wwPDB X-ray Structure Validation Summary Report (i)

May 26, 2020 – 07:42 am BST

PDB ID	:	4WZ8
Title	:	Crystal structure of human-yeast chimera acetyl coA carboxylase CT domain bound to Compound 6
Authors	:	Vajdos, F.F.
Deposited on	:	2014-11-18
Resolution	:	2.23 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 (1)) were used in the production of this report:

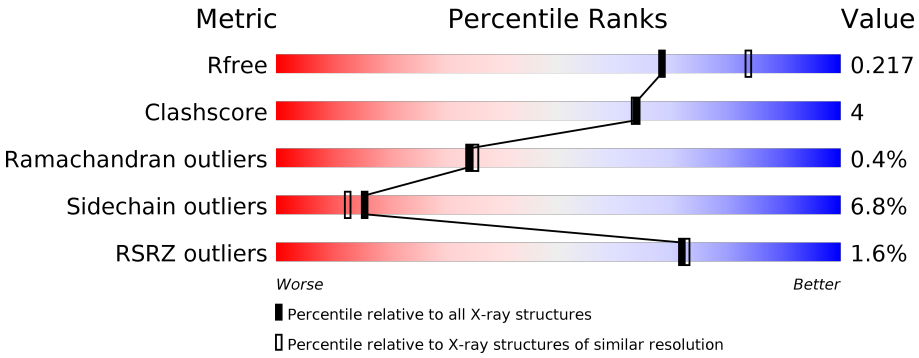
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 2.23 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	2391 (2.26-2.22)
Clashscore	141614	2539 (2.26-2.22)
Ramachandran outliers	138981	2489 (2.26-2.22)
Sidechain outliers	138945	2490 (2.26-2.22)
RSRZ outliers	127900	2353 (2.26-2.22)

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	B	769	2% 78% 12% • 9%
1	C	769	% 77% 12% • 9%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 12728 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 Acetyl-CoA carboxylase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	698	Total	C	N	O	S	0	0	0
			5579	3548	964	1050	17			
1	C	698	Total	C	N	O	S	0	0	0
			5570	3536	965	1052	17			

There are 40 discrepancies between the modelled and reference sequences:

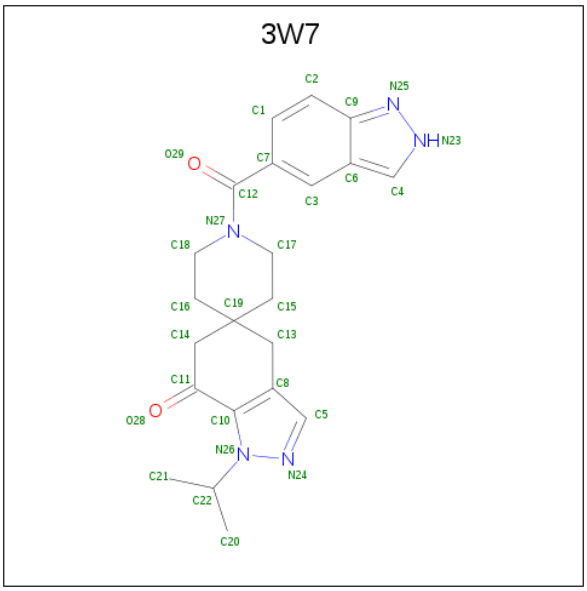
Chain	Residue	Modelled	Actual	Comment	Reference
B	1473	MET	-	initiating methionine	UNP Q00955
B	1474	ALA	-	expression tag	UNP Q00955
B	1475	SER	-	expression tag	UNP Q00955
B	1760	SER	PRO	engineered mutation	UNP Q00955
B	1762	LEU	ILE	engineered mutation	UNP Q00955
B	1765	VAL	MET	engineered mutation	UNP Q00955
B	1919	GLN	GLU	engineered mutation	UNP Q00955
B	1920	ALA	PRO	engineered mutation	UNP Q00955
B	1925	PHE	HIS	engineered mutation	UNP Q00955
B	2028	GLU	GLN	engineered mutation	UNP Q00955
B	2030	THR	MET	engineered mutation	UNP Q00955
B	2032	GLU	GLY	engineered mutation	UNP Q00955
B	2234	LEU	-	expression tag	UNP Q00955
B	2235	GLU	-	expression tag	UNP Q00955
B	2236	HIS	-	expression tag	UNP Q00955
B	2237	HIS	-	expression tag	UNP Q00955
B	2238	HIS	-	expression tag	UNP Q00955
B	2239	HIS	-	expression tag	UNP Q00955
B	2240	HIS	-	expression tag	UNP Q00955
B	2241	HIS	-	expression tag	UNP Q00955
C	1473	MET	-	initiating methionine	UNP Q00955
C	1474	ALA	-	expression tag	UNP Q00955
C	1475	SER	-	expression tag	UNP Q00955
C	1760	SER	PRO	engineered mutation	UNP Q00955
C	1762	LEU	ILE	engineered mutation	UNP Q00955

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	1765	VAL	MET	engineered mutation	UNP Q00955
C	1919	GLN	GLU	engineered mutation	UNP Q00955
C	1920	ALA	PRO	engineered mutation	UNP Q00955
C	1925	PHE	HIS	engineered mutation	UNP Q00955
C	2028	GLU	GLN	engineered mutation	UNP Q00955
C	2030	THR	MET	engineered mutation	UNP Q00955
C	2032	GLU	GLY	engineered mutation	UNP Q00955
C	2234	LEU	-	expression tag	UNP Q00955
C	2235	GLU	-	expression tag	UNP Q00955
C	2236	HIS	-	expression tag	UNP Q00955
C	2237	HIS	-	expression tag	UNP Q00955
C	2238	HIS	-	expression tag	UNP Q00955
C	2239	HIS	-	expression tag	UNP Q00955
C	2240	HIS	-	expression tag	UNP Q00955
C	2241	HIS	-	expression tag	UNP Q00955

- Molecule 2 is 1'-(2H-indazol-5-ylcarbonyl)-1-(propan-2-yl)-1,4-dihydrospiro[indazole-5,4'-piperidin]-7(6H)-one (three-letter code: 3W7) (formula: $C_{22}H_{25}N_5O_2$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	B	1	Total	C	N	O	0	0
			29	22	5	2		
2	C	1	Total	C	N	O	0	0
			29	22	5	2		

- Molecule 3 is water.

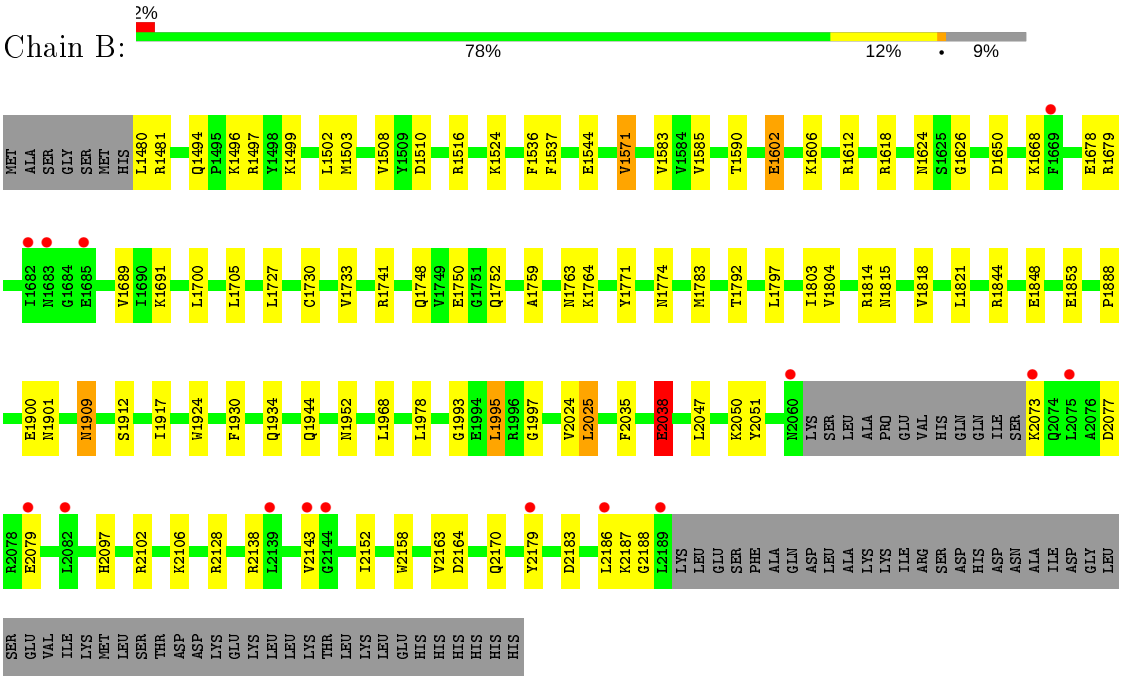
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	743	Total O 743 743	0	0
3	C	778	Total O 778 778	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA 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.

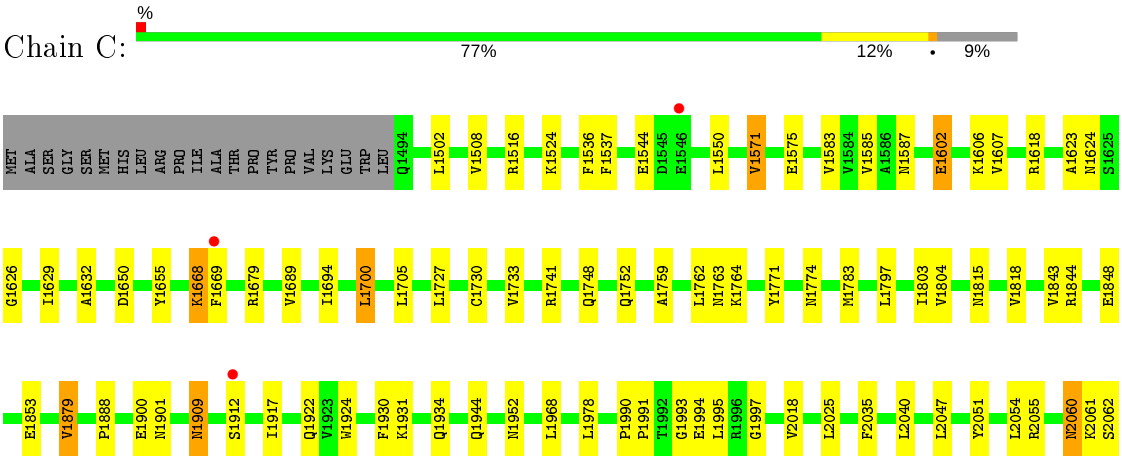
- Molecule 1: Acetyl-CoA carboxylase

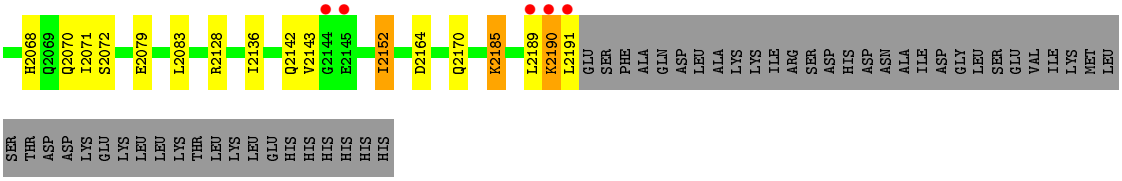
Chain B:



- Molecule 1: Acetyl-CoA carboxylase

Chain C:




H2068
Q2069
Q2070
I2071
S2072
E2079
L2083
R2128
I2136
Q2142
V2143
G2144
E2145
I2152
D2164
Q2170
K2185
L2189
K2190
L2191
GLU
SER
PHE
ALA
GLN
ASP
LEU
ALA
LYS
LYS
ILE
ARG
SER
ASP
HIS
ASP
ASN
ALA
ILE
ASP
GLY
LEU
SER
GLU
VAL
ILE
LYS
MET
LEU
SER
THR
ASP
ASP
LYS
GLU
LYS
LEU
LEU
LYS
THR
LEU
LYS
LEU
GLU
HIS
HIS
HIS
HIS
HIS
HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	94.52Å 138.03Å 184.83Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	110.59 – 2.23 92.42 – 2.23	Depositor EDS
% Data completeness (in resolution range)	97.8 (110.59-2.23) 97.8 (92.42-2.23)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.08	Depositor
$< I/\sigma(I) >$ [1]	2.24 (at 2.22Å)	Xtriage
Refinement program	BUSTER-TNT BUSTER 2.11.5	Depositor
R, R_{free}	0.177 , 0.209 0.182 , 0.217	Depositor DCC
R_{free} test set	5766 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	31.2	Xtriage
Anisotropy	0.532	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 68.1	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	0.95	EDS
Total number of atoms	12728	wwPDB-VP
Average B, all atoms (Å^2)	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.49% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: 3W7

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	B	0.51	0/5699	0.68	0/7718
1	C	0.51	0/5687	0.69	0/7698
All	All	0.51	0/11386	0.69	0/15416

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

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	B	5579	0	5519	51	0
1	C	5570	0	5513	51	0
2	B	29	0	25	1	0
2	C	29	0	25	3	0
3	B	743	0	0	3	0
3	C	778	0	0	3	0
All	All	12728	0	11082	94	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 4.

The worst 5 of 94 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:1763:ASN:HD21	1:B:1771:TYR:H	1.19	0.89
1:C:1763:ASN:HD21	1:C:1771:TYR:H	1.20	0.87
1:B:1764:LYS:HE3	2:C:2301:3W7:H25	1.40	0.84
1:B:2097:HIS:HE1	1:C:1632:ALA:H	1.28	0.80
1:B:1494:GLN:HE21	1:B:1496:LYS:H	1.35	0.74

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	B	694/769 (90%)	672 (97%)	19 (3%)	3 (0%)	34	35
1	C	696/769 (90%)	679 (98%)	15 (2%)	2 (0%)	41	44
All	All	1390/1538 (90%)	1351 (97%)	34 (2%)	5 (0%)	34	35

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	2143	VAL
1	C	2190	LYS
1	C	2143	VAL
1	B	1481	ARG
1	B	2038	GLU

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	B	594/658 (90%)	558 (94%)	36 (6%)	18 16
1	C	594/658 (90%)	549 (92%)	45 (8%)	13 9
All	All	1188/1316 (90%)	1107 (93%)	81 (7%)	16 13

5 of 81 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	1502	LEU
1	C	1606	LYS
1	C	2071	ILE
1	C	1508	VAL
1	C	1571	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 47 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	2170	GLN
1	C	1587	ASN
1	C	2060	ASN
1	C	1517	GLN
1	C	1599	GLN

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 carbohydrates in this entry.

5.6 Ligand geometry (i)

2 ligands are modelled in this entry.

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	$\#\|Z\| > 2$
2	3W7	B	2301	-	31,33,33	1.18	3 (9%)	28,50,50	1.77	6 (21%)
2	3W7	C	2301	-	31,33,33	1.15	4 (12%)	28,50,50	2.04	5 (17%)

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
2	3W7	B	2301	-	-	0/12/38/38	0/5/5/5
2	3W7	C	2301	-	-	0/12/38/38	0/5/5/5

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	2301	3W7	C10-C11	-3.60	1.44	1.50
2	C	2301	3W7	C10-C11	-2.91	1.45	1.50
2	B	2301	3W7	C12-N27	2.30	1.39	1.34
2	B	2301	3W7	C2-C9	-2.10	1.38	1.41
2	C	2301	3W7	C2-C9	-2.09	1.38	1.41

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	2301	3W7	C1-C2-C9	-6.34	112.86	120.84
2	B	2301	3W7	C1-C2-C9	-5.00	114.54	120.84
2	C	2301	3W7	C7-C3-C6	-4.56	114.21	121.24
2	C	2301	3W7	C1-C7-C3	3.95	123.93	119.23
2	B	2301	3W7	C7-C3-C6	-3.89	115.23	121.24

There are no chirality outliers.

There are no torsion outliers.

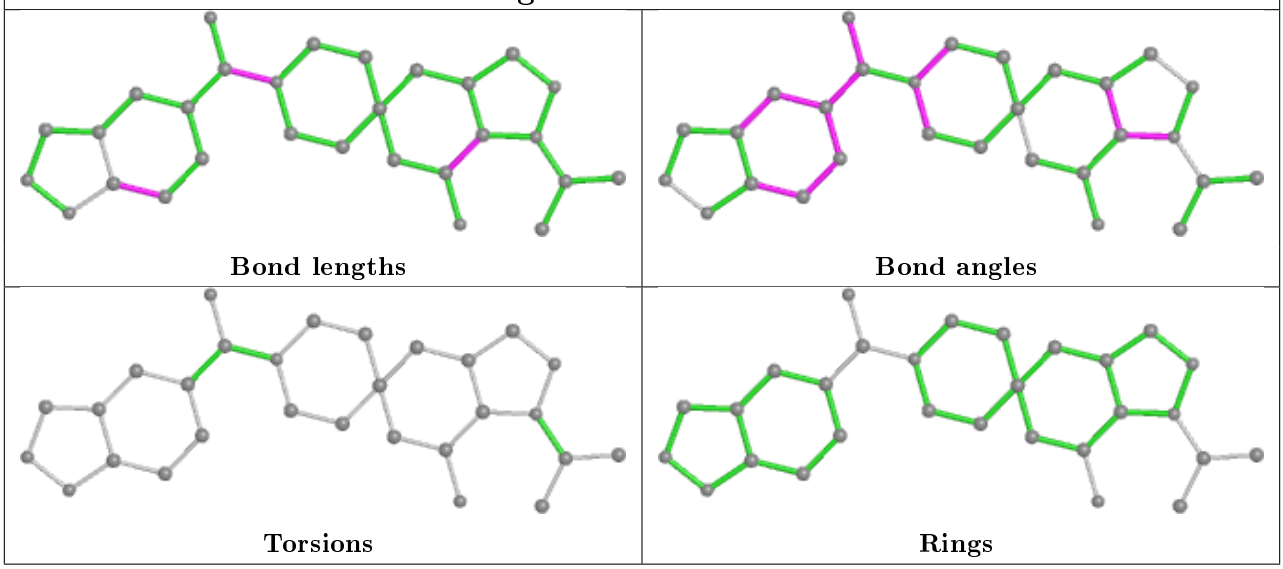
There are no ring outliers.

2 monomers are involved in 4 short contacts:

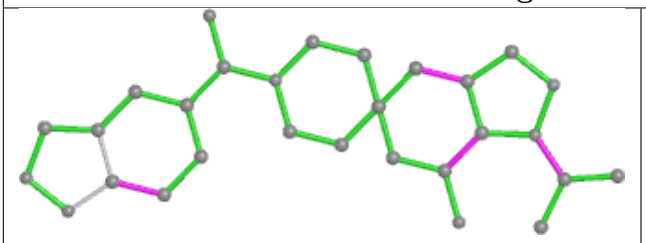
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	2301	3W7	1	0
2	C	2301	3W7	3	0

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.

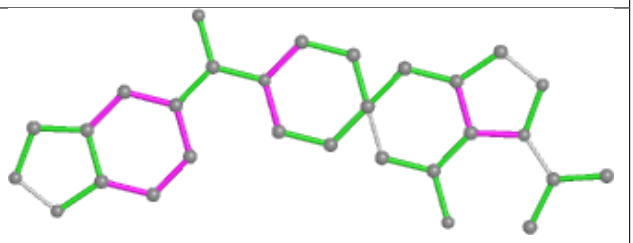
Ligand 3W7 B 2301



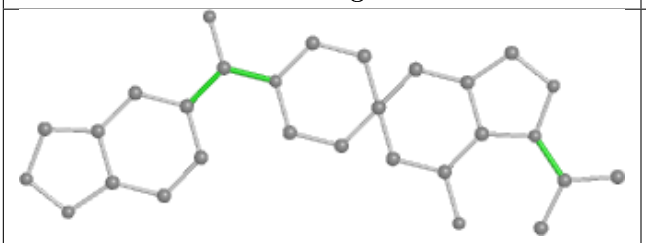
Ligand 3W7 C 2301



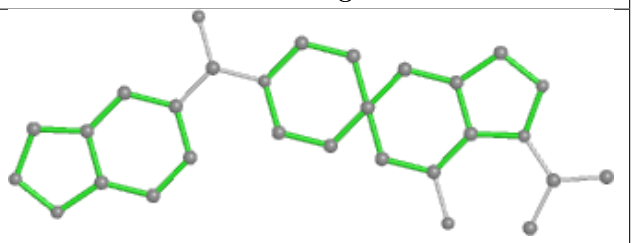
Bond lengths



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	B	698/769 (90%)	-0.09	15 (2%) 63 65	19, 35, 80, 118	0
1	C	698/769 (90%)	-0.18	8 (1%) 80 81	19, 35, 68, 126	0
All	All	1396/1538 (90%)	-0.13	23 (1%) 72 73	19, 35, 74, 126	0

The worst 5 of 23 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	2191	LEU	8.0
1	C	2144	GLY	5.6
1	C	2189	LEU	4.8
1	B	2189	LEU	4.8
1	B	2143	VAL	4.6

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 carbohydrates in this entry.

6.4 Ligands

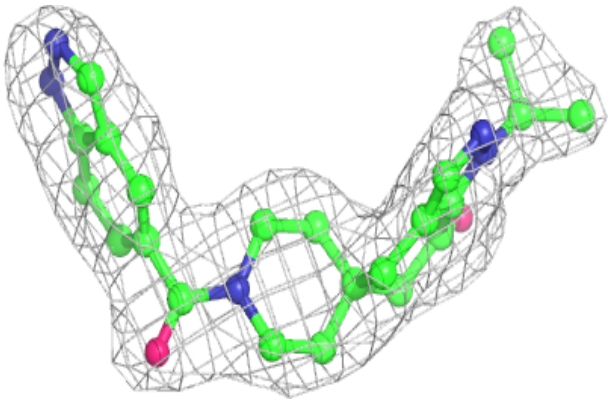
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

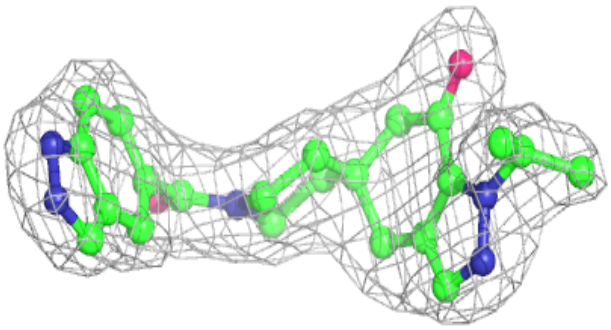
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	3W7	B	2301	29/29	0.97	0.11	27,34,45,46	0
2	3W7	C	2301	29/29	0.97	0.12	24,29,34,36	0

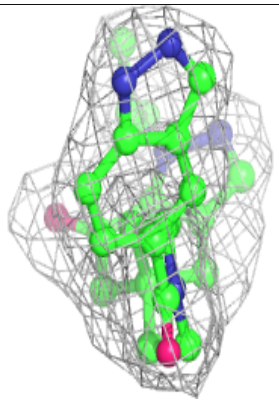
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around 3W7 B 2301:

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

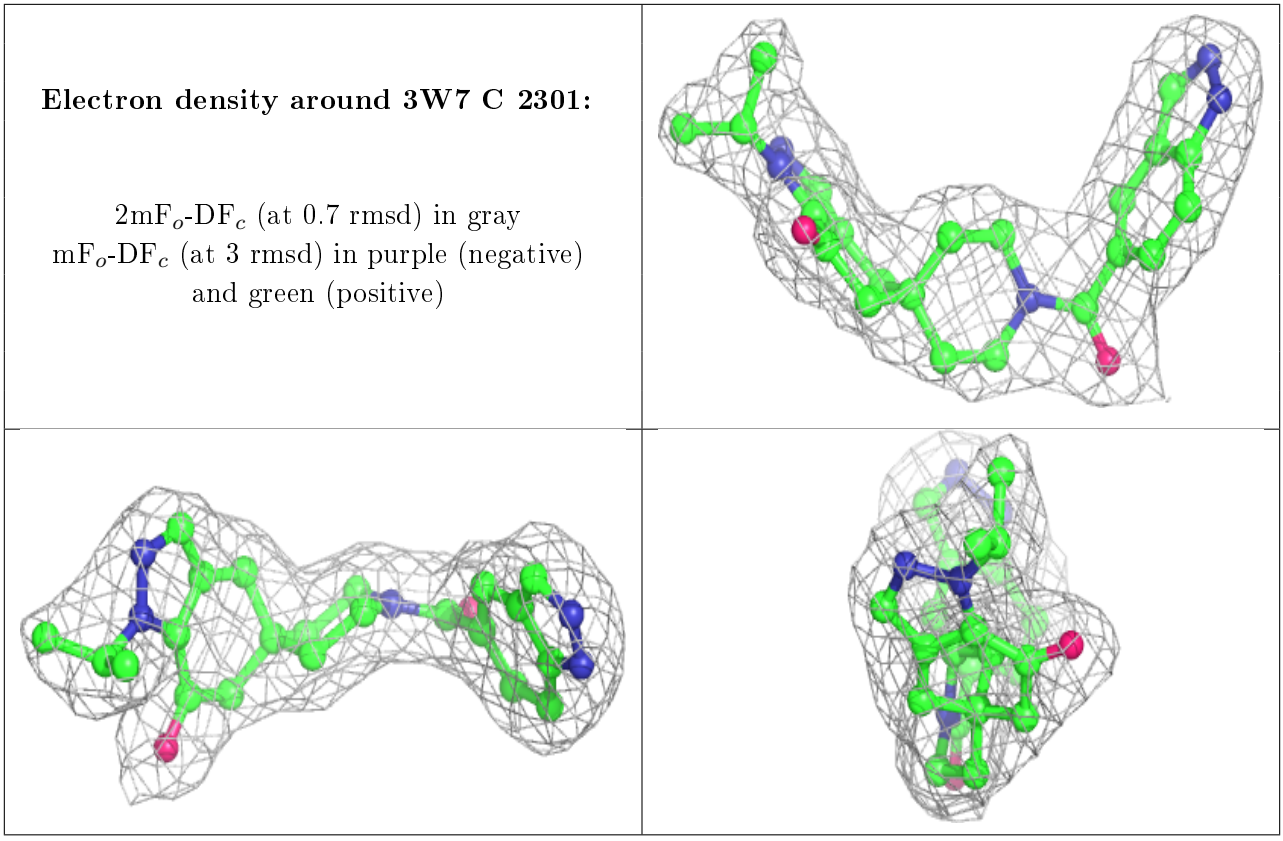






Electron density around 3W7 C 2301:

$2mF_o$-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.