Full wwPDB X-ray Structure Validation Report (i)

Sep 13, 2020 – 09:33 PM BST

PDB ID	:	5LM4
Title	:	Structure of the thermostalilized EAAT1 cryst-II mutant in complex with L-ASP and the allosteric inhibitor UCPH101
Authors	:	Canul-Tec, J.; Assal, R.; Legrand, P.; Reyes, N.
Deposited on	:	2016-07-29
Resolution	:	3.10 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 following versions of software and data (see references (i)) were used in the production of this report:

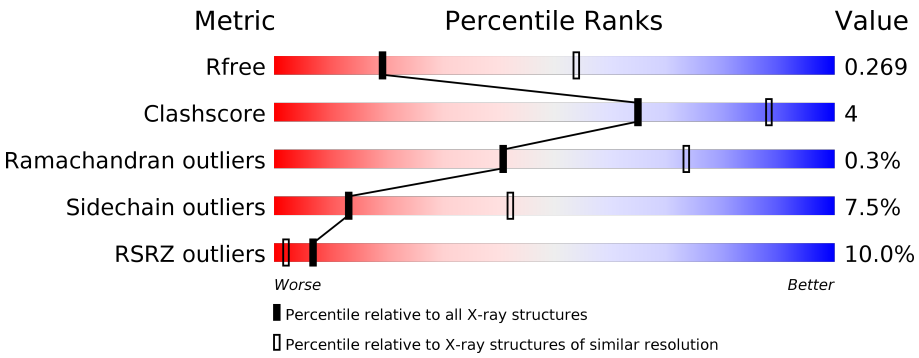
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.14.4.dev1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.14.4.dev1

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 3.10 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	1094 (3.10-3.10)
Clashscore	141614	1184 (3.10-3.10)
Ramachandran outliers	138981	1141 (3.10-3.10)
Sidechain outliers	138945	1141 (3.10-3.10)
RSRZ outliers	127900	1067 (3.10-3.10)

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	522	

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 2992 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 Excitatory amino acid transporter 1,Neutral amino acid transporter B(0),Excitatory amino acid transporter 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	389	2950	1946	471	518	15	0	0	0

There are 76 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	23	SER	ARG	engineered mutation	UNP P43003
A	44	PHE	TYR	engineered mutation	UNP P43003
A	46	ARG	PHE	engineered mutation	UNP P43003
A	50	LEU	PHE	engineered mutation	UNP P43003
A	51	LEU	VAL	engineered mutation	UNP P43003
A	56	LEU	THR	engineered mutation	UNP P43003
A	60	LEU	VAL	engineered mutation	UNP P43003
A	62	VAL	THR	engineered mutation	UNP P43003
A	63	VAL	ILE	engineered mutation	UNP P43003
A	67	LEU	THR	engineered mutation	UNP P43003
A	72	PRO	ARG	engineered mutation	UNP P43003
A	73	LEU	MET	engineered mutation	UNP P43003
A	75	PRO	TYR	engineered mutation	UNP P43003
A	82	ALA	SER	engineered mutation	UNP P43003
A	93	LYS	GLN	engineered mutation	UNP P43003
A	96	ILE	VAL	engineered mutation	UNP P43003
A	101	VAL	ILE	engineered mutation	UNP P43003
A	105	ILE	VAL	engineered mutation	UNP P43003
A	108	LEU	MET	engineered mutation	UNP P43003
A	110	SER	ALA	engineered mutation	UNP P43003
A	113	ALA	SER	engineered mutation	UNP P43003
A	118	ARG	LYS	engineered mutation	UNP P43003
A	119	LEU	MET	engineered mutation	UNP P43003
A	129	SER	THR	engineered mutation	UNP P43003
A	137	LEU	ILE	engineered mutation	UNP P43003
A	141	LEU	ILE	engineered mutation	UNP P43003

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	143	LEU	ILE	engineered mutation	UNP P43003
A	149	ALA	LYS	engineered mutation	UNP P43003
A	155	THR	ASN	engineered mutation	UNP Q15758
A	175	CYS	SER	engineered mutation	UNP Q15758
A	204	THR	ASN	engineered mutation	UNP Q15758
A	223	ILE	ALA	engineered mutation	UNP P43003
A	231	ILE	MET	engineered mutation	UNP P43003
A	232	VAL	CYS	engineered mutation	UNP P43003
A	235	ILE	PHE	engineered mutation	UNP P43003
A	236	ALA	VAL	engineered mutation	UNP P43003
A	237	LEU	ILE	engineered mutation	UNP P43003
A	239	LYS	ASN	engineered mutation	UNP P43003
A	241	GLY	LYS	engineered mutation	UNP P43003
A	246	LEU	ALA	engineered mutation	UNP P43003
A	248	VAL	ARG	engineered mutation	UNP P43003
A	249	ASP	GLU	engineered mutation	UNP P43003
A	252	ASN	ASP	engineered mutation	UNP P43003
A	258	THR	ILE	engineered mutation	UNP P43003
A	260	LYS	ARG	engineered mutation	UNP P43003
A	264	ILE	VAL	engineered mutation	UNP P43003
A	271	LEU	VAL	engineered mutation	UNP P43003
A	287	LEU	MET	engineered mutation	UNP P43003
A	288	GLU	GLY	engineered mutation	UNP P43003
A	290	LEU	ILE	engineered mutation	UNP P43003
A	295	GLY	ALA	engineered mutation	UNP P43003
A	298	MET	THR	engineered mutation	UNP P43003
A	306	VAL	LEU	engineered mutation	UNP P43003
A	309	GLY	ALA	engineered mutation	UNP P43003
A	310	LEU	VAL	engineered mutation	UNP P43003
A	316	ILE	LEU	engineered mutation	UNP P43003
A	320	ILE	VAL	engineered mutation	UNP P43003
A	326	PHE	TRP	engineered mutation	UNP P43003
A	330	ALA	GLY	engineered mutation	UNP P43003
A	332	ILE	LEU	engineered mutation	UNP P43003
A	366	ILE	VAL	engineered mutation	UNP P43003
A	388	VAL	LEU	engineered mutation	UNP P43003
A	399	TYR	PHE	engineered mutation	UNP P43003
A	402	ASP	ASN	engineered mutation	UNP P43003
A	437	ALA	SER	engineered mutation	UNP P43003
A	454	LEU	PHE	engineered mutation	UNP P43003
A	458	PHE	LEU	engineered mutation	UNP P43003
A	461	MET	THR	engineered mutation	UNP P43003

Continued on next page...

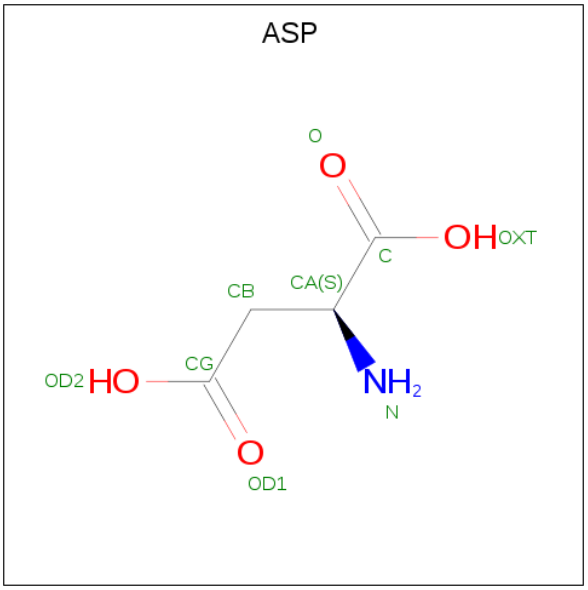
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	462	VAL	THR	engineered mutation	UNP P43003
A	468	ALA	SER	engineered mutation	UNP P43003
A	480	LYS	HIS	engineered mutation	UNP P43003
A	483	GLU	LYS	engineered mutation	UNP P43003
A	484	LYS	ASN	engineered mutation	UNP P43003
A	485	GLN	ARG	engineered mutation	UNP P43003
A	487	ALA	VAL	engineered mutation	UNP P43003
A	489	LEU	MET	engineered mutation	UNP P43003

- Molecule 2 is SODIUM ION (three-letter code: NA) (formula: Na).

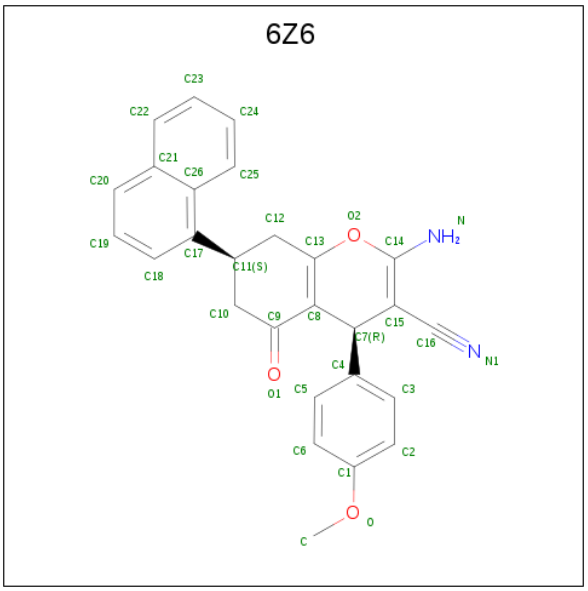
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Na		
2	A	1	1	1	0	0

- Molecule 3 is ASPARTIC ACID (three-letter code: ASP) (formula: $C_4H_7NO_4$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
3	A	1	9	4	1	4	0	0

- Molecule 4 is 2-Amino-5,6,7,8-tetrahydro-4-(4-methoxyphenyl)-7-(naphthalen-1-yl)-5-oxo-4 H-chromene-3-carbonitrile (three-letter code: 6Z6) (formula: $C_{27}H_{22}N_2O_3$).



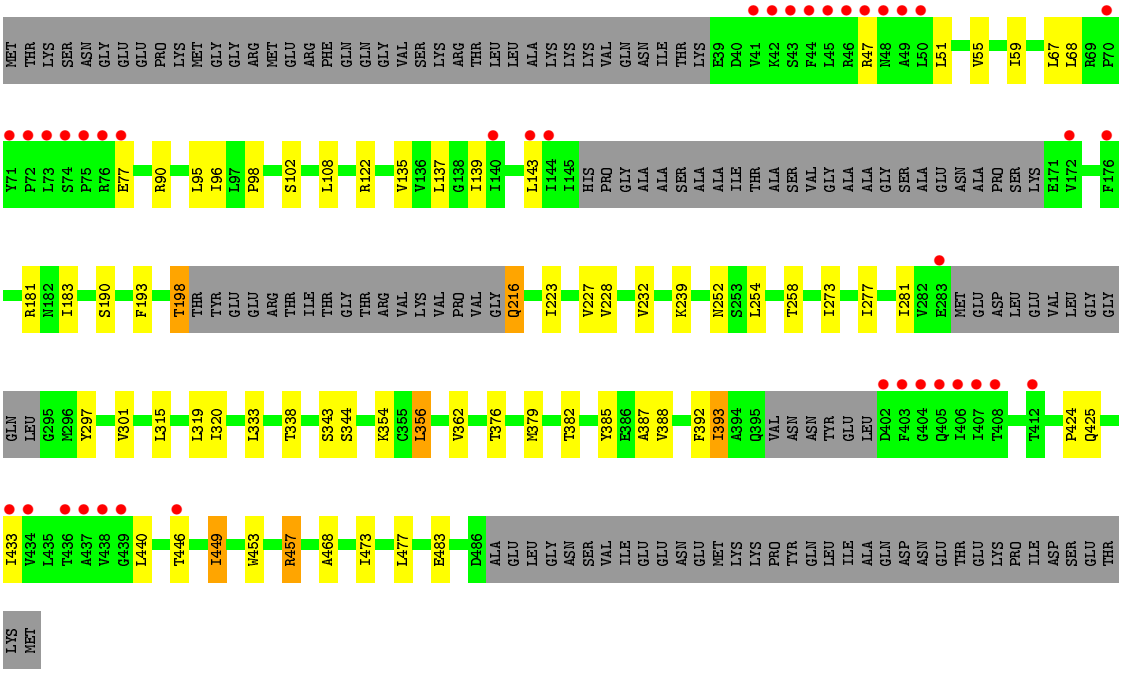
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
4	A	1	32	27	2	3	0	0

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: Excitatory amino acid transporter 1,Neutral amino acid transporter B(0),Excitatory amino acid transporter 1

Chain A:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 63	Depositor
Cell constants a, b, c, α, β, γ	123.11Å 123.11Å 89.62Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	45.80 – 3.10 45.82 – 2.90	Depositor EDS
% Data completeness (in resolution range)	75.8 (45.80-3.10) 65.4 (45.82-2.90)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.55 (at 2.91Å)	Xtriage
Refinement program	BUSTER 2.10.2	Depositor
R, R_{free}	0.217 , 0.259 0.227 , 0.269	Depositor DCC
R_{free} test set	565 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	103.7	Xtriage
Anisotropy	0.019	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 94.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	0.066 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.88	EDS
Total number of atoms	2992	wwPDB-VP
Average B, all atoms (Å^2)	111.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.13% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: NA, 6Z6

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	0.49	0/2993	0.67	0/4065

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	2950	0	3155	25	2
2	A	1	0	0	0	0
3	A	9	0	3	0	0
4	A	32	0	0	0	0
All	All	2992	0	3158	25	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 4.

All (25) 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:356:LEU:HD23	1:A:362:VAL:HG21	1.69	0.74
1:A:344:SER:HB2	1:A:424:PRO:HD3	1.76	0.68
1:A:320:ILE:HG21	1:A:473:ILE:HG23	1.79	0.65
1:A:98:PRO:HB2	1:A:258:THR:HG21	1.79	0.64
1:A:135:VAL:HG23	1:A:387:ALA:HB2	1.84	0.59
1:A:343:SER:HB2	1:A:457:ARG:HD2	1.85	0.58
1:A:102:SER:HB3	1:A:254:LEU:HD23	1.92	0.52
1:A:393:ILE:HD11	1:A:440:LEU:HB3	1.90	0.52
1:A:273:ILE:HD11	1:A:338:THR:HG22	1.93	0.51
1:A:198:THR:HA	1:A:216:GLN:HG3	1.93	0.49
1:A:281:ILE:HG22	1:A:453:TRP:CH2	2.50	0.47
1:A:301:VAL:HG22	1:A:388:VAL:HG11	1.97	0.47
1:A:95:LEU:O	1:A:98:PRO:HD2	2.15	0.46
1:A:183:ILE:HG23	1:A:228:VAL:HG21	1.96	0.46
1:A:190:SER:HA	1:A:193:PHE:CE1	2.50	0.46
1:A:223:ILE:O	1:A:227:VAL:HG23	2.19	0.43
1:A:139:ILE:O	1:A:143:LEU:HG	2.19	0.42
1:A:55:VAL:O	1:A:59:ILE:HG12	2.19	0.42
1:A:379:MET:HB3	1:A:382:THR:HB	2.01	0.42
1:A:228:VAL:O	1:A:232:VAL:HG23	2.19	0.42
1:A:356:LEU:HD11	1:A:468:ALA:HA	2.01	0.42
1:A:446:THR:HA	1:A:449:ILE:HD12	2.01	0.42
1:A:297:TYR:HB2	1:A:392:PHE:CD1	2.56	0.41
1:A:297:TYR:HB2	1:A:392:PHE:CE1	2.55	0.41
1:A:122:ARG:HD3	1:A:477:LEU:HB3	2.02	0.41

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)
1:A:47:ARG:NH2	1:A:77:GLU:OE1[4_0105]	1.75	0.45
1:A:90:ARG:NH2	1:A:181:ARG:O[3_-355]	2.08	0.12

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	379/522 (73%)	364 (96%)	14 (4%)	1 (0%)	41 73

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	449	ILE

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	321/431 (74%)	298 (93%)	23 (7%)	14 44

All (23) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	51	LEU
1	A	67	LEU
1	A	68	LEU
1	A	96	ILE
1	A	108	LEU
1	A	137	LEU
1	A	198	THR
1	A	216	GLN
1	A	239	LYS
1	A	252	ASN
1	A	277	ILE
1	A	315	LEU
1	A	319	LEU
1	A	333	LEU
1	A	354	LYS
1	A	356	LEU
1	A	376	THR
1	A	385	TYR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	393	ILE
1	A	425	GLN
1	A	433	ILE
1	A	457	ARG
1	A	483	GLU

Some 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 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	ASP	A	602	-	2,8,8	0.37	0	1,10,10	0.14	0
4	6Z6	A	603	-	36,36,36	0.37	0	51,52,52	0.71	2 (3%)

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	ASP	A	602	-	-	0/2/8/8	-
4	6Z6	A	603	-	-	1/10/44/44	0/5/5/5

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	603	6Z6	O-C1-C2	2.56	132.09	119.82
4	A	603	6Z6	O-C1-C6	-2.43	108.21	119.82

There are no chirality outliers.

All (1) torsion outliers are listed below:

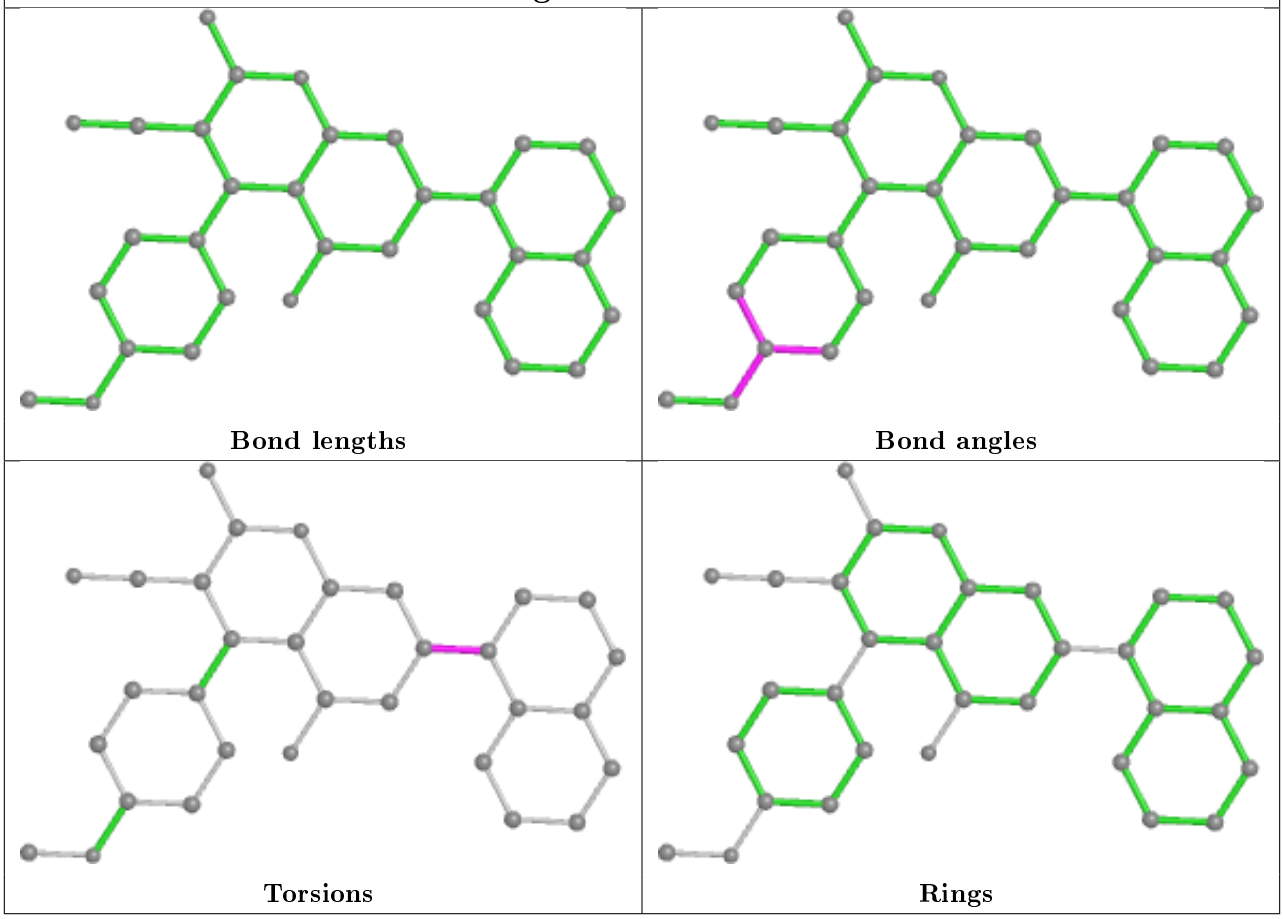
Mol	Chain	Res	Type	Atoms
4	A	603	6Z6	C10-C11-C17-C18

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 6Z6 A 603



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	389/522 (74%)	0.25	39 (10%) 7 2	60, 108, 157, 198	1 (0%)

All (39) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	46	ARG	10.8
1	A	43	SER	8.8
1	A	42	LYS	8.3
1	A	47	ARG	8.2
1	A	405	GLN	8.2
1	A	404	GLY	7.7
1	A	438	VAL	6.7
1	A	45	LEU	6.5
1	A	403	PHE	6.3
1	A	408	THR	5.9
1	A	72	PRO	5.8
1	A	407	ILE	5.4
1	A	402	ASP	5.1
1	A	44	PHE	4.9
1	A	439	GLY	4.7
1	A	73	LEU	4.6
1	A	70	PRO	4.3
1	A	433	ILE	4.2
1	A	48	ASN	4.1
1	A	41	VAL	3.9
1	A	283	GLU	3.9
1	A	71	TYR	3.8
1	A	74	SER	3.7
1	A	50	LEU	3.5
1	A	434	VAL	3.3
1	A	437	ALA	3.1
1	A	49	ALA	3.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	75	PRO	2.9
1	A	77	GLU	2.9
1	A	406	ILE	2.8
1	A	436	THR	2.7
1	A	446	THR	2.7
1	A	76	ARG	2.6
1	A	144	ILE	2.5
1	A	143	LEU	2.2
1	A	412	THR	2.2
1	A	172	VAL	2.1
1	A	176	PHE	2.1
1	A	140	ILE	2.1

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)

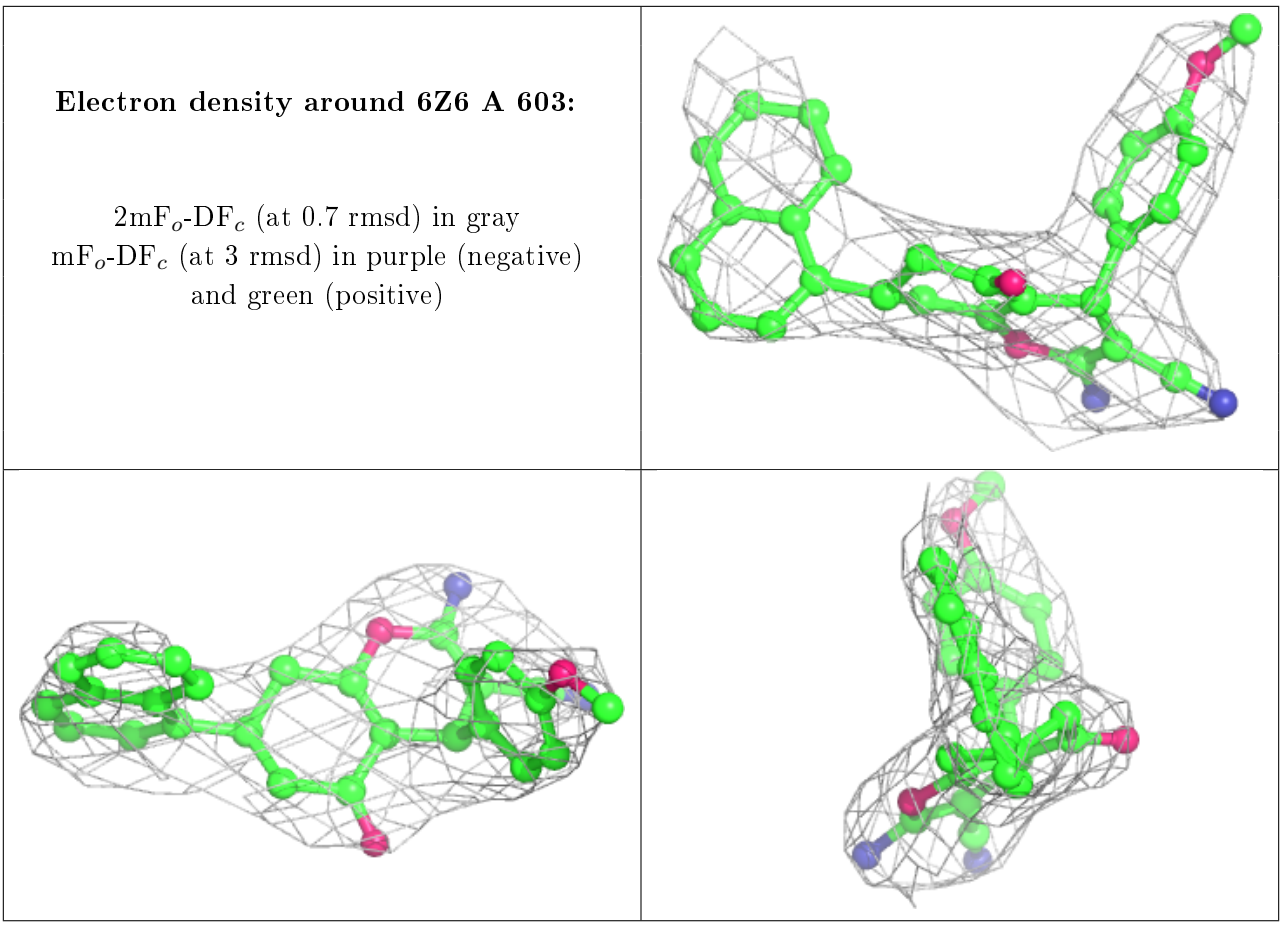
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	NA	A	601	1/1	0.92	0.12	85,85,85,85	0
4	6Z6	A	603	32/32	0.93	0.26	85,95,112,114	0
3	ASP	A	602	9/9	0.97	0.21	98,117,130,146	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 6Z6 A 603:

$2mF_o$-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.