wwPDB X-ray Structure Validation Summary Report (i)

Feb 5, 2024 – 08:22 PM EST

PDB ID	:	1ZAF
Title	:	Crystal structure of estrogen receptor beta complexed with 3-Bromo-6-hydroxy-2-(4-hydroxy-phenyl)-inden-1-one
Authors	:	McDevitt, R.E.; Malamas, M.S.; Manas, E.S.; Unwalla, R.J.; Xu, Z.B.; Miller, C.P.; Harris, H.A.
Deposited on	:	2005-04-06
Resolution	:	2.20 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 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:

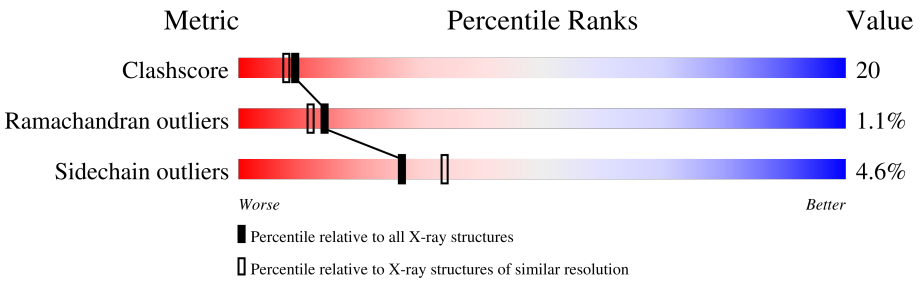
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.20 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))
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (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 >=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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	238	62% 31% · ·
1	B	238	66% 27% · ·
2	C	13	38% 31% 23% 8%
2	D	13	54% 23% 23%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 4001 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 Estrogen receptor beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	228	1801	1158	304	319	20	0	0	0
1	B	228	1801	1158	304	319	20	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	409	ALA	SER	conflict	UNP Q92731
B	409	ALA	SER	conflict	UNP Q92731

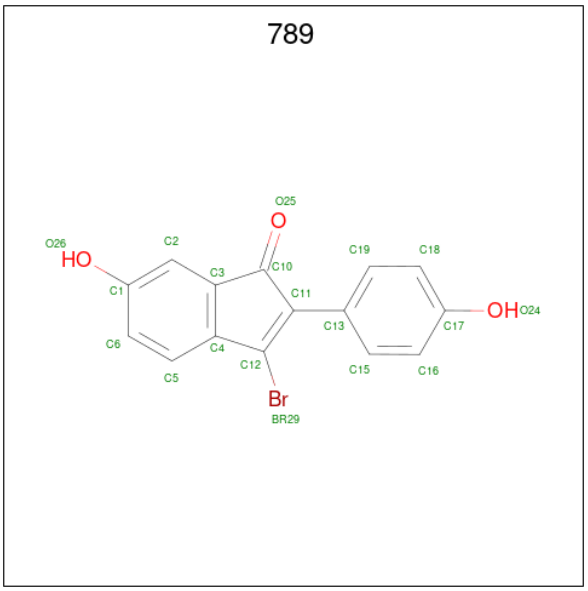
- Molecule 2 is a protein called Nuclear receptor coactivator 1.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	C	13	97	60	17	20	0	0	0
2	D	13	97	60	17	20	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	601	SER	-	cloning artifact	UNP Q15788
C	602	GLY	-	cloning artifact	UNP Q15788
D	601	SER	-	cloning artifact	UNP Q15788
D	602	GLY	-	cloning artifact	UNP Q15788

- Molecule 3 is 3-BROMO-6-HYDROXY-2-(4-HYDROXYPHENYL)-1H-INDEN-1-ONE (three-letter code: 789) (formula: $C_{15}H_9BrO_3$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	Br	C	O		
3	A	1	19	1	15	3	0	0
			Total	Br	C	O		
3	B	1	19	1	15	3	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	80	80	80	0	0
			Total	O		
4	B	81	81	81	0	0
			Total	O		
4	C	4	4	4	0	0
			Total	O		
4	D	2	2	2	0	0

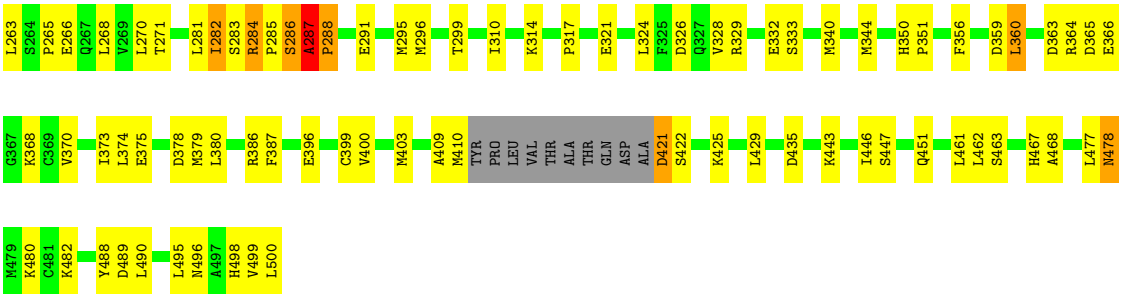
3 Residue-property plots (i)

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. 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. 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.

Note EDS was not executed.

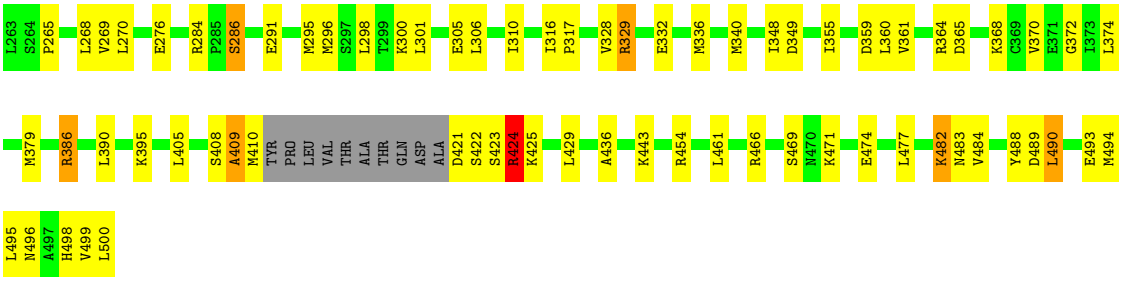
• Molecule 1: Estrogen receptor beta

Chain A: 62% 31%



• Molecule 1: Estrogen receptor beta

Chain B: 66% 27%



• Molecule 2: Nuclear receptor coactivator 1

Chain C: 38% 31% 23% 8%

• Molecule 2: Nuclear receptor coactivator 1

Chain D: 54% 23% 23%

S601
G602
S603
H604
K605
Q608
T612
T613

4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	52.28Å 89.21Å 99.33Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	14.91 – 2.20	Depositor
% Data completeness (in resolution range)	96.4 (14.91-2.20)	Depositor
R_{merge}	0.09	Depositor
R_{sym}	0.09	Depositor
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.226 , 0.286	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	4001	wwPDB-VP
Average B, all atoms (Å^2)	26.0	wwPDB-VP

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: 789

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	0.47	3/1833 (0.2%)	0.66	4/2473 (0.2%)
1	B	0.38	0/1833	1.08	6/2473 (0.2%)
2	C	1.46	4/97 (4.1%)	2.38	7/129 (5.4%)
2	D	0.55	0/97	1.86	2/129 (1.6%)
All	All	0.49	7/3860 (0.2%)	0.99	19/5204 (0.4%)

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	B	0	2
2	C	0	1
All	All	0	3

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	602	GLY	CA-C	8.03	1.64	1.51
2	C	602	GLY	N-CA	7.46	1.57	1.46
2	C	603	SER	N-CA	-6.59	1.33	1.46
1	A	422	SER	CA-CB	6.55	1.62	1.52
1	A	351	PRO	N-CD	-6.24	1.39	1.47

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	424	ARG	NE-CZ-NH1	-30.08	105.26	120.30
1	B	424	ARG	NE-CZ-NH2	23.43	132.01	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	424	ARG	CD-NE-CZ	-17.41	99.23	123.60
2	D	603	SER	C-N-CA	-17.34	78.34	121.70
2	C	604	HIS	N-CA-CB	-13.78	85.80	110.60

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	424	ARG	Sidechain
1	B	488	TYR	Sidechain
2	C	603	SER	Mainchain

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	1801	0	1882	81	0
1	B	1801	0	1882	70	1
2	C	97	0	103	7	0
2	D	97	0	103	8	0
3	A	19	0	9	0	0
3	B	19	0	9	0	0
4	A	80	0	0	6	0
4	B	81	0	0	4	1
4	C	4	0	0	0	0
4	D	2	0	0	0	0
All	All	4001	0	3988	155	1

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 20.

The worst 5 of 155 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)
2:D:604:HIS:O	2:D:608:GLN:HG3	1.45	1.16

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:285:PRO:HD3	1:A:359:ASP:HB2	1.39	1.02
1:B:423:SER:HB3	1:B:424:ARG:NH1	1.73	1.01
1:A:287:ALA:O	1:A:288:PRO:O	1.79	1.00
1:B:423:SER:HB3	1:B:424:ARG:HH12	1.23	0.99

All (1) 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:B:286:SER:O	4:B:53:HOH:O[4_555]	1.53	0.67

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	224/238 (94%)	210 (94%)	10 (4%)	4 (2%)	8 5
1	B	224/238 (94%)	215 (96%)	8 (4%)	1 (0%)	34 37
2	C	11/13 (85%)	9 (82%)	2 (18%)	0	100 100
2	D	11/13 (85%)	9 (82%)	2 (18%)	0	100 100
All	All	470/502 (94%)	443 (94%)	22 (5%)	5 (1%)	14 12

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	286	SER
1	B	409	ALA
1	A	284	ARG
1	A	287	ALA
1	A	283	SER

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	205/213 (96%)	197 (96%)	8 (4%)	32	41
1	B	205/213 (96%)	196 (96%)	9 (4%)	28	35
2	C	12/12 (100%)	11 (92%)	1 (8%)	11	11
2	D	12/12 (100%)	10 (83%)	2 (17%)	2	1
All	All	434/450 (96%)	414 (95%)	20 (5%)	27	34

5 of 20 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	482	LYS
2	C	609	LEU
2	D	613	THR
2	D	604	HIS
1	A	478	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
2	C	604	HIS
2	C	608	GLN
1	B	308	HIS
1	B	407	ASN
1	B	470	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)

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	$\#\|Z\| > 2$
3	789	A	201	-	20,21,21	2.23	11 (55%)	27,31,31	1.46	5 (18%)
3	789	B	202	-	20,21,21	2.29	12 (60%)	27,31,31	1.43	5 (18%)

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	789	A	201	-	-	0/4/20/20	0/3/3/3
3	789	B	202	-	-	0/4/20/20	0/3/3/3

The worst 5 of 23 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	202	789	C2-C1	3.92	1.45	1.39
3	A	201	789	C3-C10	-3.49	1.41	1.48
3	B	202	789	C5-C4	3.46	1.45	1.39
3	A	201	789	C2-C1	3.37	1.44	1.39
3	B	202	789	C3-C10	-3.27	1.41	1.48

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	201	789	C3-C10-C11	4.35	109.86	106.21
3	B	202	789	C3-C10-C11	4.23	109.76	106.21

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	202	789	C13-C11-C10	2.79	126.87	122.69
3	A	201	789	C13-C11-C10	2.72	126.76	122.69
3	B	202	789	C4-C3-C10	-2.55	105.83	108.95

There are no chirality outliers.

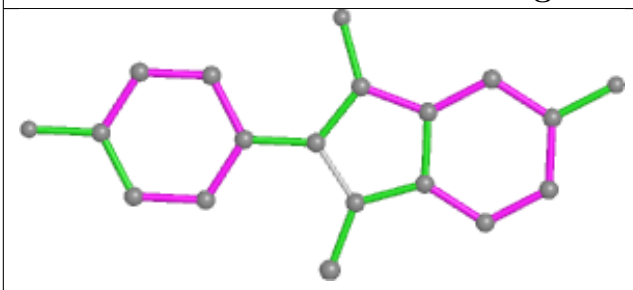
There are no torsion outliers.

There are no ring outliers.

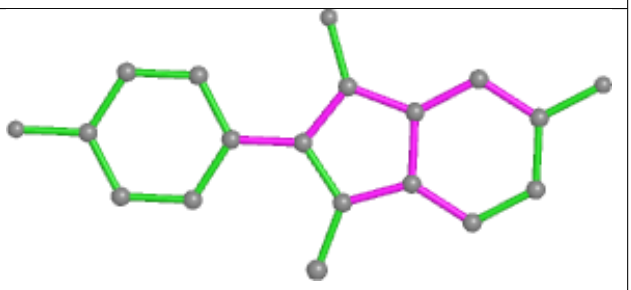
No monomer is involved in short contacts.

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.

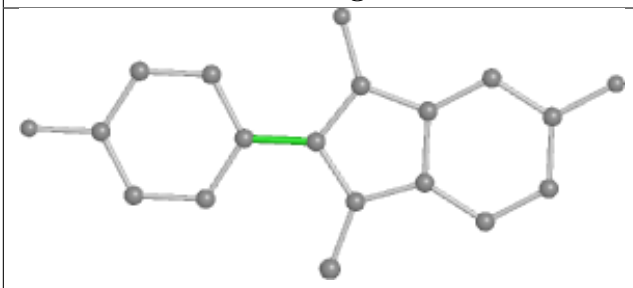
Ligand 789 A 201



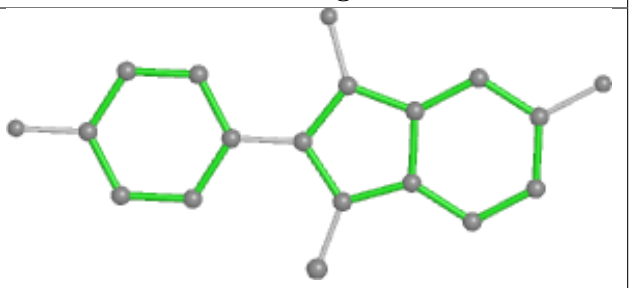
Bond lengths



Bond angles

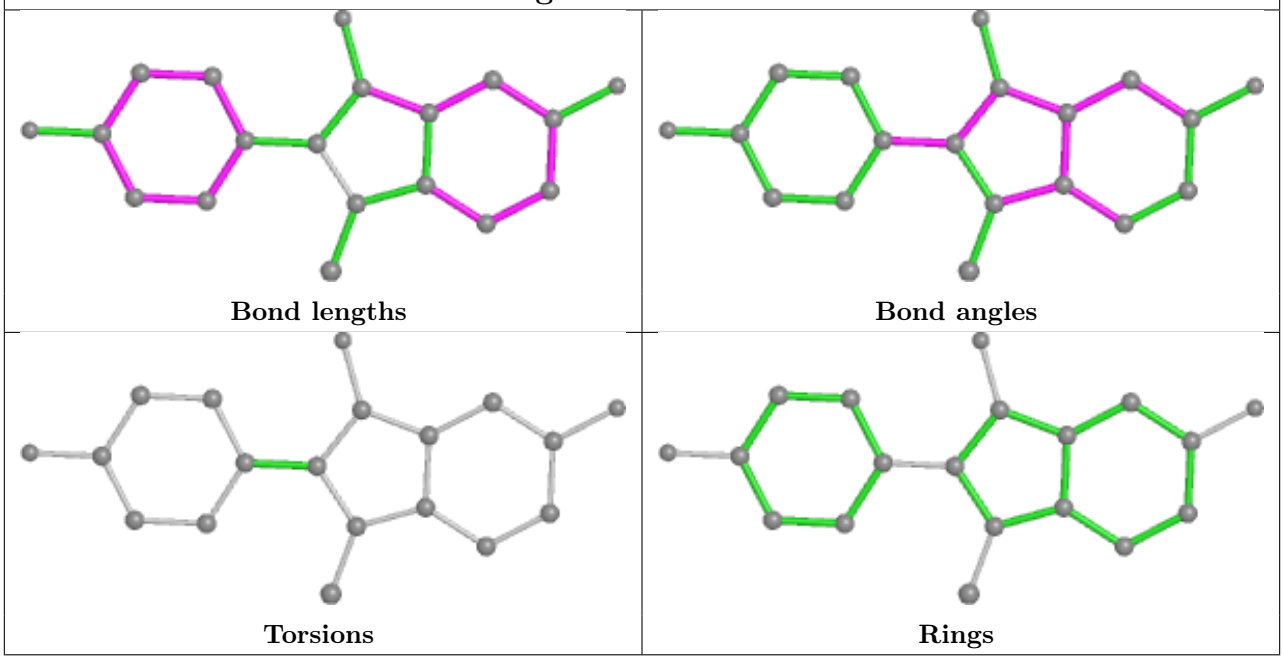


Torsions



Rings

Ligand 789 B 202



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)

EDS was not executed - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains (i)

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates (i)

EDS was not executed - this section is therefore empty.

6.4 Ligands (i)

EDS was not executed - this section is therefore empty.

6.5 Other polymers (i)

EDS was not executed - this section is therefore empty.