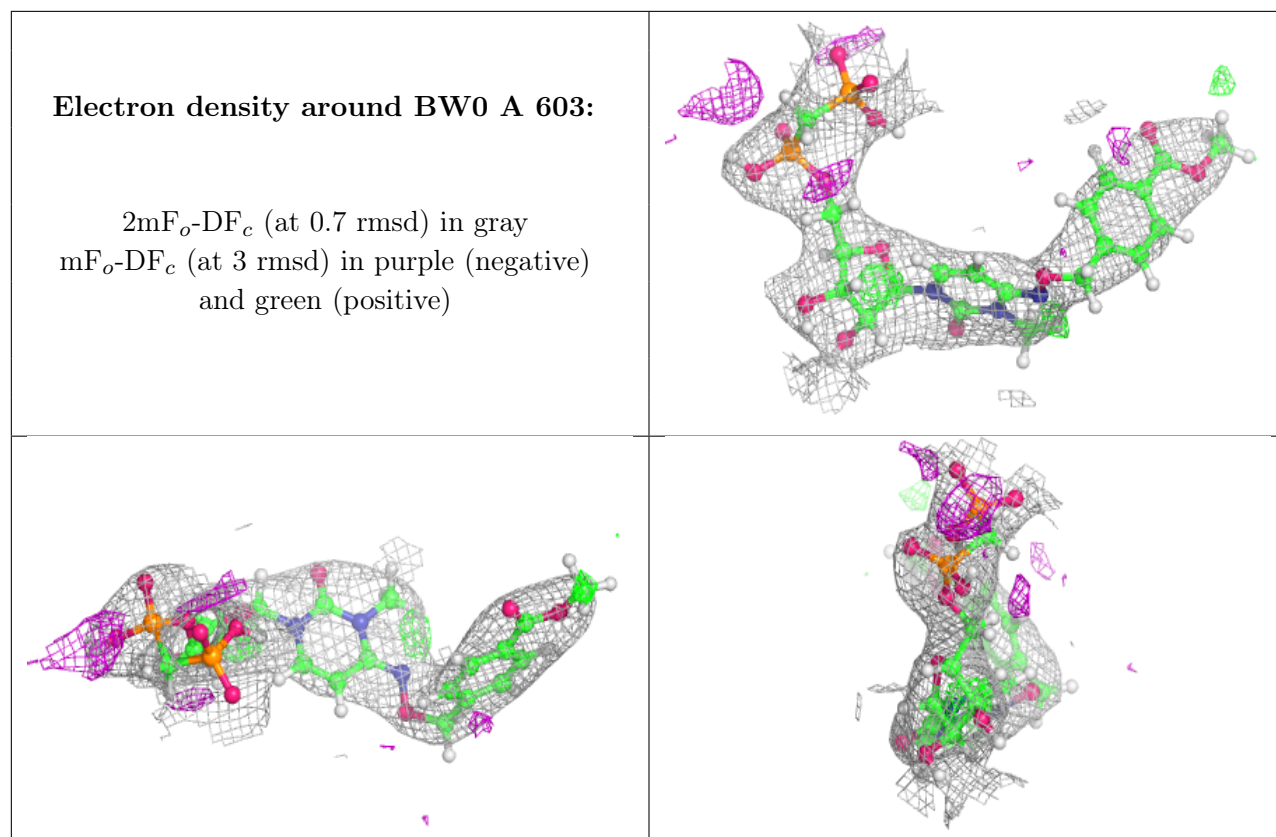
There are no monosaccharides in this entry.

## 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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
3	BW0	A	603	38/38	0.93	0.15	33,66,75,76	64
2	ZN	A	601	1/1	0.98	0.07	79,79,79,79	0
2	ZN	A	602	1/1	0.99	0.07	69,69,69,69	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.



## 6.5 Other polymers [i](#)

There are no such residues in this entry.