WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Jun 14, 2020 – 03:03 am BST

PDB ID	:	1I2D
Title	:	CRYSTAL STRUCTURE OF ATP SULFURYLASE FROM PENICILLIUM CHRYSOGENUM
Authors	:	MacRae, I.J.; Segel, I.H.; Fisher, A.J.
Deposited on	:	2001-02-07
Resolution	:	2.81 Å(reported)

The following versions of software and data (see references (1)) were used in the production of this report:

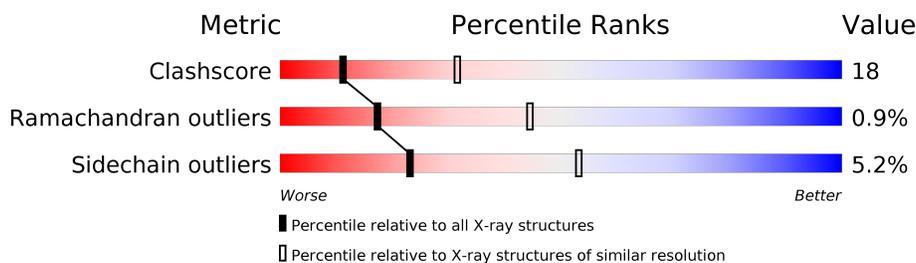|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 2.81 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	Percentile Ranks	Value
Clashscore		18
Ramachandran outliers		0.9%
Sidechain outliers		5.2%

Worse   *Better*

■ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	4060 (2.84-2.80)
Ramachandran outliers	138981	3978 (2.84-2.80)
Sidechain outliers	138945	3980 (2.84-2.80)

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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	573	70% 27% •
1	B	573	69% 27% •
1	C	573	67% 30% •

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 13851 atoms, of which 3 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.
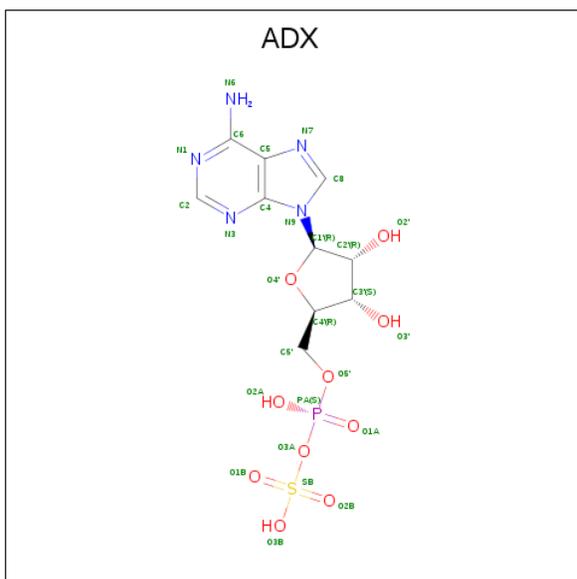
- Molecule 1 is a protein called ATP SULFURYLASE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	572	Total	C	H	N	O	S	7	0	0
			4469	2812	1	813	831	12			
1	B	572	Total	C	H	N	O	S	0	0	0
			4469	2812	1	813	831	12			
1	C	572	Total	C	H	N	O	S	7	0	0
			4469	2812	1	813	831	12			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	100	GLY	ALA	SEE REMARK 999	UNP Q12650
B	100	GLY	ALA	SEE REMARK 999	UNP Q12650
C	100	GLY	ALA	SEE REMARK 999	UNP Q12650

- Molecule 2 is ADENOSINE-5'-PHOSPHOSULFATE (three-letter code: ADX) (formula: $C_{10}H_{14}N_5O_{10}PS$).

ADX

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
2	A	1	Total 27	C 10	N 5	O 10	P 1	S 1	0	0
2	A	1	Total 27	C 10	N 5	O 10	P 1	S 1	0	0
2	B	1	Total 27	C 10	N 5	O 10	P 1	S 1	0	0
2	B	1	Total 27	C 10	N 5	O 10	P 1	S 1	0	0
2	C	1	Total 27	C 10	N 5	O 10	P 1	S 1	0	0
2	C	1	Total 27	C 10	N 5	O 10	P 1	S 1	0	0

- Molecule 3 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	178	Total 178	O 178	0	0
3	B	87	Total 87	O 87	0	0
3	C	17	Total 17	O 17	0	0

3 Residue-property plots ⓘ

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. 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.
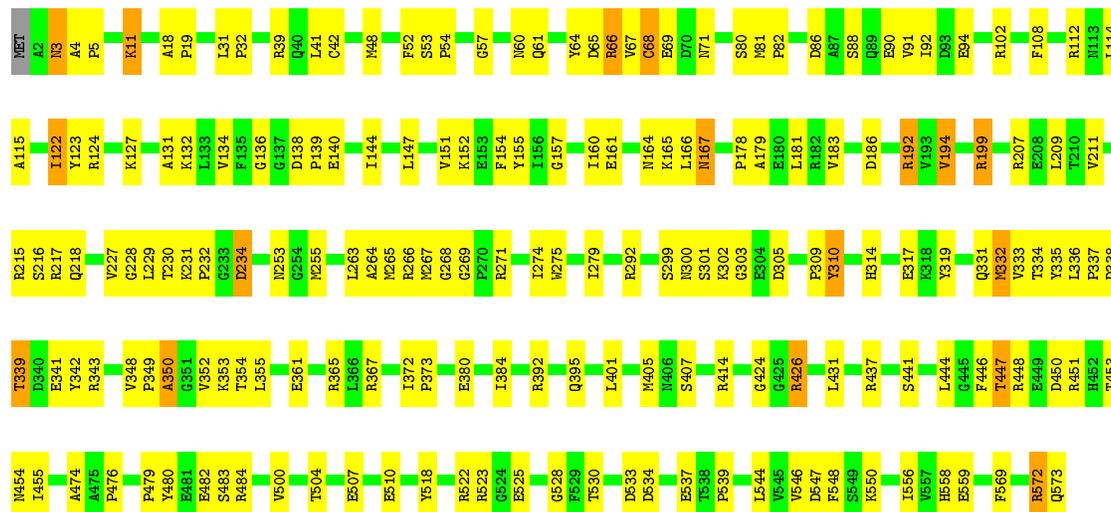
• Molecule 1: ATP SULFURYLASE
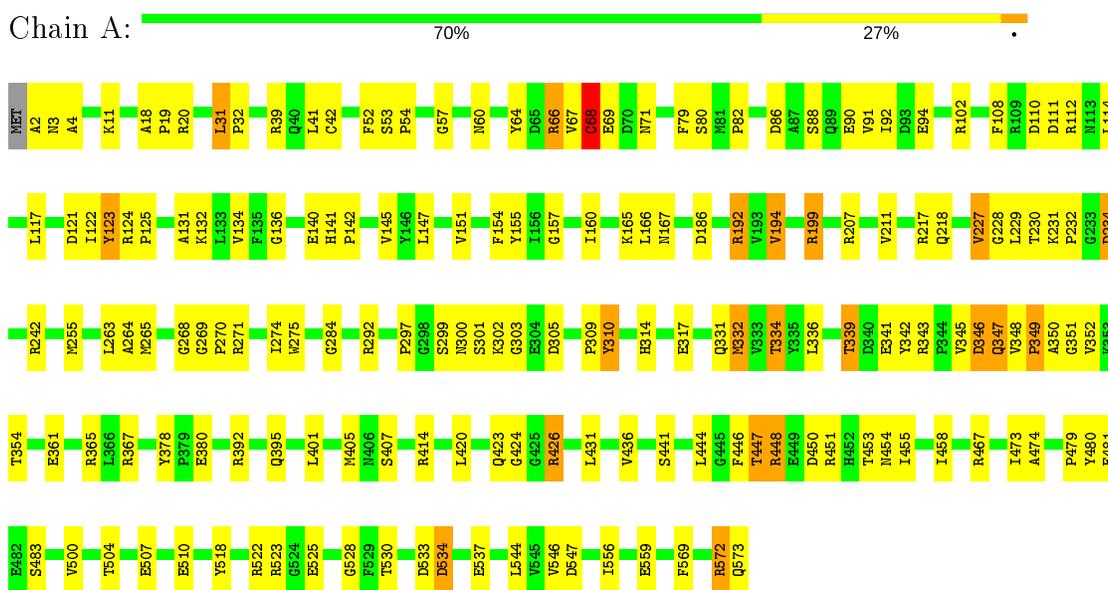
Chain A:

• Molecule 1: ATP SULFURYLASE
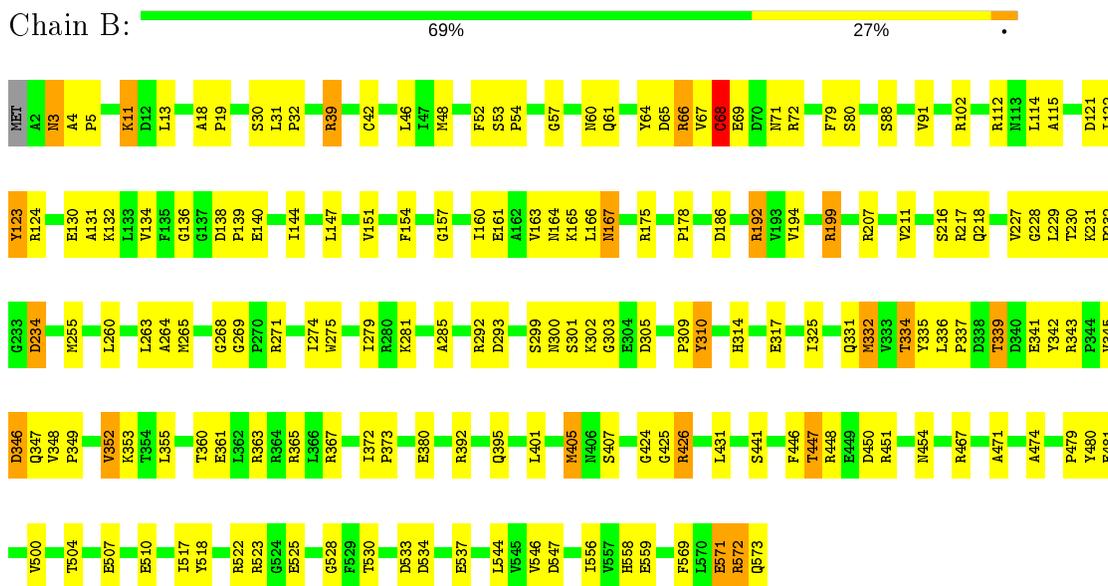
Chain B:

• Molecule 1: ATP SULFURYLASE

Chain C:

4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α, β, γ	135.67Å 162.09Å 273.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.08 – 2.81	Depositor
% Data completeness (in resolution range)	97.5 (29.08-2.81)	Depositor
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.207 , 0.245	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	13851	wwPDB-VP
Average B, all atoms (Å^2)	65.0	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: ADX

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	0.84	2/4563 (0.0%)	0.95	4/6191 (0.1%)
1	B	0.79	2/4563 (0.0%)	0.91	6/6191 (0.1%)
1	C	0.58	1/4563 (0.0%)	0.81	3/6191 (0.0%)
All	All	0.75	5/13689 (0.0%)	0.89	13/18573 (0.1%)

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	68	CYS	CB-SG	13.60	2.05	1.82
1	B	68	CYS	CB-SG	11.76	2.02	1.82
1	C	68	CYS	CB-SG	6.77	1.93	1.82
1	B	481	GLU	CG-CD	5.90	1.60	1.51
1	A	481	GLU	CG-CD	5.35	1.59	1.51

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	68	CYS	CA-CB-SG	7.96	128.33	114.00
1	B	68	CYS	CA-CB-SG	7.49	127.48	114.00
1	B	424	GLY	N-CA-C	6.69	129.82	113.10
1	A	424	GLY	N-CA-C	6.63	129.69	113.10
1	C	424	GLY	N-CA-C	6.45	129.23	113.10

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	4468	1	4394	154	0
1	B	4468	1	4394	161	0
1	C	4468	1	4394	164	0
2	A	54	0	25	3	0
2	B	54	0	25	4	0
2	C	54	0	26	5	0
3	A	178	0	0	7	0
3	B	87	0	0	9	0
3	C	17	0	0	3	0
All	All	13848	3	13258	480	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 18.

The worst 5 of 480 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:68:CYS:CB	1:B:68:CYS:SG	2.02	1.46
1:A:68:CYS:CB	1:A:68:CYS:SG	2.05	1.45
1:C:134:VAL:HA	1:C:271:ARG:HH11	1.22	1.02
1:B:134:VAL:HA	1:B:271:ARG:HH11	1.23	1.00
1:A:134:VAL:HA	1:A:271:ARG:HH11	1.23	1.00

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	A	570/573 (100%)	530 (93%)	33 (6%)	7 (1%)	13	37
1	B	570/573 (100%)	519 (91%)	46 (8%)	5 (1%)	17	44
1	C	570/573 (100%)	525 (92%)	41 (7%)	4 (1%)	22	51
All	All	1710/1719 (100%)	1574 (92%)	120 (7%)	16 (1%)	17	44

5 of 16 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	350	ALA
1	C	350	ALA
1	A	68	CYS
1	A	310	TYR
1	A	349	PRO

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	464/474 (98%)	440 (95%)	24 (5%)	23	53
1	B	464/474 (98%)	438 (94%)	26 (6%)	21	49
1	C	464/474 (98%)	441 (95%)	23 (5%)	24	55
All	All	1392/1422 (98%)	1319 (95%)	73 (5%)	23	53

5 of 73 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	186	ASP
1	B	293	ASP
1	C	426	ARG
1	B	216	SER
1	B	334	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	347	GLN
1	B	454	ASN
1	C	454	ASN
1	B	167	ASN
1	C	501	HIS

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 carbohydrates in this entry.

5.6 Ligand geometry (i)

6 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	ADX	A	574	-	25,29,29	1.83	6 (24%)	26,45,45	1.42	5 (19%)
2	ADX	C	579	-	25,29,29	2.14	6 (24%)	26,45,45	1.54	6 (23%)
2	ADX	B	576	-	25,29,29	1.58	4 (16%)	26,45,45	1.58	5 (19%)
2	ADX	A	575	-	25,29,29	1.19	1 (4%)	26,45,45	1.73	7 (26%)
2	ADX	C	578	-	25,29,29	2.04	6 (24%)	26,45,45	1.50	5 (19%)
2	ADX	B	577	-	25,29,29	1.41	2 (8%)	26,45,45	1.53	6 (23%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral

wwPDB X-ray Structure Validation Summary Report

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	ADX	A	574	-	-	2/6/32/32	0/3/3/3
2	ADX	C	579	-	-	6/6/32/32	0/3/3/3
2	ADX	B	576	-	-	5/6/32/32	0/3/3/3
2	ADX	A	575	-	-	2/6/32/32	0/3/3/3
2	ADX	C	578	-	-	5/6/32/32	0/3/3/3
2	ADX	B	577	-	-	0/6/32/32	0/3/3/3

The worst 5 of 25 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	574	ADX	O4'-C1'	5.85	1.49	1.41
2	C	579	ADX	PA-O3A	5.62	1.63	1.59
2	C	578	ADX	O4'-C1'	5.57	1.48	1.41
2	C	579	ADX	O4'-C1'	5.51	1.48	1.41
2	B	577	ADX	O4'-C1'	4.41	1.47	1.41

The worst 5 of 34 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	575	ADX	O3'-C3'-C2'	-3.26	101.28	111.82
2	A	575	ADX	C3'-C2'-C1'	3.13	105.69	100.98
2	C	578	ADX	C5-C6-N1	-3.08	113.36	120.35
2	A	575	ADX	O4'-C1'-C2'	-3.04	102.49	106.93
2	B	576	ADX	C5-C6-N6	3.01	124.92	120.35

There are no chirality outliers.

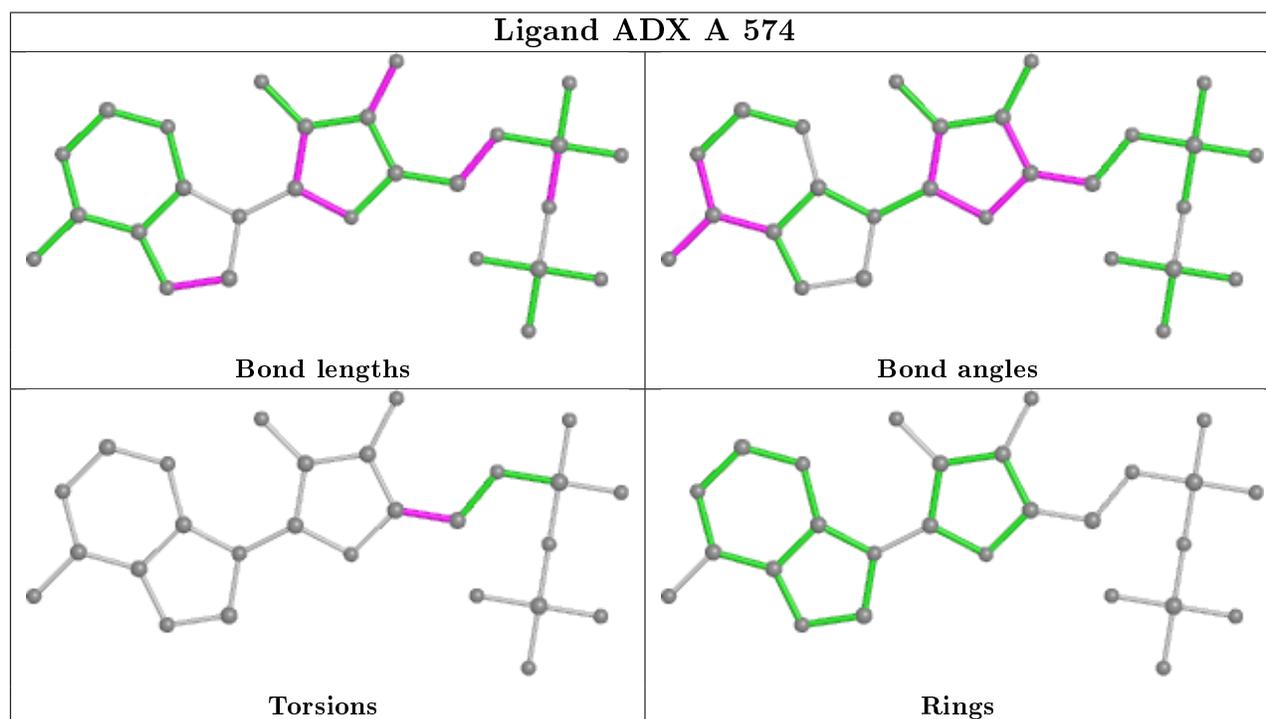5 of 20 torsion outliers are listed below:

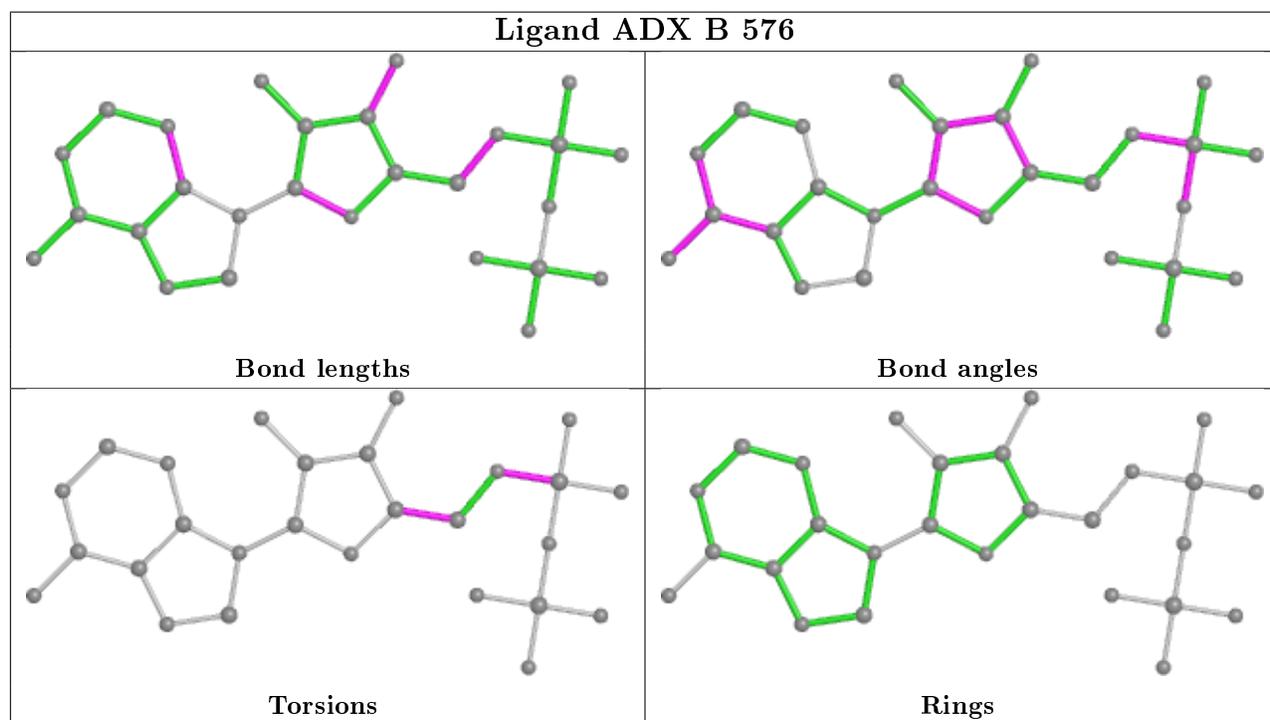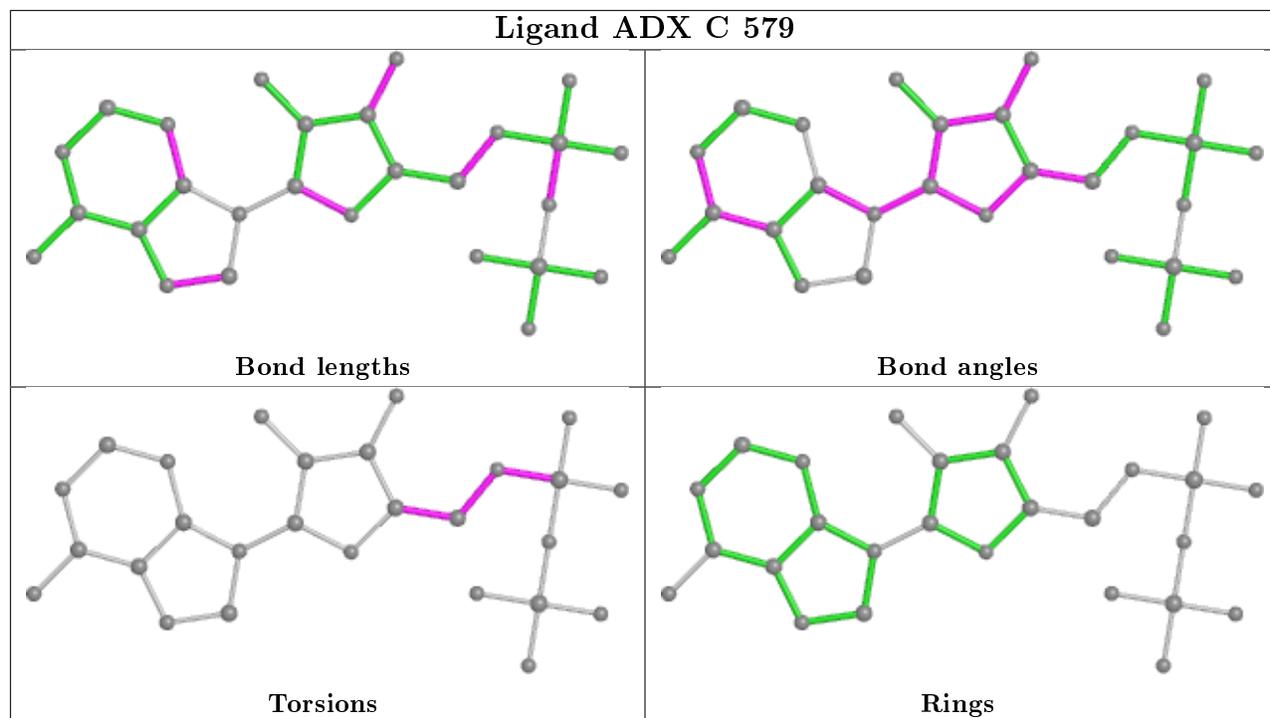| Mol	Chain	Res	Type	Atoms
2	C	579	ADX	C5'-O5'-PA-O2A
2	C	579	ADX	O4'-C4'-C5'-O5'
2	B	576	ADX	C5'-O5'-PA-O1A
2	B	576	ADX	C5'-O5'-PA-O3A
2	C	578	ADX	C5'-O5'-PA-O1A

There are no ring outliers.
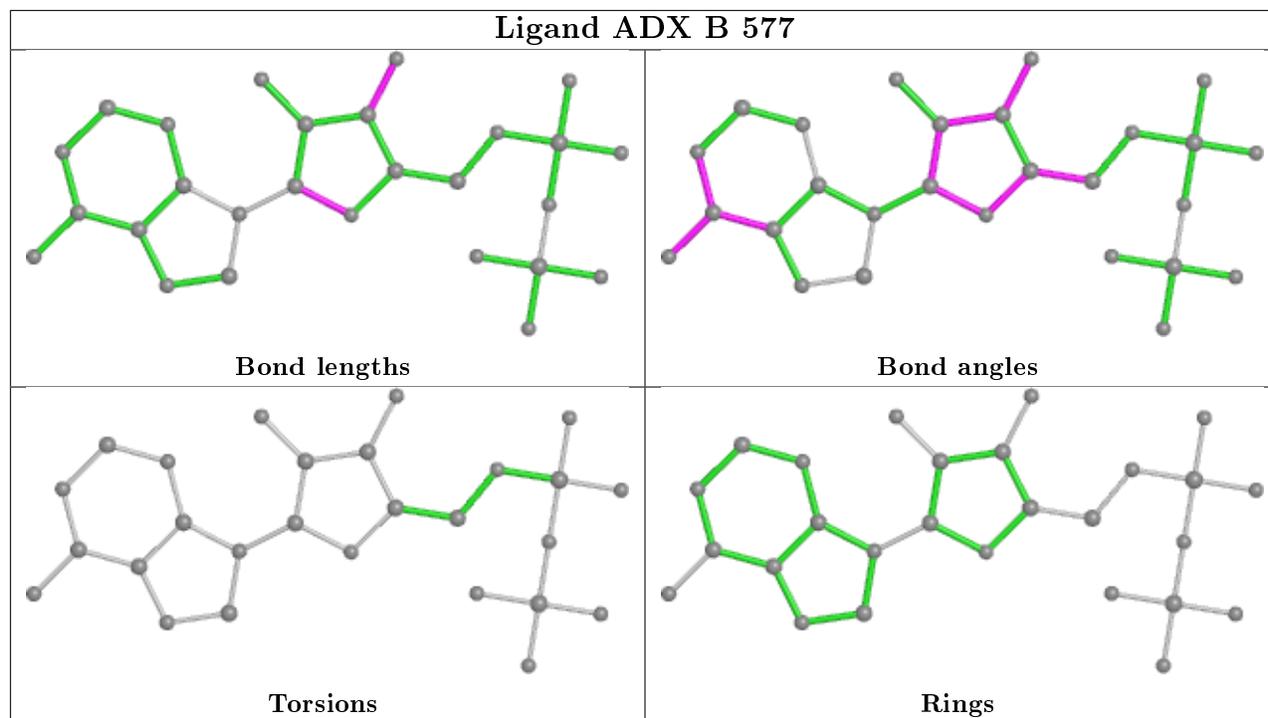
6 monomers are involved in 12 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	574	ADX	1	0
2	C	579	ADX	4	0
2	B	576	ADX	1	0
2	A	575	ADX	2	0
2	C	578	ADX	1	0
2	B	577	ADX	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 ADX A 574

Bond lengths

Bond angles

Torsions

Rings

Ligand ADX C 579

Bond lengths

Bond angles

Torsions

Rings

Ligand ADX B 576

Bond lengths

Bond angles

Torsions

Rings

Ligand ADX A 575

Bond lengths

Bond angles

Torsions

Rings

Ligand ADX C 578

Bond lengths

Bond angles

Torsions

Rings

Ligand ADX B 577

Bond lengths

Bond angles

Torsions

Rings
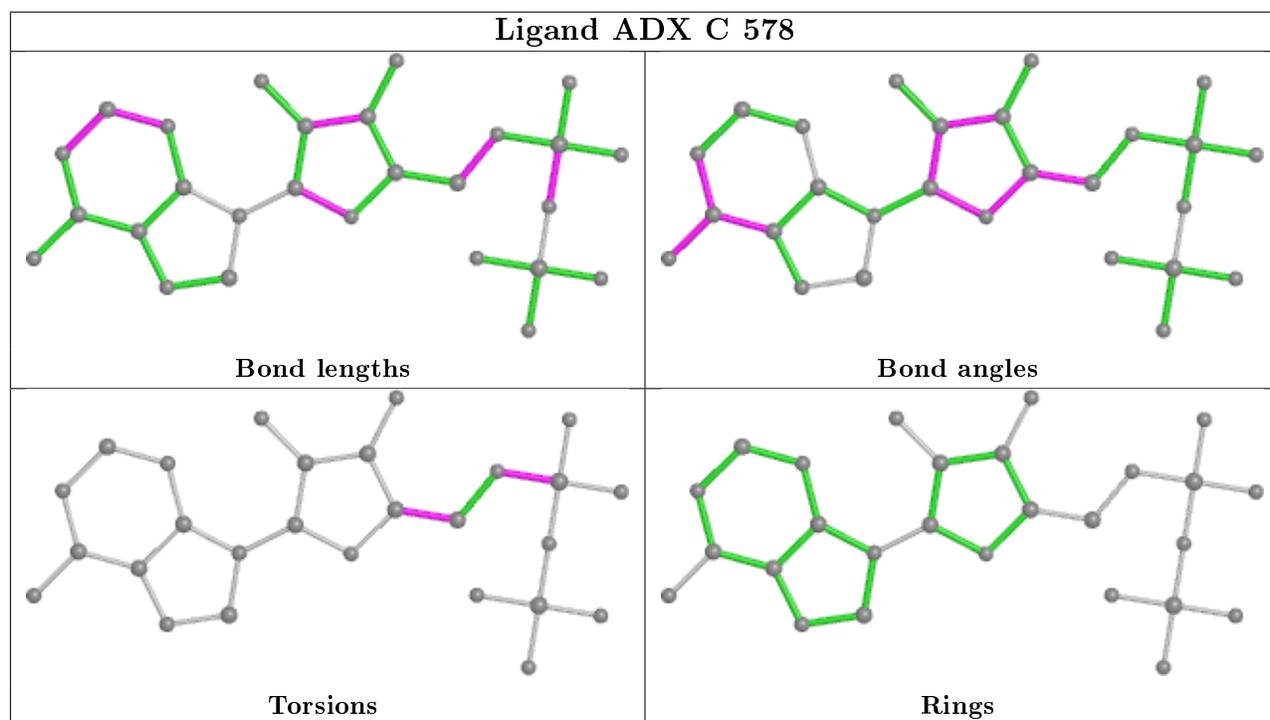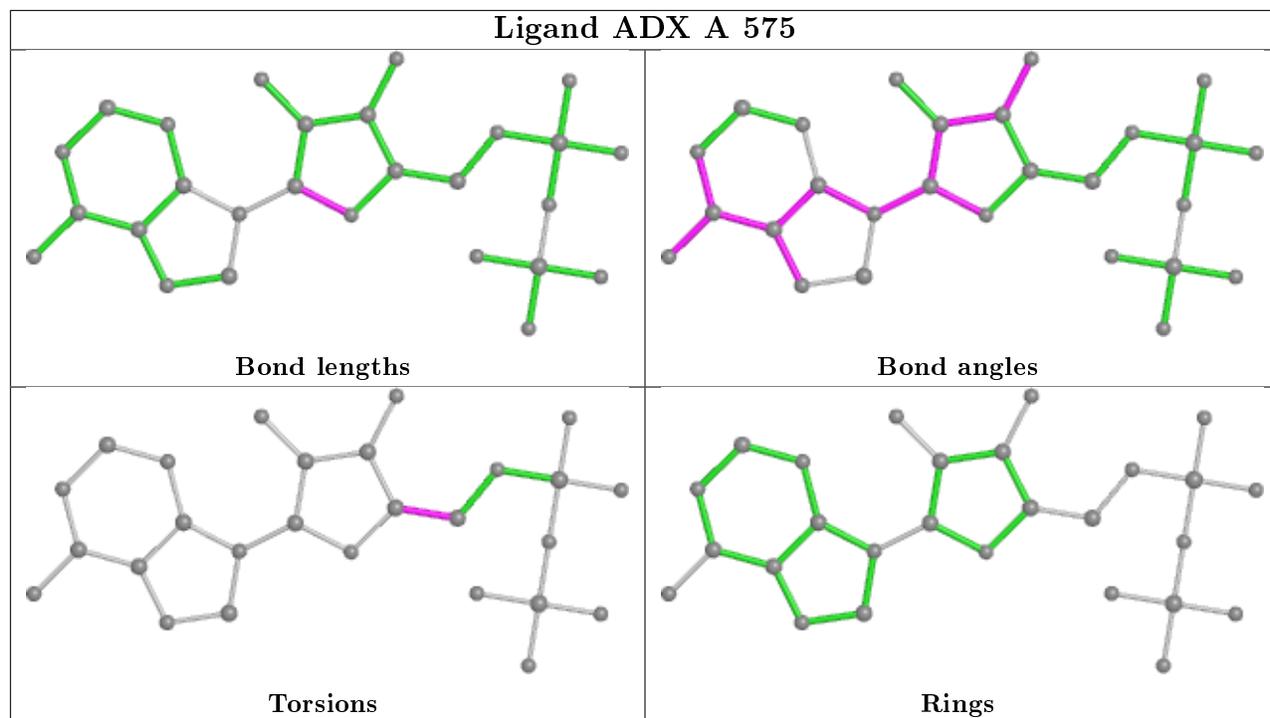
5.7 Other polymers (i)

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.