wwPDB X-ray Structure Validation Summary Report (i)

May 16, 2020 – 08:36 pm BST

PDB ID	:	1DO0
Title	:	ORTHORHOMBIC CRYSTAL FORM OF HEAT SHOCK LOCUS U (HSLU) FROM ESCHERICHIA COLI
Authors	:	Bochtler, M.; Hartmann, C.; Song, H.K.; Bourenkov, G.P.; Bartunik, H.D.
Deposited on	:	1999-12-18
Resolution	:	3.00 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:

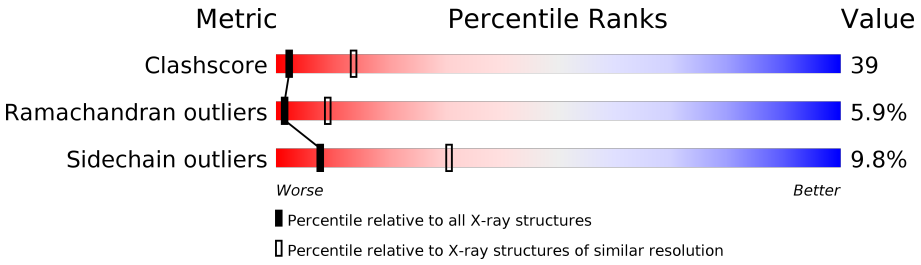
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, 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 3.00 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	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)

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	442	47% 38% 7% 8%
1	B	442	43% 40% 9% 8%
1	C	442	42% 42% 7% 8%
1	D	442	48% 38% 6% 8%
1	E	442	42% 41% 8% 8%
1	F	442	42% 42% 8% 8%

2 Entry composition

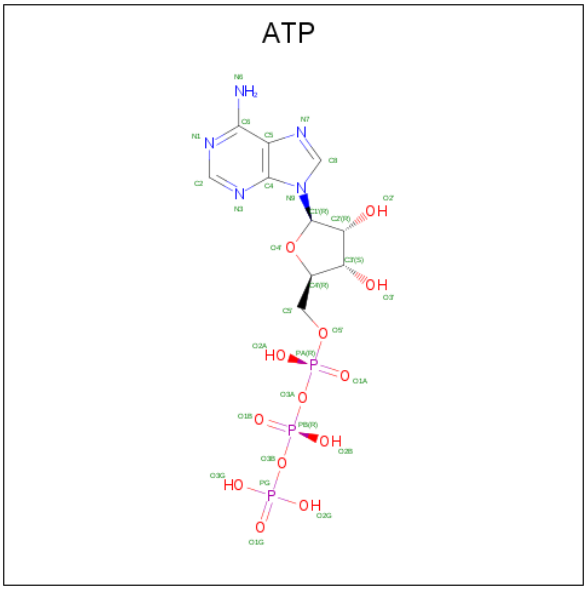
There are 4 unique types of molecules in this entry. The entry contains 19366 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 PROTEIN (HEAT SHOCK LOCUS U).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	406	Total	C	N	O	S	628	0	0
			3205	2001	570	624	10			
1	B	406	Total	C	N	O	S	143	0	0
			3205	2001	570	624	10			
1	C	406	Total	C	N	O	S	677	0	0
			3205	2001	570	624	10			
1	D	406	Total	C	N	O	S	579	0	0
			3205	2001	570	624	10			
1	E	406	Total	C	N	O	S	135	0	0
			3205	2001	570	624	10			
1	F	406	Total	C	N	O	S	600	0	0
			3205	2001	570	624	10			

- Molecule 2 is ADENOSINE-5'-TRIPHOSPHATE (three-letter code: ATP) (formula: $C_{10}H_{16}N_5O_{13}P_3$).

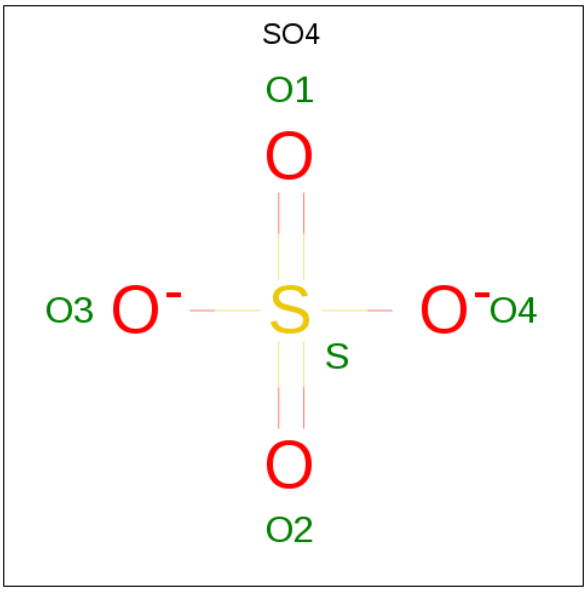


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
2	B	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
2	D	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
2	E	1	Total	C	N	O	P	0	0
			31	10	5	13	3		

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Mg	0	0
			1	1		
3	E	1	Total	Mg	0	0
			1	1		

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	C	1	Total	O	S	0	0
			5	4	1		
4	F	1	Total	O	S	0	0
			5	4	1		

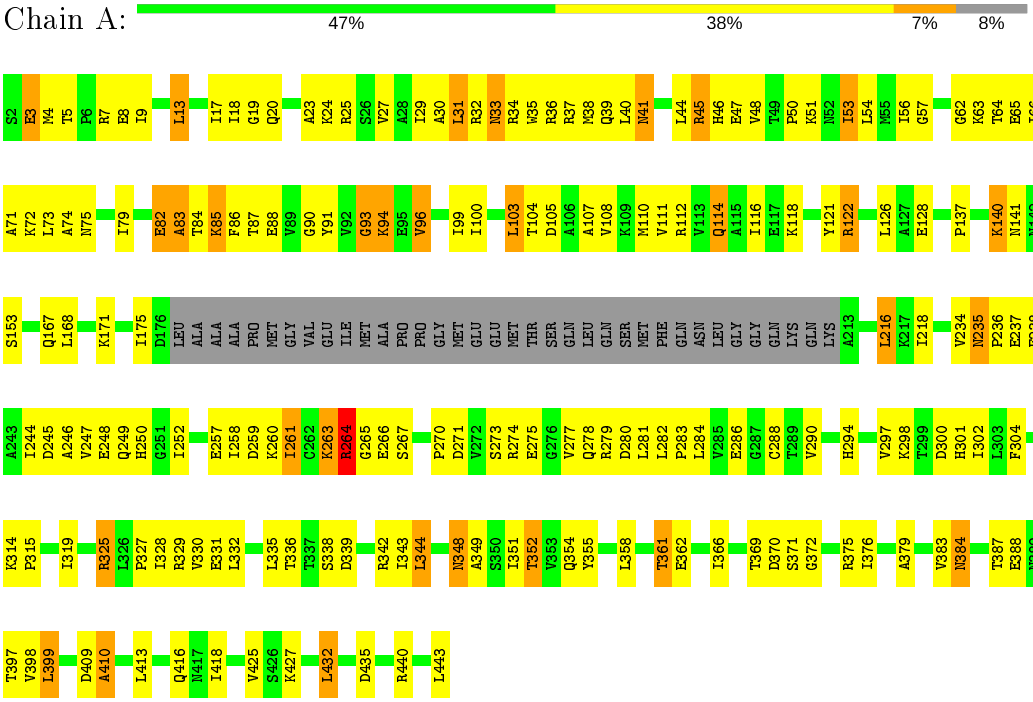
3 Residue-property plots (i)

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. 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.

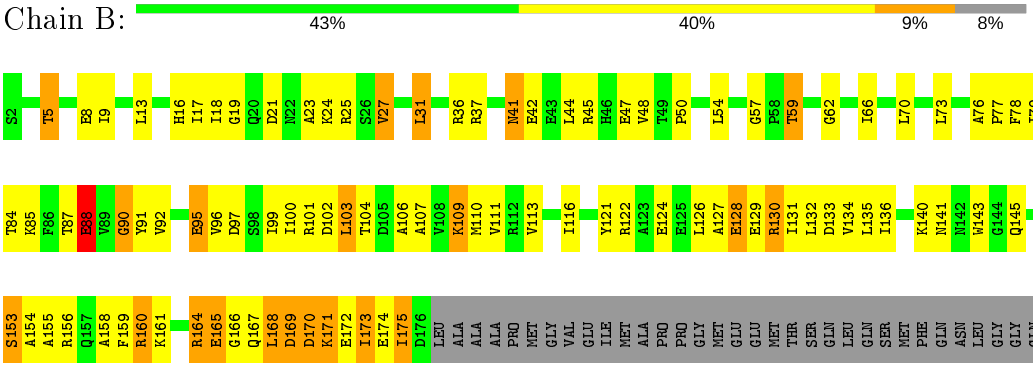
- Molecule 1: PROTEIN (HEAT SHOCK LOCUS U)

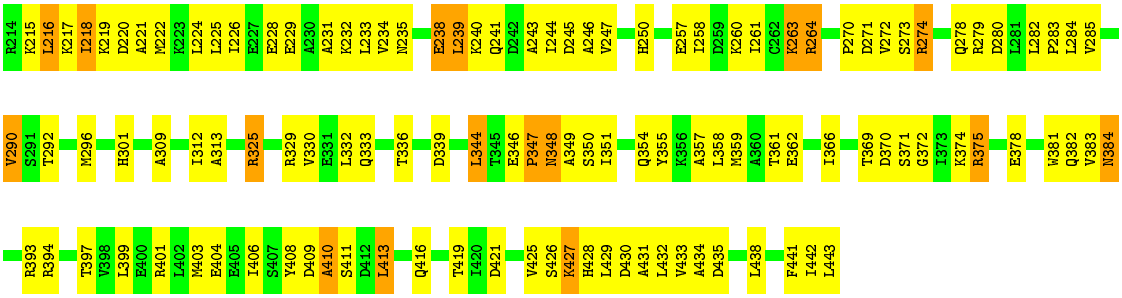
Chain A:



- Molecule 1: PROTEIN (HEAT SHOCK LOCUS U)

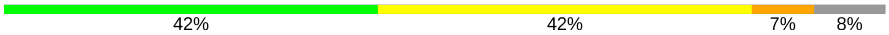
Chain B:

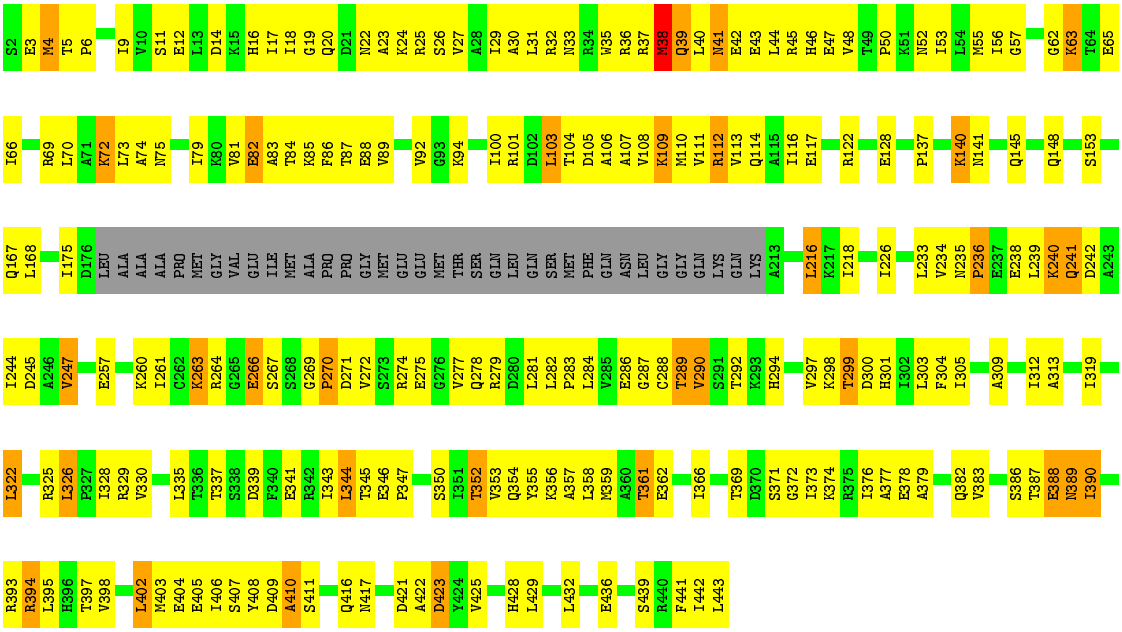




• Molecule 1: PROTEIN (HEAT SHOCK LOCUS U)

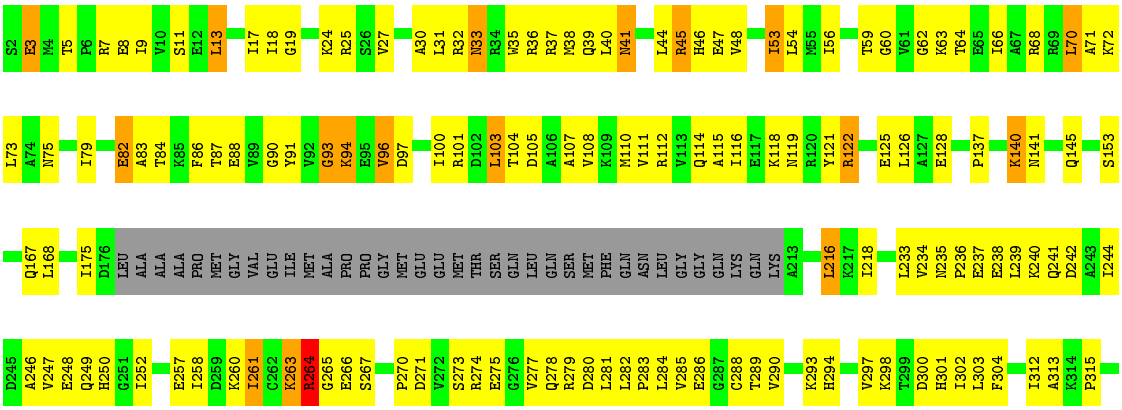
Chain C:

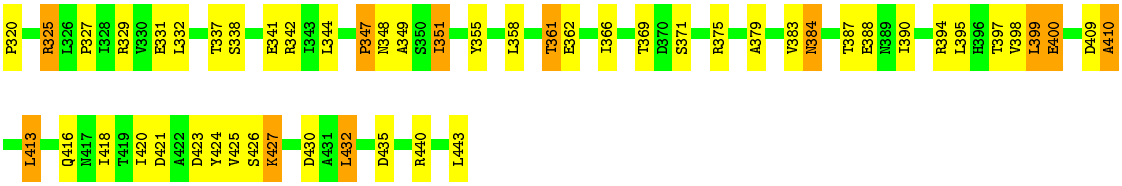




• Molecule 1: PROTEIN (HEAT SHOCK LOCUS U)

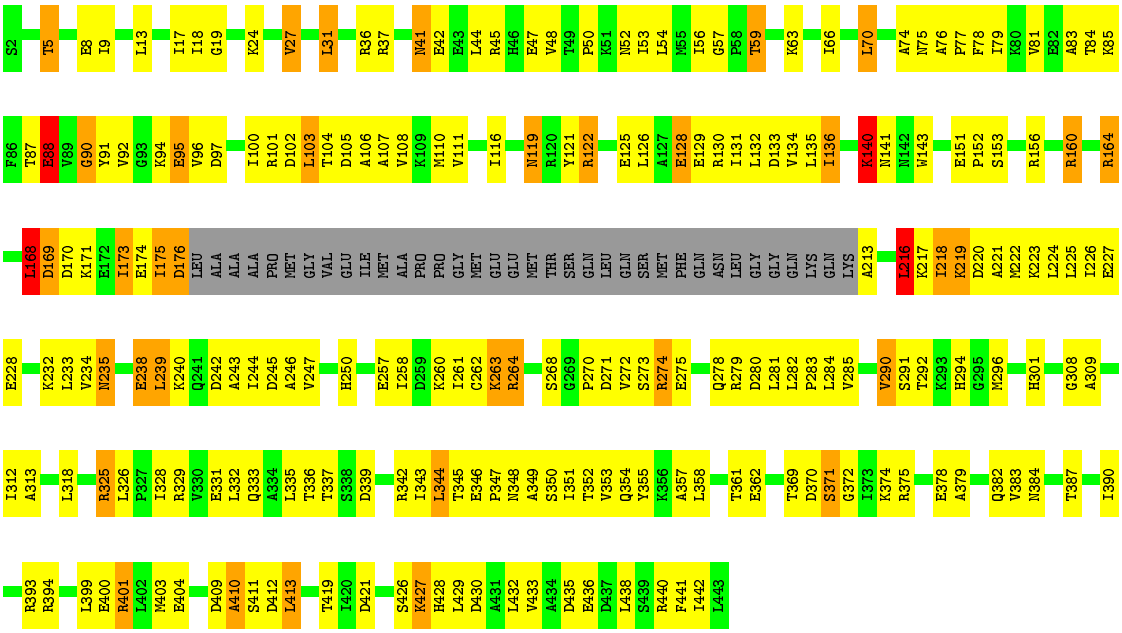
Chain D:





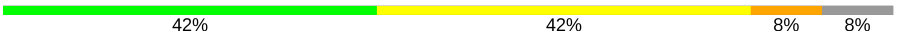
• Molecule 1: PROTEIN (HEAT SHOCK LOCUS U)

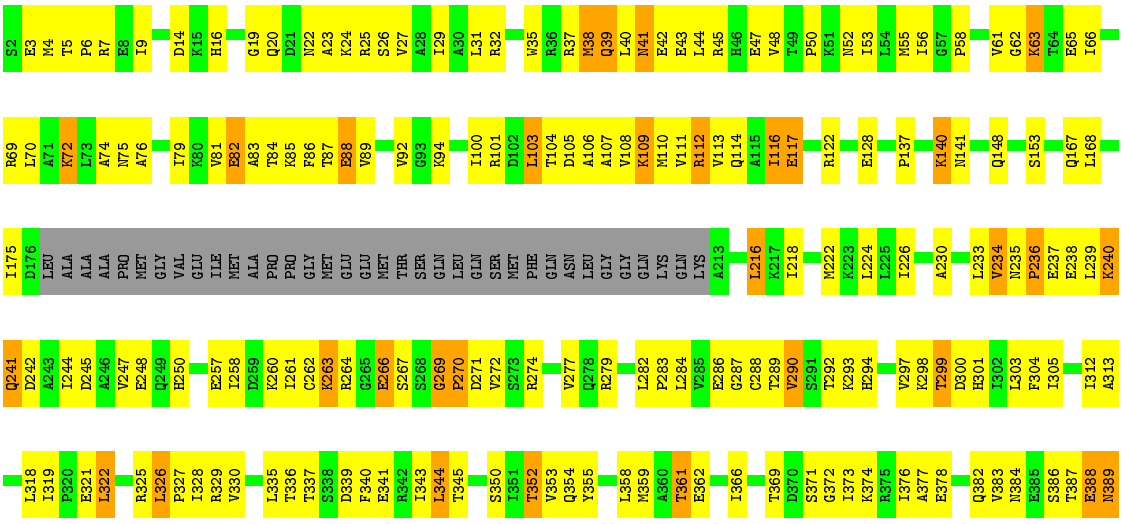
Chain E:



• Molecule 1: PROTEIN (HEAT SHOCK LOCUS U)

Chain F:






I390
G391
A392
R393
R394
L395
H396
T397
V398
L399
E400
R401
L402
M403
E404
E405
I406
S407
Y408
D409
A410
S411
Q416
N417
A422
V425
S426
K427
H428
L429
D430
A431
L432
V433
E436
S439
R440
F441
I442
L443

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 2	Depositor
Cell constants a, b, c, α, β, γ	208.15Å 167.70Å 108.19Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	25.00 – 3.00	Depositor
% Data completeness (in resolution range)	96.8 (25.00-3.00)	Depositor
R_{merge}	0.08	Depositor
R_{sym}	7.90	Depositor
Refinement program	CNS 0.9	Depositor
R, R_{free}	0.294 , 0.342	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	19366	wwPDB-VP
Average B, all atoms (Å^2)	55.0	wwPDB-VP

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: MG, SO4, ATP

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	0.42	0/3244	0.65	0/4373
1	B	0.48	0/3244	0.79	7/4373 (0.2%)
1	C	0.41	0/3244	0.64	0/4373
1	D	0.42	0/3244	0.64	0/4373
1	E	0.46	0/3244	0.73	4/4373 (0.1%)
1	F	0.42	0/3244	0.64	0/4373
All	All	0.43	0/19464	0.68	11/26238 (0.0%)

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	130	ARG	CG-CD-NE	14.79	142.86	111.80
1	B	109	LYS	CB-CG-CD	-9.55	86.78	111.60
1	B	165	GLU	CB-CA-C	-8.04	94.33	110.40
1	B	160	ARG	NE-CZ-NH2	6.98	123.79	120.30
1	E	160	ARG	NE-CZ-NH2	6.98	123.79	120.30

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	A	3205	0	3265	215	0
1	B	3205	0	3265	239	0
1	C	3205	0	3265	231	0
1	D	3205	0	3265	209	0
1	E	3205	0	3265	231	0
1	F	3205	0	3265	229	0
2	A	31	0	12	5	0
2	B	31	0	12	3	0
2	D	31	0	12	5	0
2	E	31	0	12	4	0
3	B	1	0	0	0	0
3	E	1	0	0	0	0
4	C	5	0	0	0	0
4	F	5	0	0	0	0
All	All	19366	0	19638	1287	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 39.

The worst 5 of 1287 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:E:168:LEU:O	1:E:169:ASP:CB	1.77	1.20
1:B:169:ASP:O	1:B:170:ASP:CG	1.84	1.16
1:C:312:ILE:HG13	1:C:313:ALA:H	1.20	1.07
1:E:168:LEU:O	1:E:169:ASP:HB2	1.28	1.06
1:B:169:ASP:O	1:B:170:ASP:OD1	1.74	1.06

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	402/442 (91%)	328 (82%)	50 (12%)	24 (6%)	1 9
1	B	402/442 (91%)	337 (84%)	46 (11%)	19 (5%)	2 14
1	C	402/442 (91%)	297 (74%)	72 (18%)	33 (8%)	1 4
1	D	402/442 (91%)	332 (83%)	49 (12%)	21 (5%)	2 12
1	E	402/442 (91%)	343 (85%)	42 (10%)	17 (4%)	3 16
1	F	402/442 (91%)	305 (76%)	68 (17%)	29 (7%)	1 5
All	All	2412/2652 (91%)	1942 (80%)	327 (14%)	143 (6%)	1 9

5 of 143 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	83	ALA
1	A	116	ILE
1	A	167	GLN
1	A	264	ARG
1	B	153	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	348/376 (93%)	316 (91%)	32 (9%)	9 34
1	B	348/376 (93%)	312 (90%)	36 (10%)	7 28
1	C	348/376 (93%)	316 (91%)	32 (9%)	9 34
1	D	348/376 (93%)	317 (91%)	31 (9%)	9 35
1	E	348/376 (93%)	303 (87%)	45 (13%)	4 19
1	F	348/376 (93%)	319 (92%)	29 (8%)	11 39
All	All	2088/2256 (93%)	1883 (90%)	205 (10%)	8 30

5 of 205 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	388	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	237	GLU
1	F	240	LYS
1	C	402	LEU
1	D	53	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 54 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	389	ASN
1	D	119	ASN
1	F	354	GLN
1	C	416	GLN
1	D	22	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 carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 8 ligands modelled in this entry, 2 are monoatomic - leaving 6 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	$\#\|Z\|>2$
4	SO4	C	800	-	4,4,4	1.89	2 (50%)	6,6,6	0.96	0
4	SO4	F	805	-	4,4,4	1.81	1 (25%)	6,6,6	0.99	0
2	ATP	A	900	-	26,33,33	0.83	0	31,52,52	0.90	1 (3%)
2	ATP	E	915	3	26,33,33	0.77	0	31,52,52	0.88	0
2	ATP	D	910	-	26,33,33	0.77	0	31,52,52	0.87	0
2	ATP	B	905	3	26,33,33	0.82	0	31,52,52	0.85	1 (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
2	ATP	D	910	-	-	2/18/38/38	0/3/3/3
2	ATP	E	915	3	-	2/18/38/38	0/3/3/3
2	ATP	A	900	-	-	2/18/38/38	0/3/3/3
2	ATP	B	905	3	-	2/18/38/38	0/3/3/3

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	C	800	SO4	O1-S	3.12	1.62	1.46
4	F	805	SO4	O1-S	3.00	1.62	1.46
4	C	800	SO4	O3-S	-2.06	1.31	1.47

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	900	ATP	C5-C6-N6	2.16	123.63	120.35
2	B	905	ATP	C5-C6-N6	2.06	123.48	120.35

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

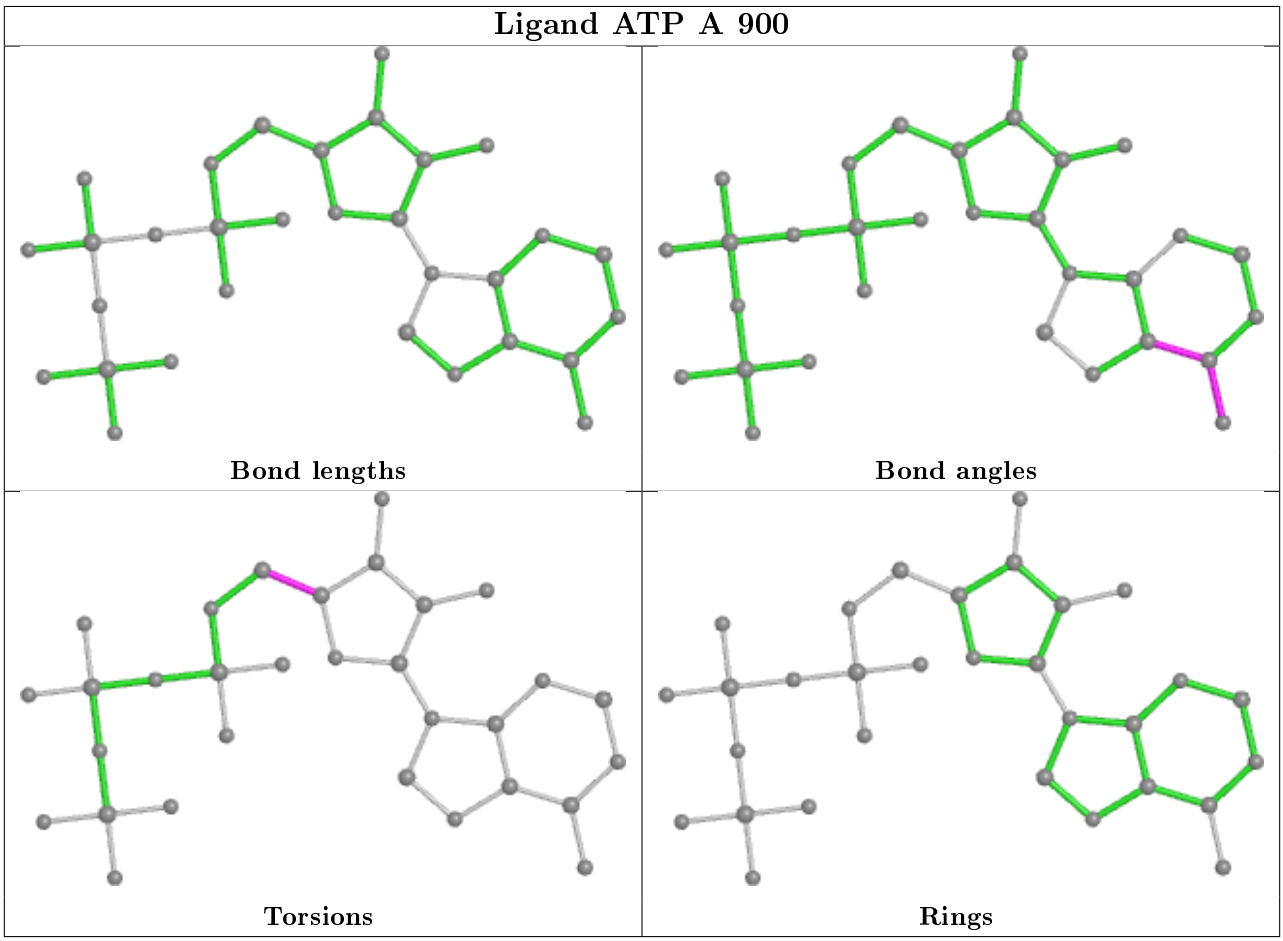
Mol	Chain	Res	Type	Atoms
2	A	900	ATP	O4'-C4'-C5'-O5'
2	E	915	ATP	O4'-C4'-C5'-O5'
2	D	910	ATP	O4'-C4'-C5'-O5'
2	B	905	ATP	O4'-C4'-C5'-O5'
2	A	900	ATP	C3'-C4'-C5'-O5'

There are no ring outliers.

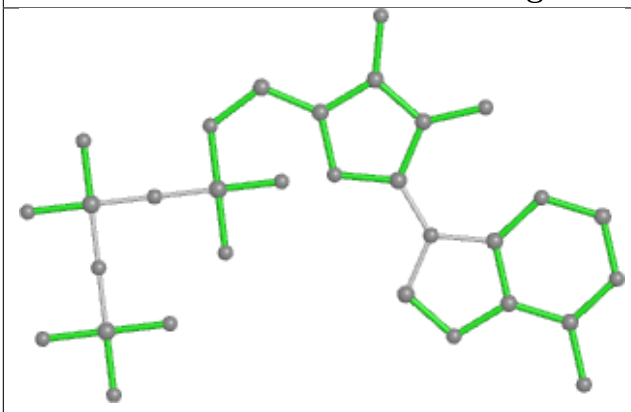
4 monomers are involved in 17 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	900	ATP	5	0
2	E	915	ATP	4	0
2	D	910	ATP	5	0
2	B	905	ATP	3	0

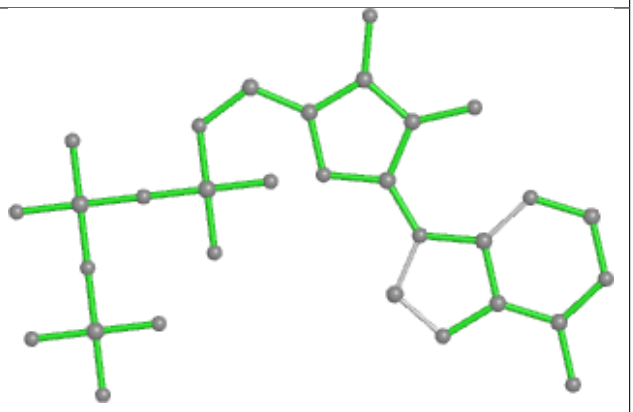
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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.



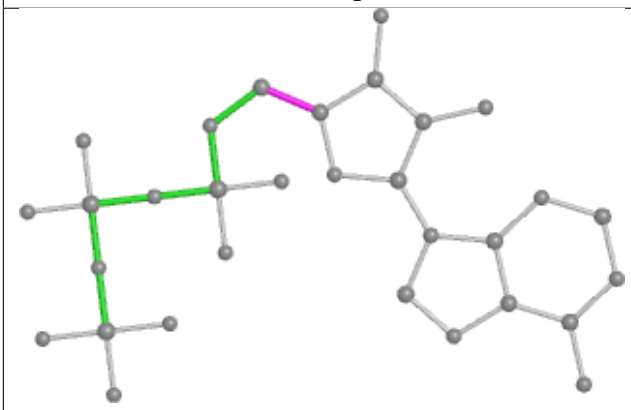
Ligand ATP E 915



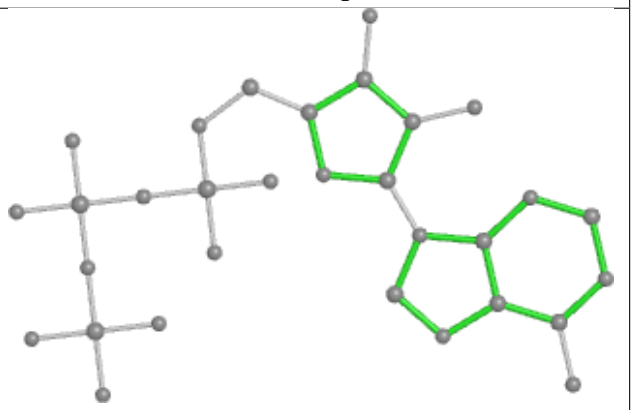
Bond lengths



Bond angles

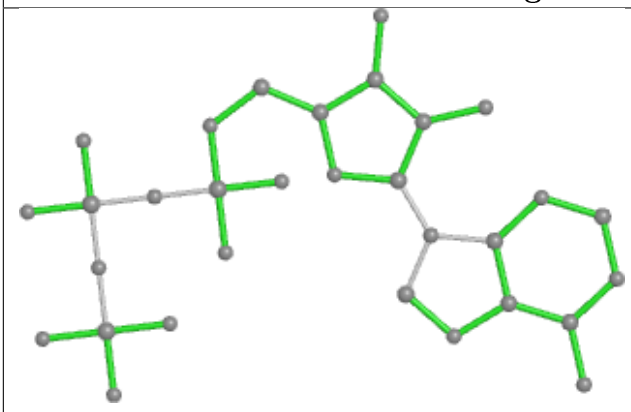


Torsions

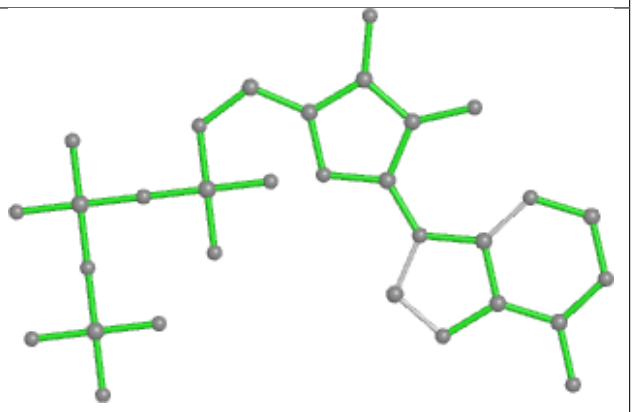


Rings

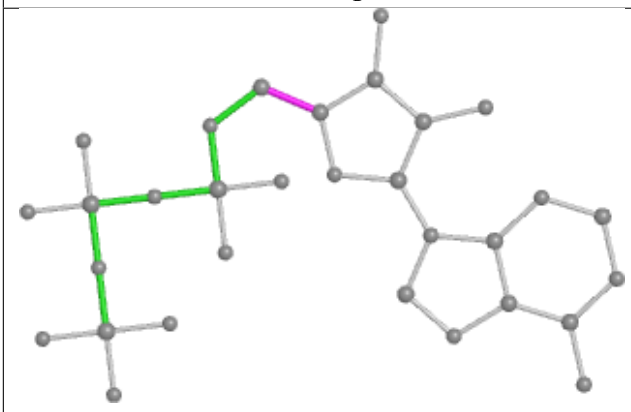
Ligand ATP D 910



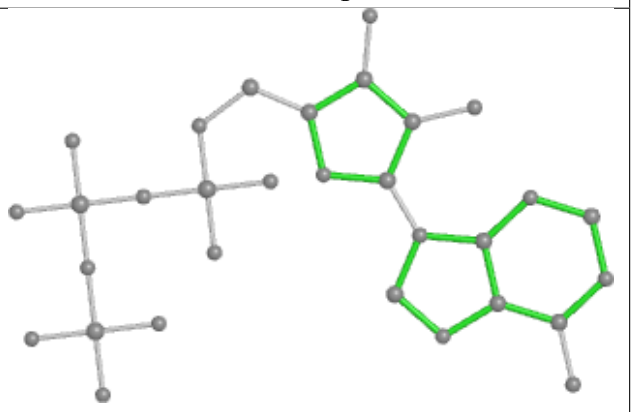
Bond lengths



Bond angles

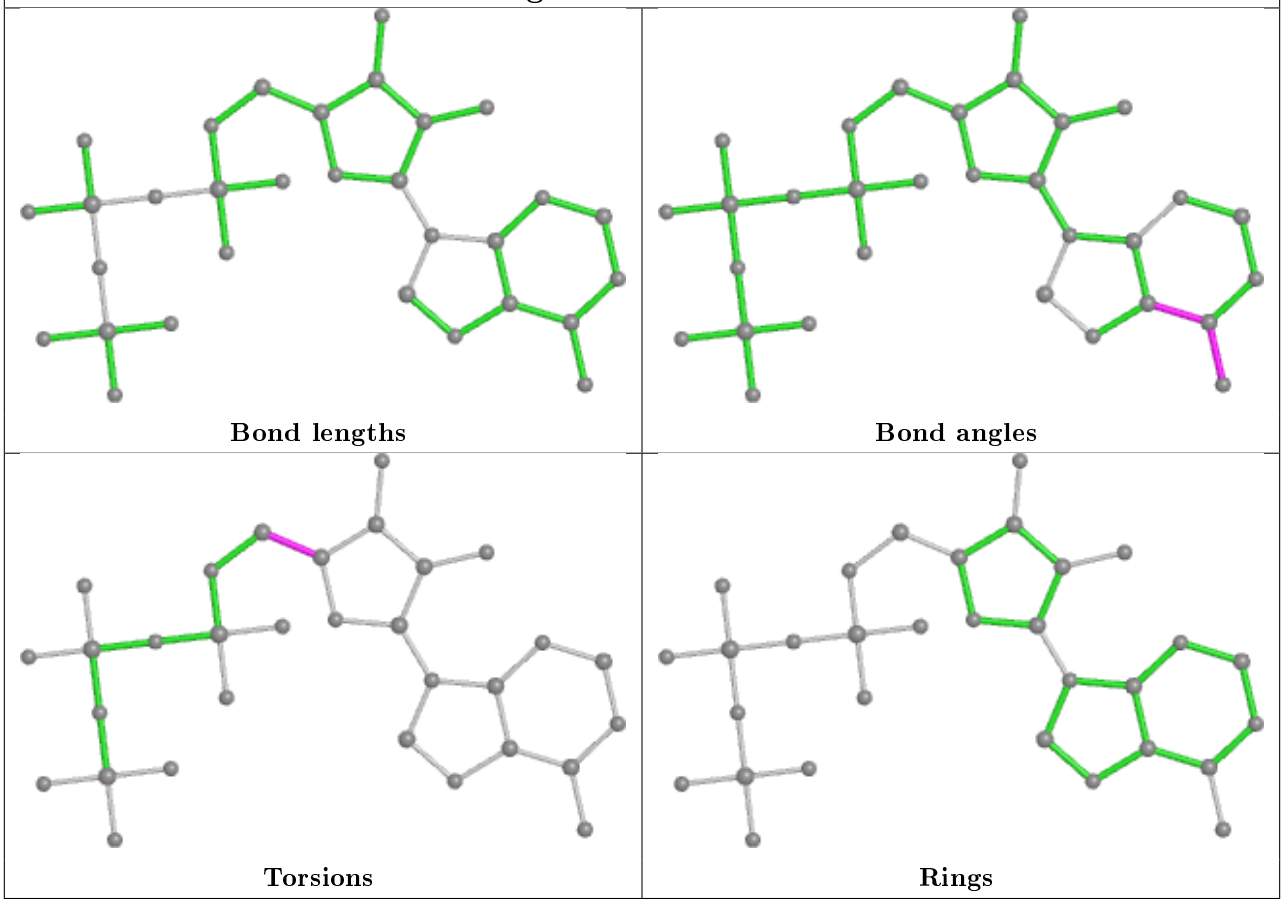


Torsions



Rings

Ligand ATP B 905



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

6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

6.4 Ligands

EDS was not executed - this section is therefore empty.

6.5 Other polymers

EDS was not executed - this section is therefore empty.