Full wwPDB X-ray Structure Validation Report (i)

May 25, 2020 – 07:46 am BST

PDB ID	:	6YIK
Title	:	Crystal structure of the CREBBP bromodomain in complex with a tetrahydroquinoxaline ligand
Authors	:	Picaud, S.; Brand, M.; Tobias, K.; von Delft, F.; Arrowsmith, C.H.; Edwards, A.M.; Bountra, C.; Conway, S.; Filippakopoulos, P.
Deposited on	:	2020-04-01
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 (i) symbol.

The following versions of software and data (see references (i)) were used in the production of this report:

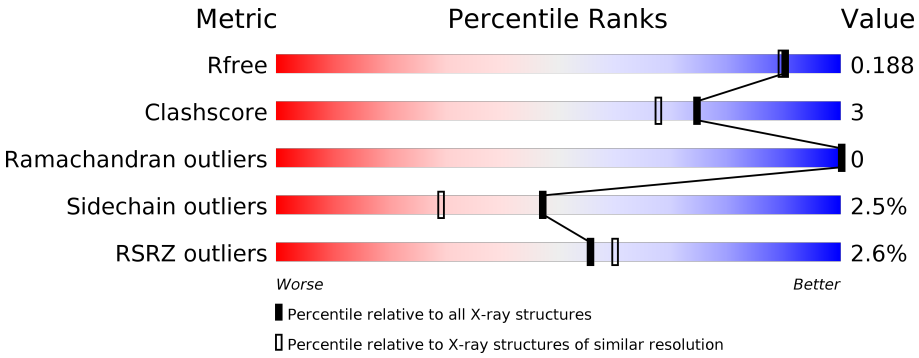
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (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 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}	130704	4298 (1.70-1.70)
Clashscore	141614	4695 (1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (1.70-1.70)
RSRZ outliers	127900	4222 (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 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	119	82% 13%
1	B	119	3% 85% 9%
1	C	119	5% 87% 8% 5%

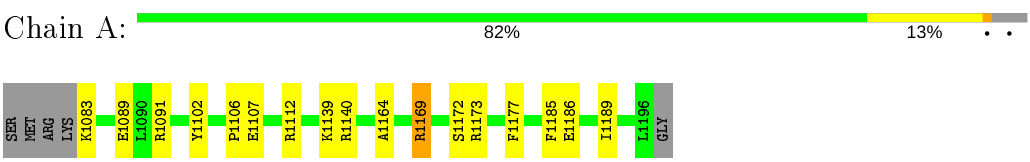
ENTRY-COMPOSITION INFOmissingINFO

2 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.

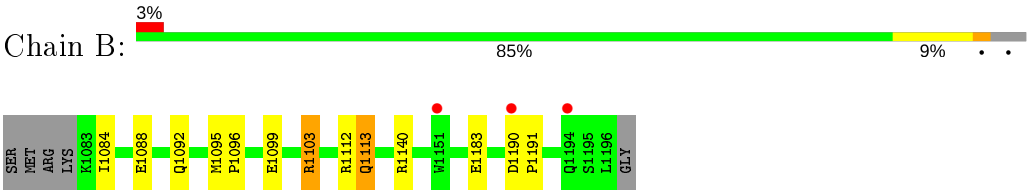
- Molecule 1: CREBBP

Chain A:



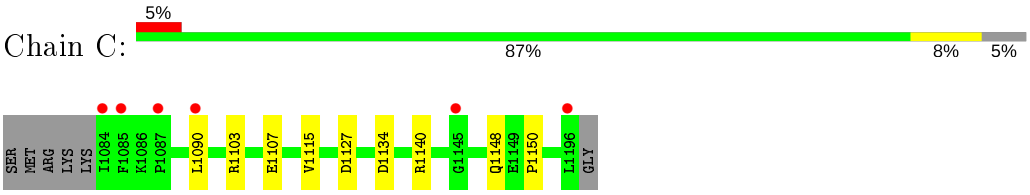
- Molecule 1: CREBBP

Chain B:



- Molecule 1: CREBBP

Chain C:



3 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	80.23Å 43.91Å 88.27Å 90.00° 110.06° 90.00°	Depositor
Resolution (Å)	48.15 – 1.70 48.15 – 1.70	Depositor EDS
% Data completeness (in resolution range)	99.3 (48.15-1.70) 98.9 (48.15-1.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.05	Depositor
$< I/\sigma(I) >$ [1]	2.12 (at 1.70Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.156 , 0.184 0.169 , 0.188	Depositor DCC
R_{free} test set	1958 reflections (3.07%)	wwPDB-VP
Wilson B-factor (Å^2)	26.3	Xtriage
Anisotropy	0.235	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 47.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	3453	wwPDB-VP
Average B, all atoms (Å^2)	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.47% 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.

4 Model quality (i)

4.1 Standard geometry (i)

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

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	1.06	3/1027 (0.3%)	1.34	8/1396 (0.6%)
1	B	1.06	1/993 (0.1%)	1.15	3/1349 (0.2%)
1	C	1.01	1/967 (0.1%)	1.06	3/1318 (0.2%)
All	All	1.04	5/2987 (0.2%)	1.19	14/4063 (0.3%)

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

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1186	GLU	CD-OE2	-8.48	1.16	1.25
1	A	1172	SER	CB-OG	-6.72	1.33	1.42
1	B	1088	GLU	CD-OE1	6.43	1.32	1.25
1	A	1172	SER	CA-CB	-6.22	1.43	1.52
1	C	1107	GLU	CD-OE1	-5.24	1.19	1.25

All (14) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	1091	ARG	NE-CZ-NH2	-11.23	114.69	120.30
1	A	1102[A]	TYR	CB-CA-C	-8.37	93.66	110.40
1	A	1102[B]	TYR	CB-CA-C	-8.37	93.66	110.40
1	A	1173	ARG	NE-CZ-NH1	7.07	123.83	120.30
1	A	1169	ARG	CB-CG-CD	6.62	128.82	111.60

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	1112	ARG	NE-CZ-NH2	6.54	123.57	120.30
1	A	1169	ARG	CD-NE-CZ	6.39	132.55	123.60
1	C	1127	ASP	CB-CG-OD2	-6.27	112.66	118.30
1	A	1091	ARG	NE-CZ-NH1	6.04	123.32	120.30
1	C	1103	ARG	NE-CZ-NH1	5.72	123.16	120.30
1	B	1103	ARG	CB-CG-CD	5.58	126.10	111.60
1	B	1113	GLN	CB-CG-CD	5.38	125.59	111.60
1	A	1169	ARG	NE-CZ-NH2	5.23	122.91	120.30
1	C	1134	ASP	CB-CG-OD1	5.03	122.82	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1112	ARG	Mainchain

4.2 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	986	0	969	7	0
1	B	962	0	946	7	0
1	C	940	0	904	2	0
2	A	90	0	0	5	0
2	B	30	0	0	0	0
2	C	60	0	0	1	0
3	A	150	0	0	2	3
3	B	143	0	0	3	3
3	C	92	0	0	1	0
All	All	3453	0	2819	19	3

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 (19) 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:1140:ARG:HD3	3:A:1429:HOH:O	1.78	0.83
1:B:1183:GLU:CG	3:B:1415:HOH:O	2.34	0.74
1:A:1139:LYS:CE	2:A:1202[A]:OSQ:C12	2.68	0.71
1:A:1164:ALA:HB2	3:A:1407:HOH:O	1.92	0.68
1:B:1190:ASP:N	1:B:1191:PRO:HD2	2.12	0.65
1:B:1092:GLN:NE2	3:B:1302:HOH:O	2.31	0.63
1:A:1139:LYS:HE2	2:A:1202[A]:OSQ:C12	2.32	0.59
1:C:1115:VAL:HG23	3:C:1320:HOH:O	2.02	0.59
1:A:1185:PHE:O	1:A:1189:ILE:HG22	2.10	0.52
1:B:1190:ASP:N	1:B:1191:PRO:CD	2.75	0.49
1:A:1139:LYS:HE3	2:A:1202[A]:OSQ:C12	2.41	0.49
1:B:1140[A]:ARG:HD3	3:B:1342:HOH:O	2.12	0.48
2:A:1202[B]:OSQ:N4	2:A:1202[B]:OSQ:N2	2.65	0.44
1:B:1099:GLU:O	1:B:1103:ARG:HG3	2.19	0.43
2:C:1202:OSQ:O1	2:C:1202:OSQ:N4	2.49	0.43
1:A:1107:GLU:HA	1:A:1177:PHE:CD2	2.53	0.43
2:A:1202[B]:OSQ:C20	2:A:1202[B]:OSQ:F2	2.58	0.42
1:B:1095:MET:N	1:B:1096:PRO:CD	2.83	0.41
1:C:1090:LEU:HD21	1:C:1150:PRO:HD3	2.02	0.41

All (3) 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)
3:A:1423:HOH:O	3:B:1375:HOH:O[2_645]	1.85	0.35
3:A:1334:HOH:O	3:B:1417:HOH:O[2_645]	1.86	0.34
3:A:1414:HOH:O	3:B:1417:HOH:O[2_645]	1.86	0.34

4.3 Torsion angles (i)

4.3.1 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	115/119 (97%)	115 (100%)	0	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	113/119 (95%)	113 (100%)	0	0	100	100
1	C	112/119 (94%)	112 (100%)	0	0	100	100
All	All	340/357 (95%)	340 (100%)	0	0	100	100

There are no Ramachandran outliers to report.

4.3.2 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	111/112 (99%)	107 (96%)	4 (4%)	35	16
1	B	107/112 (96%)	105 (98%)	2 (2%)	57	41
1	C	101/112 (90%)	99 (98%)	2 (2%)	55	38
All	All	319/336 (95%)	311 (98%)	8 (2%)	47	29

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1083	LYS
1	A	1089	GLU
1	A	1106	PRO
1	A	1169	ARG
1	B	1084	ILE
1	B	1113	GLN
1	C	1140	ARG
1	C	1148	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	1113	GLN
1	C	1162	ASN

4.3.3 RNA

There are no RNA molecules in this entry.

4.4 Non-standard residues in protein, DNA, RNA chains

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

4.5 Carbohydrates

There are no carbohydrates in this entry.

4.6 Ligand geometry

Of 6 ligands modelled in this entry, 6 could not be matched to an existing wwPDB Chemical Component Dictionary definition at this stage - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

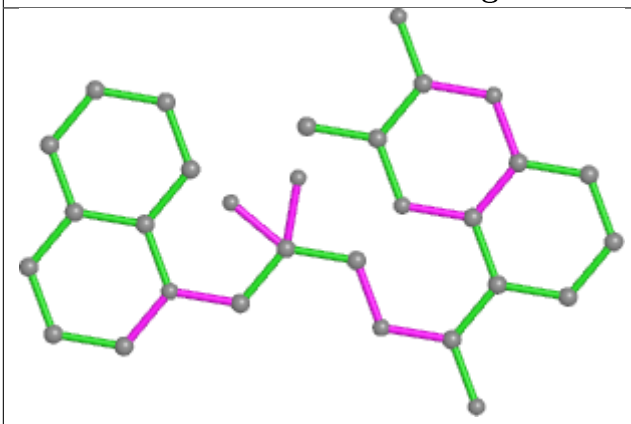
There are no torsion outliers.

There are no ring outliers.

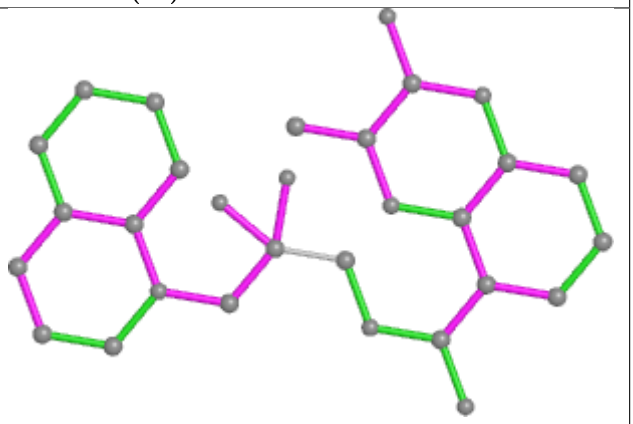
No monomer is involved in short contacts.

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 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.

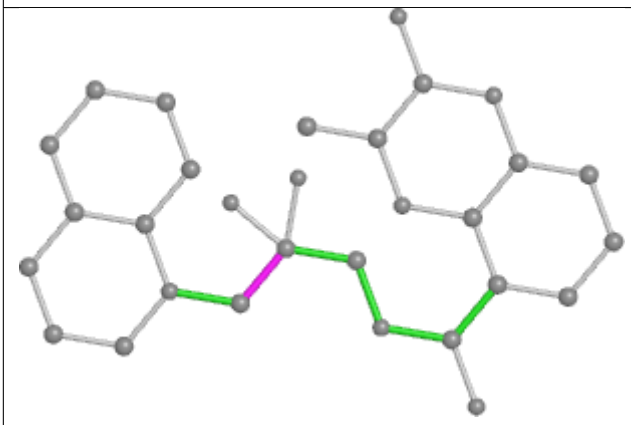
Ligand OSQ A 1202 (A)



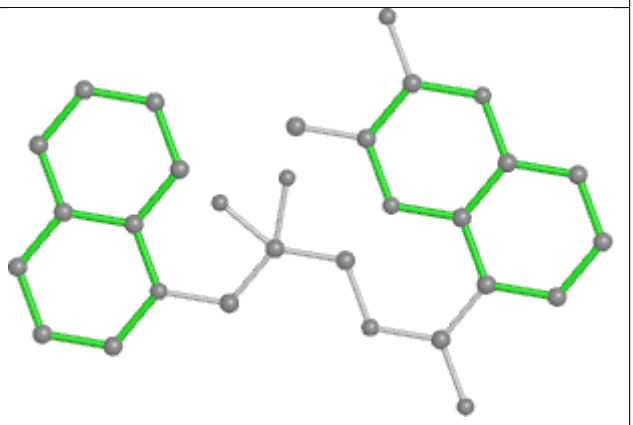
Bond lengths



Bond angles

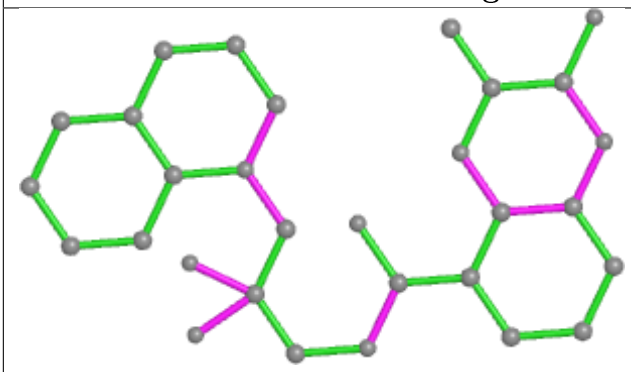


Torsions

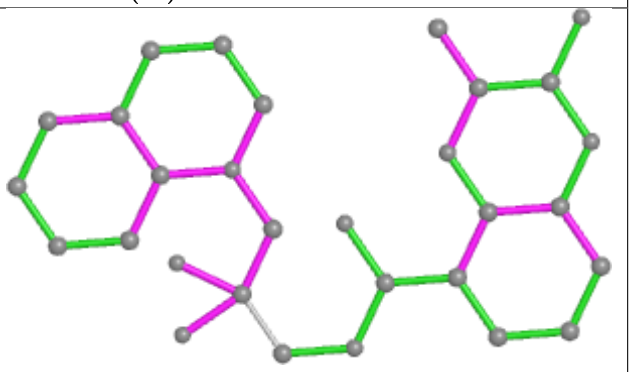


Rings

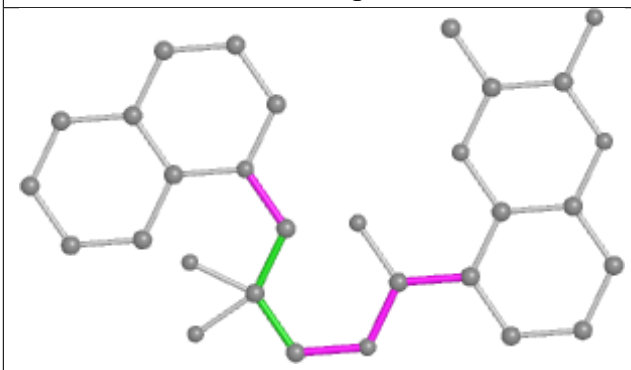
Ligand OSQ A 1202 (B)



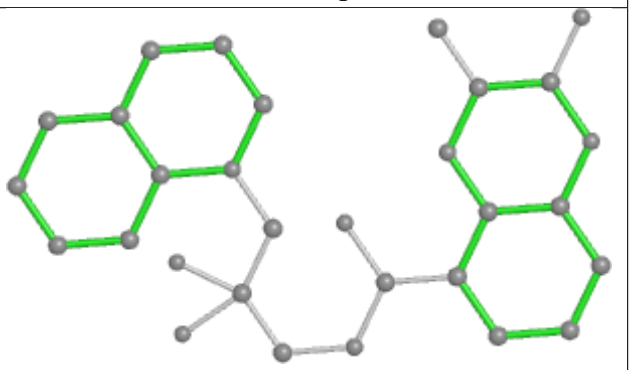
Bond lengths



Bond angles

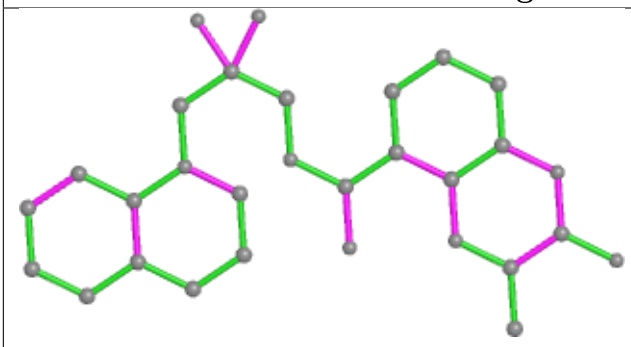


Torsions

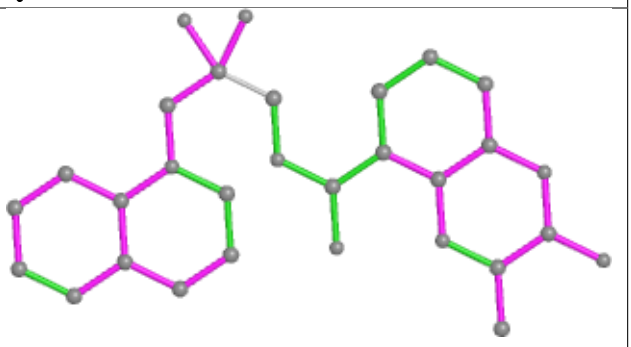


Rings

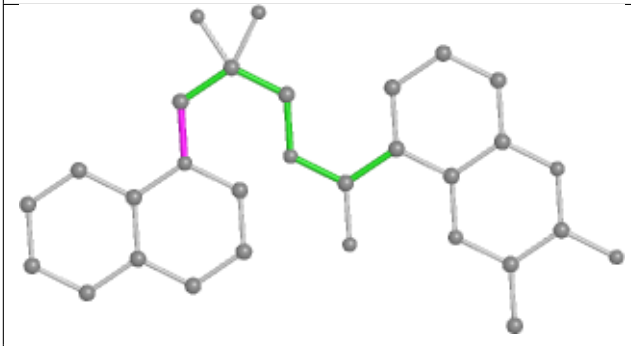
Ligand OSQ B 1201



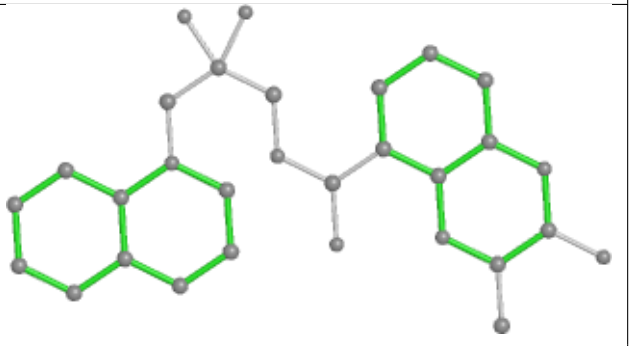
Bond lengths



Bond angles

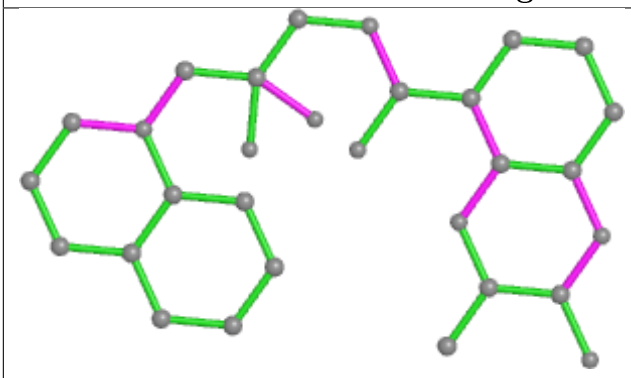


Torsions

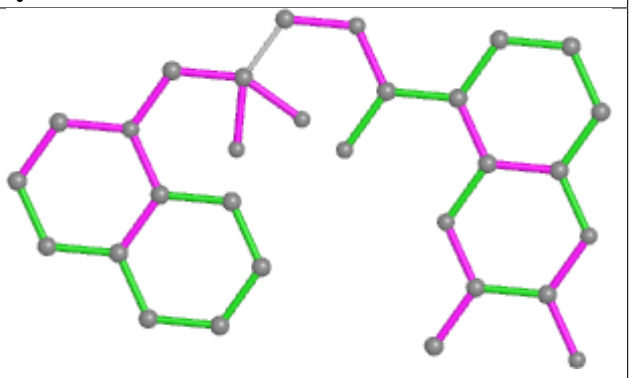


Rings

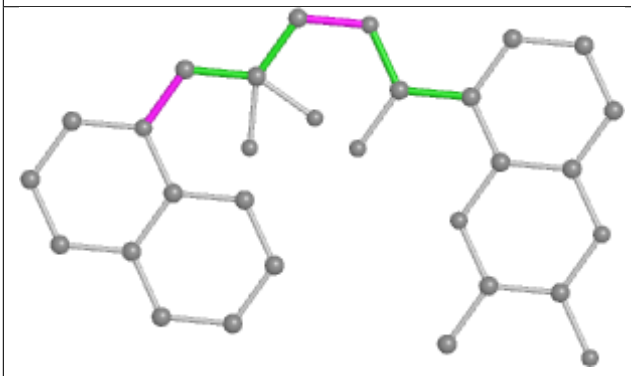
Ligand OSQ C 1202



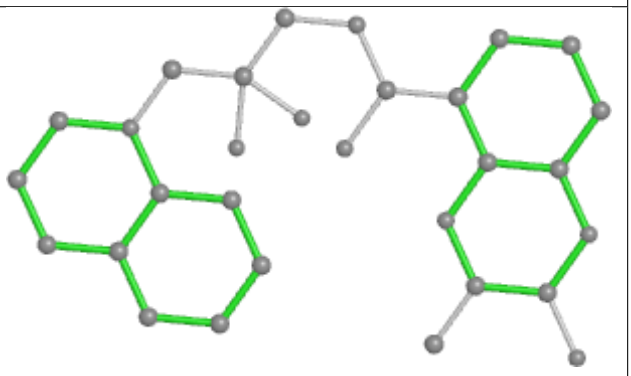
Bond lengths



Bond angles

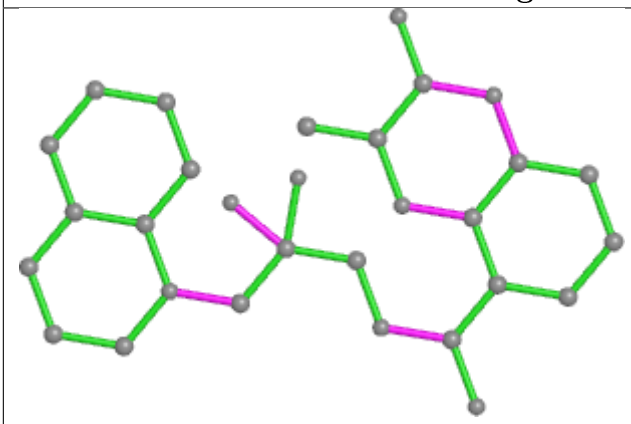


Torsions

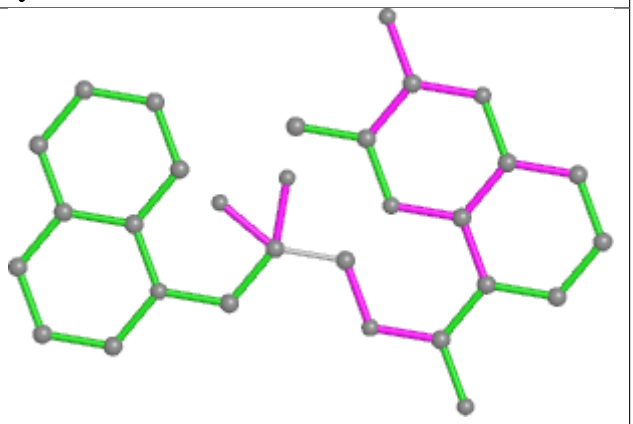


Rings

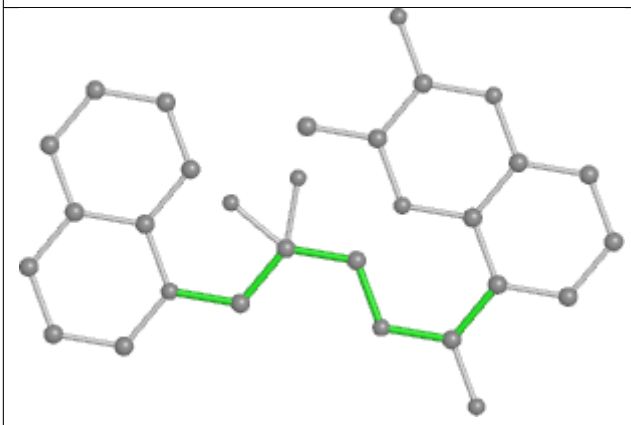
Ligand OSQ C 1201



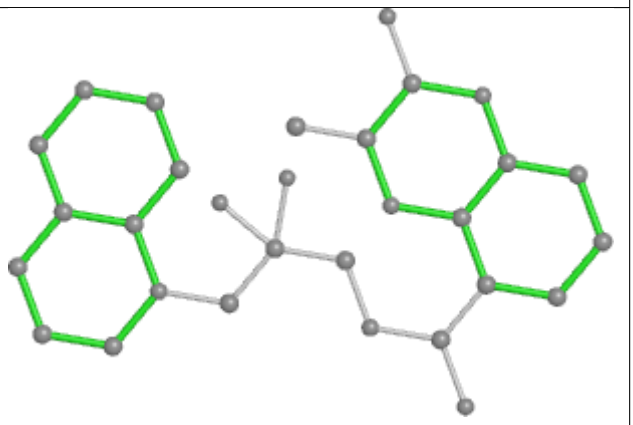
Bond lengths



Bond angles

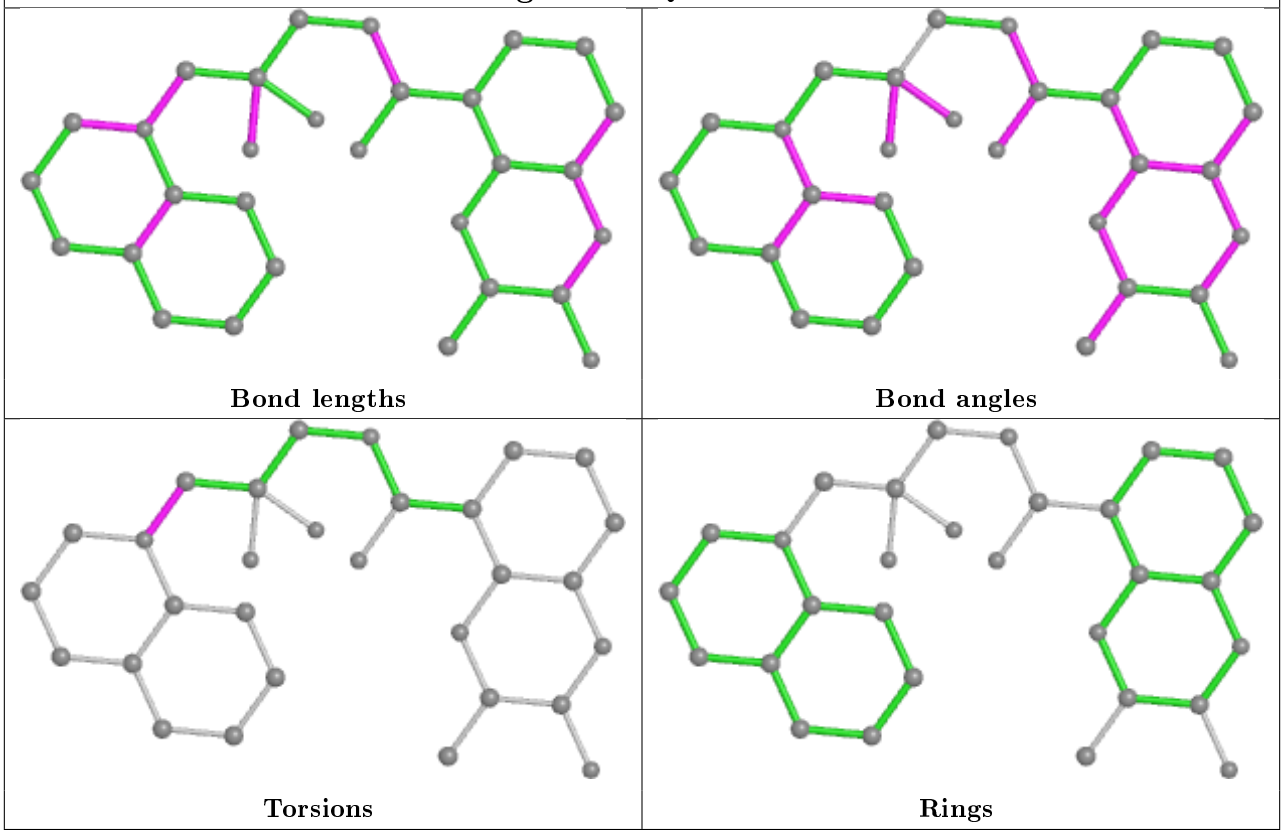


Torsions



Rings

Ligand OSQ A 1201



4.7 Other polymers (i)

There are no such residues in this entry.

4.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

5 Fit of model and data

5.1 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	A	114/119 (95%)	-0.51	0 100 100	18, 25, 52, 75	0
1	B	114/119 (95%)	-0.40	3 (2%) 56 60	21, 28, 58, 84	0
1	C	113/119 (94%)	-0.26	6 (5%) 26 29	25, 36, 72, 94	0
All	All	341/357 (95%)	-0.39	9 (2%) 56 60	18, 30, 60, 94	0

All (9) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	1196	LEU	3.4
1	C	1085	PHE	3.4
1	C	1145	GLY	3.3
1	B	1194	GLN	2.8
1	B	1151	TRP	2.4
1	C	1090	LEU	2.2
1	C	1087	PRO	2.2
1	C	1084	ILE	2.1
1	B	1190	ASP	2.0

5.2 Non-standard residues in protein, DNA, RNA chains

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

5.3 Carbohydrates

There are no carbohydrates in this entry.

5.4 Ligands

LIGAND-RSR INFOmissingINFO

5.5 Other polymers (i)

There are no such residues in this entry.