

Full wwPDB X-ray Structure Validation Report (i)

Oct 15, 2023 – 01:08 PM EDT

PDB ID	:	8D8Z
Title	:	Crystal structure of ChoE N147A mutant in complex with thiocholine and
		chloride
Authors	:	Pham, V.D.; Shi, R.
Deposited on		
Resolution	:	1.38 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 (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 (1)) were used in the production of this report:

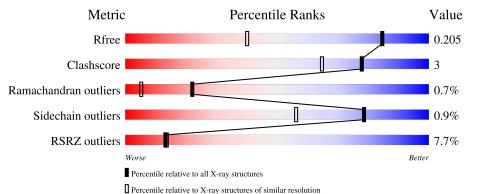
MolProbity Mogul Xtriage (Phenix) EDS	:	4.02b-467 1.8.5 (274361), CSD as541be (2020) 1.13 2.36
buster-report Percentile statistics Refmac	: : :	1.1.7 (2018) 20191225.v01 (using entries in the PDB archive December 25th 2019) 5.8.0158 7.0.044 (Gargrove)
Ideal geometry (DNA, RNA) Validation Pipeline (wwPDB-VP)		Parkinson et al. (1996) 2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure: $X\text{-}RAY \, DIFFRACTION$

The reported resolution of this entry is 1.38 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	$egin{array}{c} { m Whole \ archive} \ (\#{ m Entries}) \end{array}$	${f Similar\ resolution}\ (\#{ m Entries,\ resolution\ range}({ m \AA}))$
R_{free}	130704	2907 (1.40-1.36)
Clashscore	141614	3037 (1.40-1.36)
Ramachandran outliers	138981	2970 (1.40-1.36)
Sidechain outliers	138945	2969 (1.40-1.36)
RSRZ outliers	127900	2846 (1.40-1.36)

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	А	287	89%	11%
1	В	287	92%	8%



2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 5199 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 ChoE.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace		
1	А	287		C 1412		0 412	${ m S}{ m 5}$	0	4	0
1	В	287	Total 2259	C 1423	N 419	0 412	$\frac{S}{5}$	0	7	0

There are 2 discrepancies between the modelled and reference sequences:

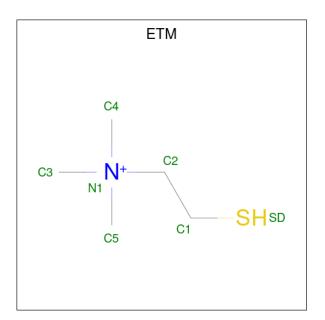
Chain	Residue	Modelled	Actual	Comment	Reference
А	147	ALA	ASN	engineered mutation	UNP Q9HUP2
В	147	ALA	ASN	engineered mutation	UNP Q9HUP2

• Molecule 2 is CHLORIDE ION (three-letter code: CL) (formula: Cl) (labeled as "Ligand of Interest" by depositor).

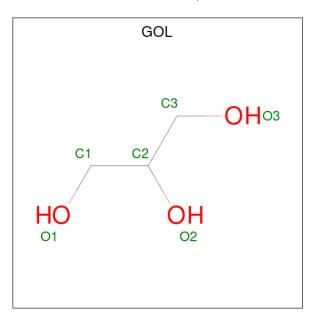
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	А	1	Total Cl 1 1	0	0
2	В	1	Total Cl 1 1	0	0

• Molecule 3 is 2-(TRIMETHYLAMMONIUM)ETHYL THIOL (three-letter code: ETM) (formula: C₅H₁₄NS) (labeled as "Ligand of Interest" by depositor).





Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	А	1	$\begin{array}{cccc} \text{Total} & \text{C} & \text{N} & \text{S} \\ 7 & 5 & 1 & 1 \end{array}$	0	0
3	В	1	$\begin{array}{cccc} \text{Total} & \text{C} & \text{N} & \text{S} \\ 7 & 5 & 1 & 1 \end{array}$	0	0



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	А	1	$\begin{array}{ccc} \text{Total} & \text{C} & \text{O} \\ 6 & 3 & 3 \end{array}$	0	0
4	В	1	$\begin{array}{ccc} \text{Total} & \text{C} & \text{O} \\ 6 & 3 & 3 \end{array}$	0	0



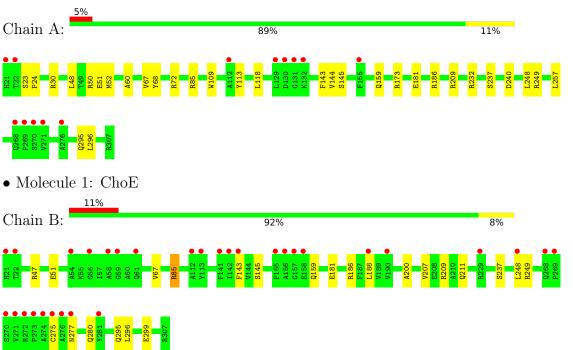
• Molecule 5 is water.

Mol	Chain	Residues	esidues Atoms		AltConf
5	А	368	Total O 368 368	0	0
5	В	298	Total O 298 298	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: ChoE



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants	45.67Å 81.80Å 81.10Å	Depositor
a, b, c, α , β , γ	90.00° 99.58° 90.00°	Depositor
Resolution (Å)	35.95 - 1.38	Depositor
Resolution (A)	35.92 - 1.38	EDS
% Data completeness	91.3 (35.95-1.38)	Depositor
(in resolution range)	91.3 (35.92-1.38)	EDS
R _{merge}	(Not available)	Depositor
R _{sym}	0.04	Depositor
$< I/\sigma(I) > 1$	$1.88 (at 1.38 \text{\AA})$	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
B B.	0.168 , 0.200	Depositor
R, R_{free}	0.176 , 0.205	DCC
R_{free} test set	5476 reflections (4.97%)	wwPDB-VP
Wilson B-factor $(Å^2)$	15.5	Xtriage
Anisotropy	0.556	Xtriage
Bulk solvent $k_{sol}(e/Å^3), B_{sol}(Å^2)$	0.37, 45.0	EDS
L-test for twinning ²	$ \langle L \rangle = 0.50, \langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	5199	wwPDB-VP
Average B, all atoms $(Å^2)$	21.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 6.17% of the height of the origin peak. No significant pseudotranslation is detected.

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.



¹Intensities estimated from amplitudes.

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: GOL, CL, ETM

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	Bo	nd lengths	Bond angles	
IVIOI	Unain	RMSZ	# Z > 5	RMSZ	# Z > 5
1	А	0.79	2/2312~(0.1%)	0.99	10/3145~(0.3%)
1	В	0.79	0/2337	0.96	5/3178~(0.2%)
All	All	0.79	2/4649~(0.0%)	0.98	15/6323~(0.2%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	А	237[A]	SER	CA-CB	-5.32	1.45	1.52
1	А	237[B]	SER	CA-CB	-5.32	1.45	1.52

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
1	А	186	ARG	NE-CZ-NH2	-7.43	116.58	120.30
1	А	173	ARG	NE-CZ-NH2	7.40	124.00	120.30
1	В	186	ARG	NE-CZ-NH2	-7.23	116.68	120.30
1	А	249	ARG	NE-CZ-NH2	-6.71	116.95	120.30
1	А	209	ARG	NE-CZ-NH1	-6.38	117.11	120.30
1	В	249	ARG	NE-CZ-NH1	6.33	123.47	120.30
1	А	209	ARG	NE-CZ-NH2	5.89	123.25	120.30
1	А	240	ASP	CB-CG-OD2	-5.81	113.07	118.30
1	А	249	ARG	NE-CZ-NH1	5.69	123.15	120.30
1	А	30	ARG	NE-CZ-NH2	-5.61	117.49	120.30
1	А	232	ARG	NE-CZ-NH1	-5.50	117.55	120.30
1	В	209	ARG	NE-CZ-NH2	5.48	123.04	120.30
1	В	209	ARG	NE-CZ-NH1	-5.31	117.64	120.30
1	В	85	ARG	NE-CZ-NH2	-5.18	117.71	120.30
1	А	30	ARG	NE-CZ-NH1	5.07	122.84	120.30

All (15) bond angle outliers are listed below:

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	А	2246	0	2177	13	0
1	В	2259	0	2206	13	0
2	А	1	0	0	0	0
2	В	1	0	0	0	0
3	А	7	0	14	0	0
3	В	7	0	14	0	0
4	А	6	0	8	0	0
4	В	6	0	8	1	0
5	А	368	0	0	2	0
5	В	298	0	0	2	0
All	All	5199	0	4427	26	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.

All (26) 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]:LEU:HD11	1:B:296:LEU:HD13	1.64	0.79
1:B:188[B]:LEU:HD13	1:B:237[B]:SER:HB3	1.72	0.69
1:B:277:ASN:ND2	1:B:280:GLN:OE1	2.25	0.67
1:A:159:GLN:HB2	5:A:632:HOH:O	1.98	0.64
1:B:275:CYS:HB3	5:B:737:HOH:O	1.98	0.63
1:B:159:GLN:HB2	5:B:520:HOH:O	2.01	0.60
1:B:277:ASN:OD1	1:B:277:ASN:N	2.35	0.59
1:B:47:ARG:NH1	1:B:51:GLU:OE2	2.38	0.57
1:B:295[A]:GLN:NE2	1:B:299:GLU:OE2	2.24	0.54
1:A:85:ARG:NH2	1:A:295[A]:GLN:OE1	2.43	0.51
1:B:188[B]:LEU:CD1	1:B:237[B]:SER:HB3	2.39	0.48
1:B:188[B]:LEU:HD13	1:B:237[B]:SER:CB	2.42	0.48
1:A:248[B]:LEU:HD11	1:A:296:LEU:HD13	1.96	0.48
1:A:248[A]:LEU:HD21	1:A:296:LEU:HD13	1.97	0.47



Atom-1	Atom-1 Atom-2		Clash overlap (Å)
1:A:51:GLU:OE1	1:A:113:TYR:OH	2.31	0.46
1:B:85:ARG:NH2	1:B:295[A]:GLN:OE1	2.48	0.46
1:A:23[A]:SER:HB2	1:A:24:PRO:CD	2.46	0.45
1:A:68:TYR:HB3	1:A:72:ARG:O	2.16	0.45
1:A:50:ARG:NE	5:A:515:HOH:O	2.50	0.43
1:B:200:ALA:HB1	4:B:403:GOL:H11	2.01	0.42
1:B:207:VAL:O	1:B:211:GLN:HG3	2.20	0.42
1:A:23[A]:SER:HB2	1:A:24:PRO:HD2	2.01	0.42
1:A:248[A]:LEU:HD22	1:A:257:LEU:HD12	2.01	0.42
1:A:118:LEU:HD22	1:A:144:VAL:HG11	2.02	0.41
1:A:48:LEU:O	1:A:52:MET:HG3	2.20	0.41
1:A:60:ALA:HA	1:A:109:TRP:CZ3	2.55	0.41

Continued from previous page...

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	А	289/287~(101%)	279~(96%)	8(3%)	2(1%)	22 5
1	В	292/287~(102%)	282~(97%)	8(3%)	2(1%)	22 5
All	All	581/574~(101%)	561 (97%)	16 (3%)	4 (1%)	22 5

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	А	145	SER
1	В	145	SER
1	А	67	VAL
1	В	67	VAL



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	А	224/220~(102%)	222~(99%)	2(1%)	78 56
1	В	227/220~(103%)	225~(99%)	2(1%)	78 56
All	All	451/440~(102%)	447~(99%)	4 (1%)	78 56

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	А	143	PHE
1	А	181	GLU
1	В	143	PHE
1	В	181	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	А	159	GLN
1	А	215	GLN
1	В	159	GLN
1	В	163	GLN
1	В	215	GLN

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, 2 are monoatomic - leaving 4 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		B	ond leng	gths	B	ond ang	gles
	туре	Unam	nes	LIIIK	Counts	RMSZ	# Z >2	Counts	RMSZ	# Z >2
4	GOL	В	403	-	$5,\!5,\!5$	0.26	0	$5,\!5,\!5$	0.58	0
3	ETM	А	402	-	$6,\!6,\!6$	0.30	0	7,8,8	0.29	0
3	ETM	В	402	-	$6,\!6,\!6$	0.33	0	7,8,8	0.40	0
4	GOL	А	403	-	$5,\!5,\!5$	0.23	0	$5,\!5,\!5$	0.51	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
4	GOL	В	403	-	-	1/4/4/4	-
3	ETM	А	402	-	-	0/4/4/4	-
3	ETM	В	402	-	-	1/4/4/4	-
4	GOL	А	403	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	В	402	ETM	SD-C1-C2-N1
4	А	403	GOL	O1-C1-C2-O2
4	А	403	GOL	O1-C1-C2-C3



Continued from previous page...

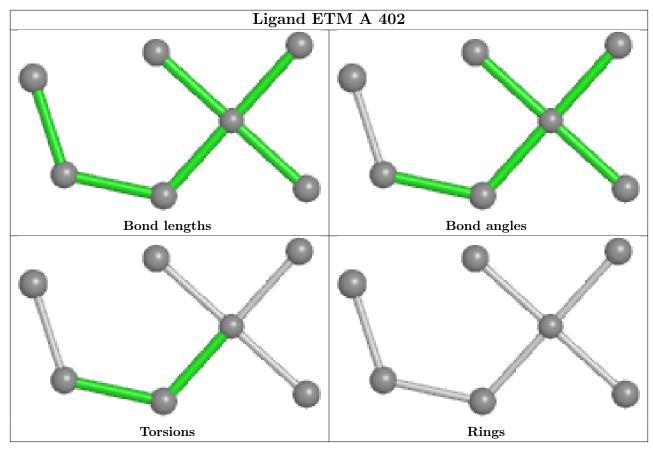
Mol	Chain	Res	Type	Atoms
4	В	403	GOL	C1-C2-C3-O3

There are no ring outliers.

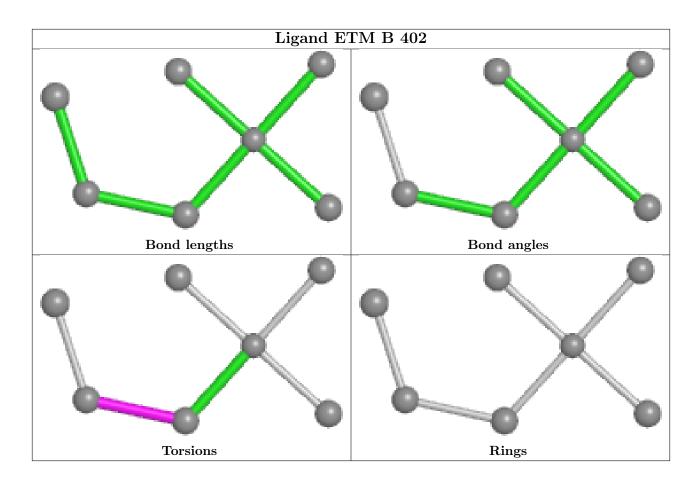
1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	В	403	GOL	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.







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, 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	$\langle RSRZ \rangle$	#RSRZ>2	$OWAB(Å^2)$	Q < 0.9
1	А	287/287~(100%)	0.32	13 (4%) 33 35	9, 16, 30, 50	0
1	В	287/287~(100%)	0.61	31 (10%) 5 6	13, 19, 40, 63	0
All	All	574/574~(100%)	0.46	44 (7%) 13 14	9, 18, 37, 63	0

All (44) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	А	131	GLY	6.3
1	В	158	GLU	5.9
1	В	269	PRO	5.4
1	В	21	HIS	5.2
1	В	276	ALA	5.1
1	В	273	PRO	4.7
1	А	130	ASP	4.4
1	В	112	ALA	4.1
1	В	270	SER	3.7
1	В	268	GLN	3.7
1	А	269	PRO	3.6
1	В	275	CYS	3.3
1	А	21	HIS	3.2
1	В	272	ARG	3.2
1	В	59	GLY	3.1
1	В	113	TYR	3.1
1	В	22	THR	3.0
1	В	274	ALA	3.0
1	В	188[A]	LEU	2.9
1	А	268	GLN	2.9
1	В	277	ASN	2.7
1	А	270	SER	2.7
1	В	143	PHE	2.7
1	В	155	PHE	2.7



Mol	Chain	Res	Type	RSRZ
1	А	112	ALA	2.6
1	В	54	ALA	2.5
1	А	271	VAL	2.5
1	В	142	ILE	2.5
1	А	129	LEU	2.5
1	А	155	PHE	2.5
1	А	132	LYS	2.4
1	В	156	ALA	2.4
1	В	248[A]	LEU	2.4
1	В	190	VAL	2.4
1	А	22	THR	2.3
1	В	56	GLY	2.2
1	В	281	TYR	2.1
1	В	141	PHE	2.1
1	В	229	ARG	2.1
1	В	271	VAL	2.1
1	А	276	ALA	2.1
1	В	157	GLY	2.1
1	В	58	ALA	2.1
1	В	61	GLN	2.0

Continued from previous page...

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	$\mathbf{B} ext{-factors}(\mathrm{\AA}^2)$	Q<0.9
4	GOL	В	403	6/6	0.71	0.18	36,44,49,50	0
3	ETM	В	402	7/7	0.81	0.20	41,42,44,45	0
3	ETM	А	402	7/7	0.83	0.19	33,33,37,39	0



2

CL

0

Contr							
Mol	Type	Chain	Res	Atoms	RSCC	RSR	$B-factors(Å^2)$
4	GOL	А	403	6/6	0.85	0.18	34,39,42,42
2	CL	В	401	1/1	0.97	0.05	23,23,23,23

1/1

401

Continued from previous page...

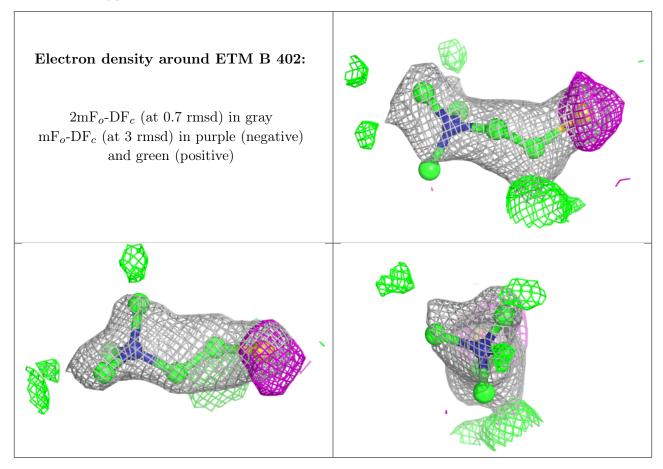
А

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.

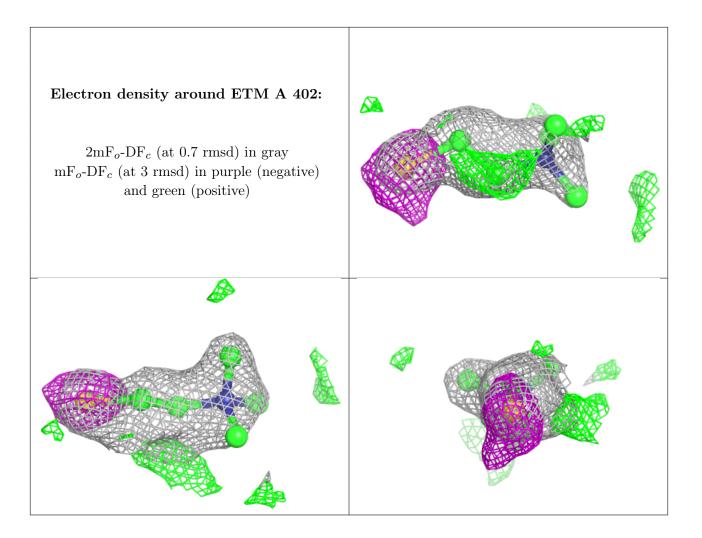
0.03

19, 19, 19, 19, 19

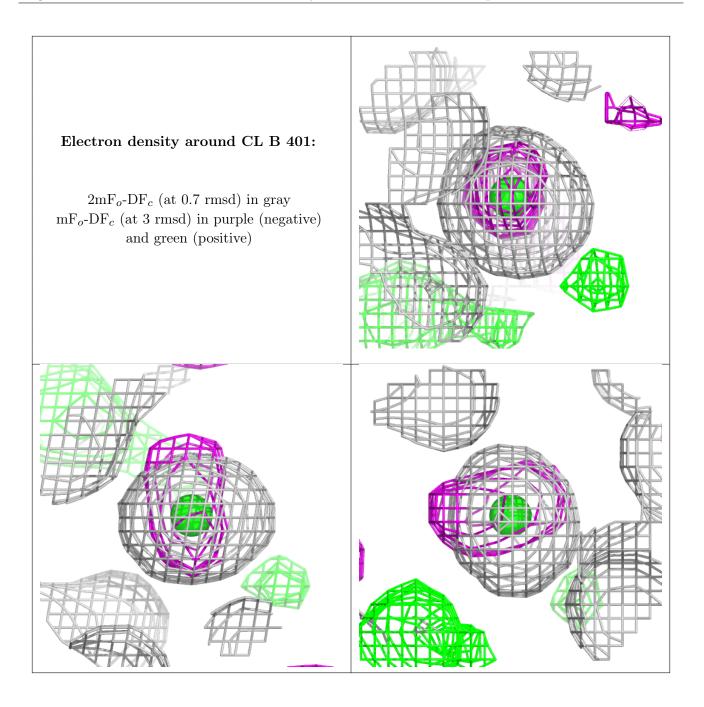
0.99



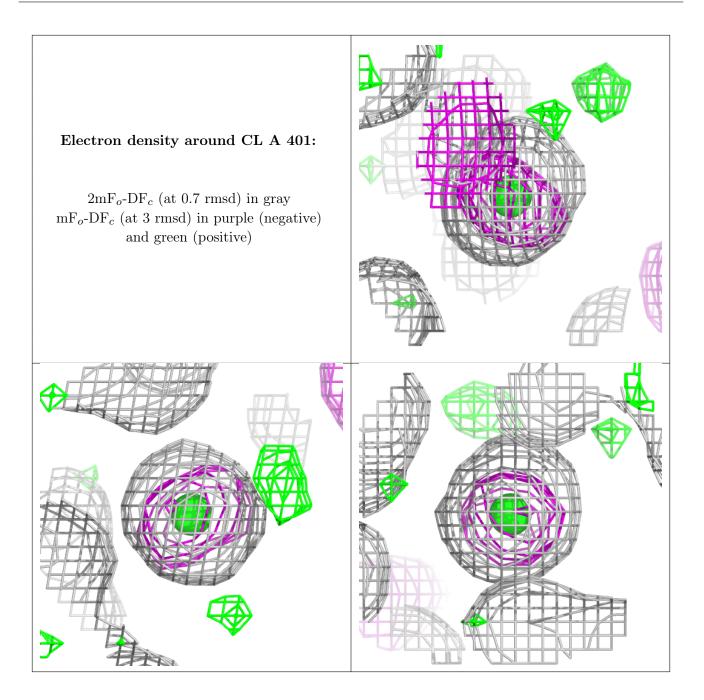












6.5 Other polymers (i)

There are no such residues in this entry.

