wwPDB X-ray Structure Validation Summary Report (i)

May 14, 2020 – 06:06 pm BST

PDB ID	:	1DO2
Title	:	TRIGONAL 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	:	4.00 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:

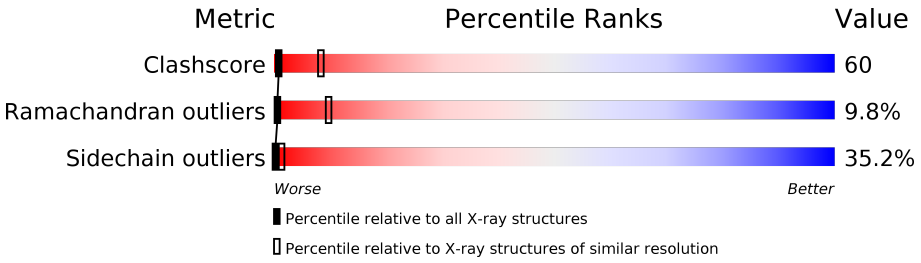
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, 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 4.00 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	1148 (4.30-3.70)
Ramachandran outliers	138981	1108 (4.30-3.70)
Sidechain outliers	138945	1099 (4.30-3.70)

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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	442	17% 52% 21% • 8%
1	B	442	22% 50% 18% • 8%
1	C	442	16% 50% 24% • 8%
1	D	442	21% 53% 16% • 8%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	ANP	A	900	-	-	X	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	ANP	C	905	-	-	X	-

2 Entry composition

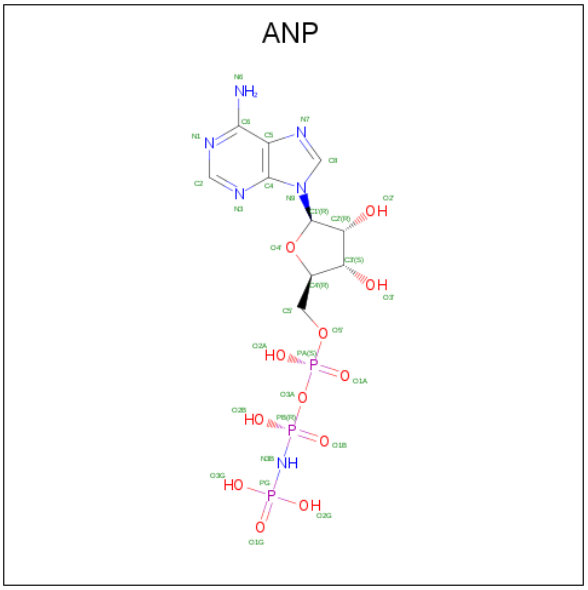
There are 2 unique types of molecules in this entry. The entry contains 12926 atoms, of which 0 are hydrogens and 0 are deuteriums.

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	407	Total	C	N	O	S	126	0	0
			3216	2008	574	624	10			
1	B	407	Total	C	N	O	S	492	0	0
			3216	2008	574	624	10			
1	C	407	Total	C	N	O	S	103	0	0
			3216	2008	574	624	10			
1	D	407	Total	C	N	O	S	484	0	0
			3216	2008	574	624	10			

- Molecule 2 is PHOSPHOAMINOPHOSPHONIC ACID-ADENYLATE ESTER (three-letter code: ANP) (formula: $C_{10}H_{17}N_6O_{12}P_3$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			31	10	6	12	3		
2	C	1	Total	C	N	O	P	0	0
			31	10	6	12	3		

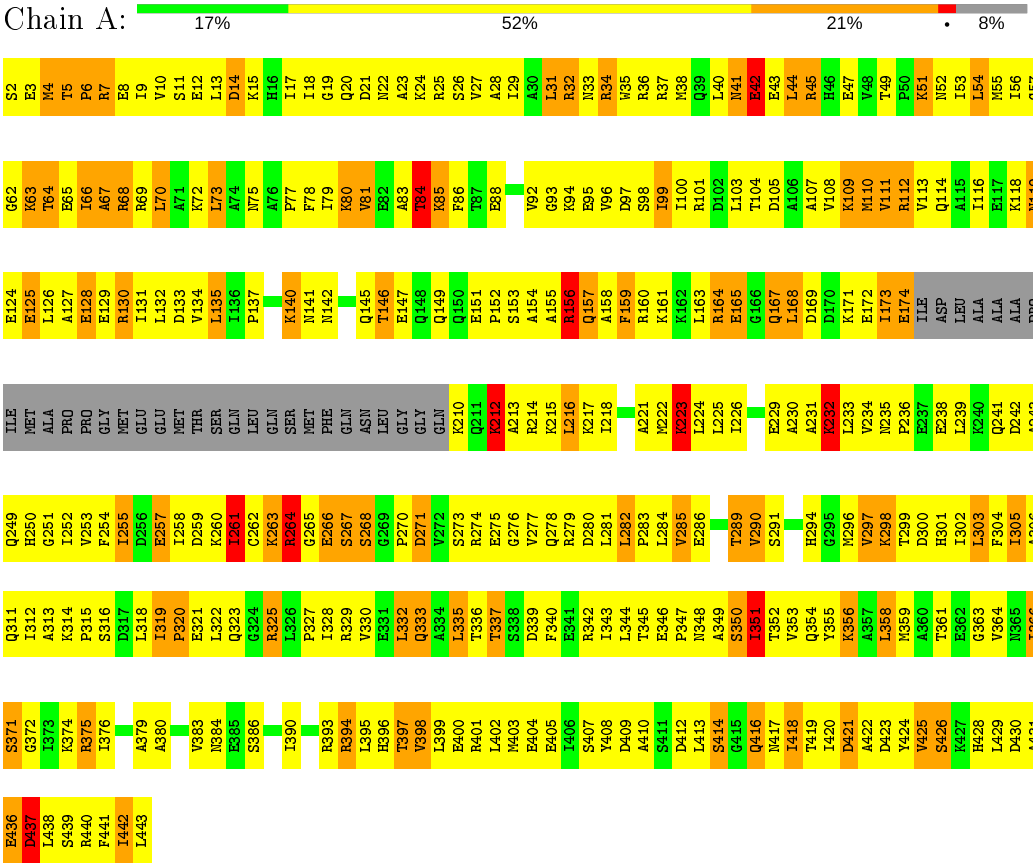
3 Residue-property plots (i)

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.

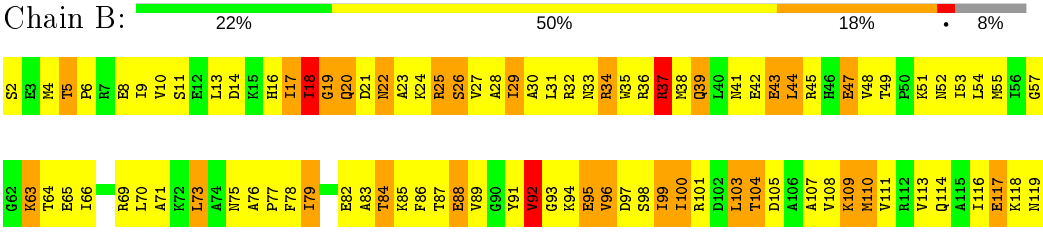
- Molecule 1: PROTEIN (HEAT SHOCK LOCUS U)

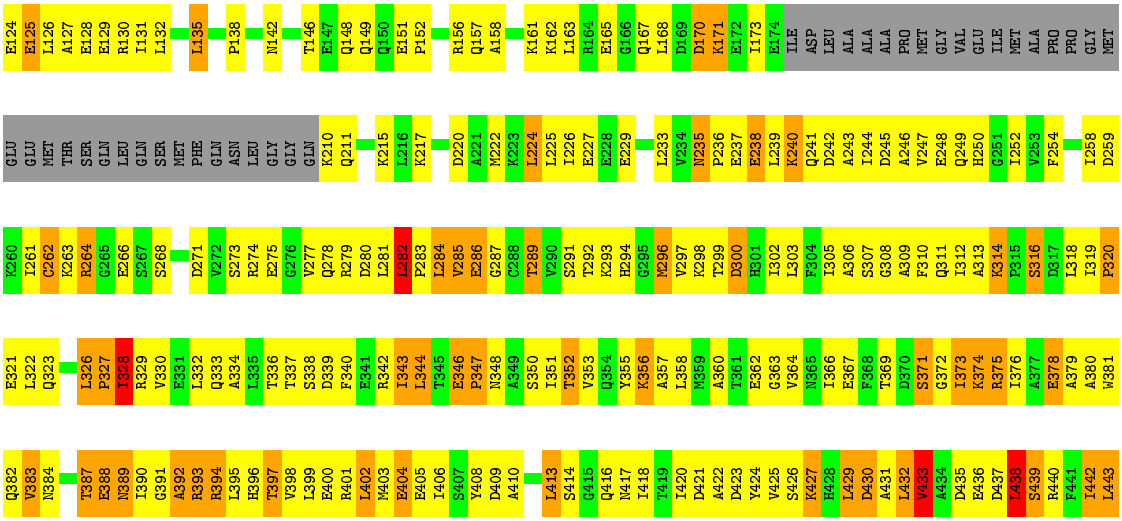
Chain A:



- Molecule 1: PROTEIN (HEAT SHOCK LOCUS U)

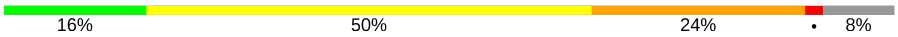
Chain B:

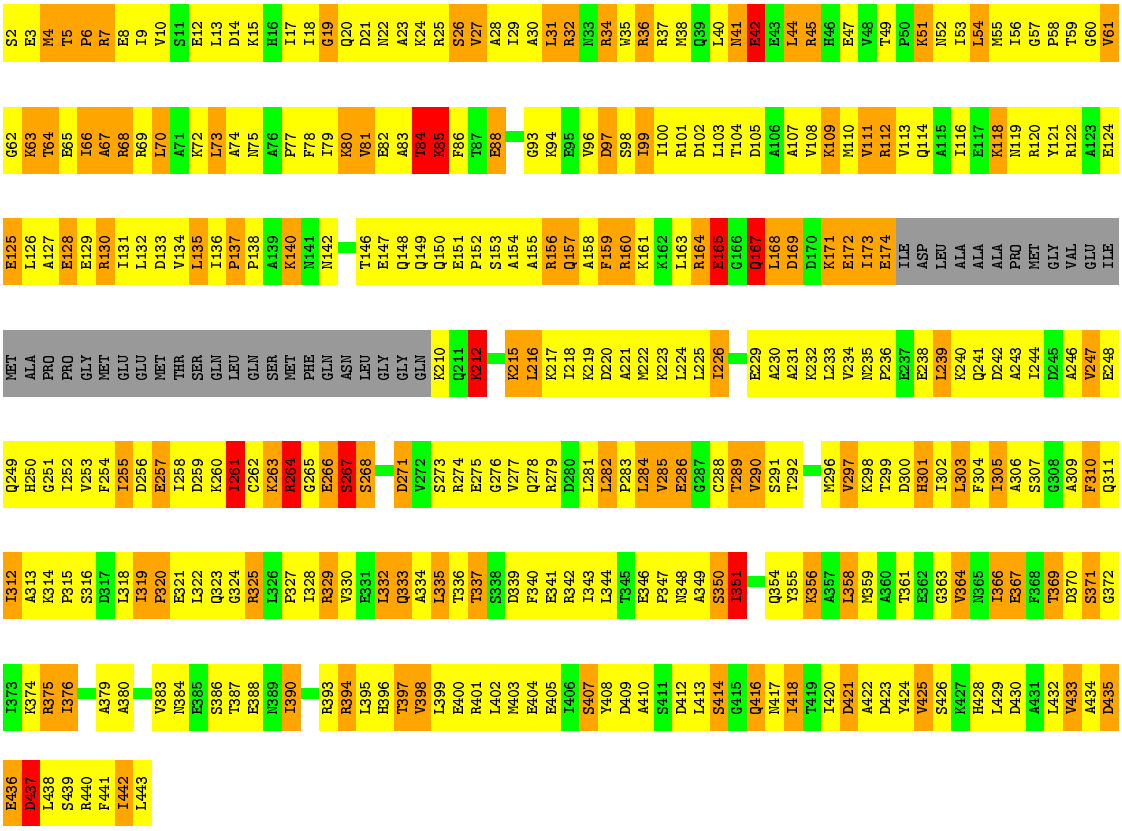




• Molecule 1: PROTEIN (HEAT SHOCK LOCUS U)

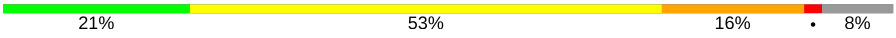
Chain C:

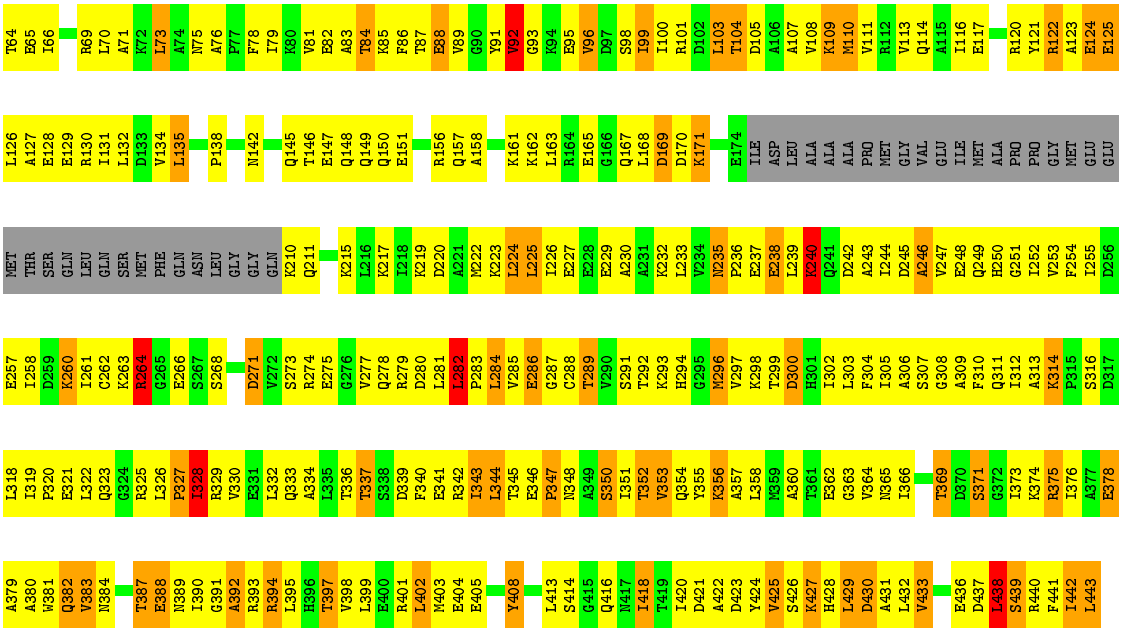




• Molecule 1: PROTEIN (HEAT SHOCK LOCUS U)

Chain D:




T64 E65 I66 R69 L70 A71 K72 L73 A74 N75 A76 P77 F78 I79 K80 V81 E82 A83 T84 K85 F86 T87 E88 V89 G90 Y91 V92 G93 K94 E95 V96 D97 S98 I99 I100 R101 D102 L103 T104 D105 A106 A107 V108 K109 M110 V111 R112 V113 Q114 A115 I116 E117 R120 Y121 R122 A123 E124 E125
L126 A127 E128 E129 R130 I131 L132 D133 V134 L135 P138 N142 Q145 T146 E147 Q148 Q149 Q150 E151 R156 Q157 A158 K161 K162 L163 R164 E165 G166 Q167 L168 D169 D170 K171 E174 ILE ASP LEU ALA ALA ALA PRO MET GLY VAL GLU ILE MET ALA PRO PRO GLY MET GLU GLU
MET THR SER GLN LEU GLN SER MET PHE GLN ASN LEU GLY GLY GLN K210 Q211 K215 L216 K217 I218 K219 D220 A221 M222 K223 L224 L225 I226 E227 E228 E229 A230 A231 K232 L233 V234 N235 P236 E237 E238 L239 K240 Q241 D242 A243 I244 D245 A246 V247 E248 Q249 H250 G251 I252 V253 F254 I255 D256
E257 I258 D259 K260 I261 C262 K263 R264 G265 E266 S267 S268 D271 V272 S273 R274 E275 G276 V277 Q278 R279 D280 L281 L282 P283 L284 V285 E286 G287 C288 T289 V290 S291 T292 K293 H294 G295 M296 V297 K298 T299 D300 H301 I302 L303 F304 I305 A306 S307 G308 A309 F310 Q311 I312 A313 K314 P315 S316 D317
L318 I319 P320 E321 L322 Q323 G324 R325 L326 P327 I328 R329 V330 E331 L332 Q333 A334 L335 T336 T337 S338 D339 F340 E341 R342 I343 L344 T345 E346 P347 N348 A349 S350 I351 T352 V353 Q354 Y355 K356 A357 L358 M359 A360 T361 E362 G363 V364 N365 I366 T369 D370 S371 G372 I373 K374 R375 I376 A377 E378
A379 A380 W381 Q382 V383 N384 T387 E388 N389 I390 G391 A392 R393 R394 L395 H396 T397 V398 L399 E400 R401 L402 M403 E404 E405 Y408 L413 S414 G415 Q416 N417 I418 T419 I420 D421 A422 D423 Y424 V425 S426 K427 H428 L429 D430 A431 L432 V433 E436 D437 L438 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 3 2 1	Depositor
Cell constants a, b, c, α, β, γ	201.78Å 201.78Å 171.63Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	15.00 – 4.00	Depositor
% Data completeness (in resolution range)	95.8 (15.00-4.00)	Depositor
R_{merge}	0.10	Depositor
R_{sym}	10.50	Depositor
Refinement program	X-PLOR 3.851	Depositor
R, R_{free}	0.229 , 0.285	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	12926	wwPDB-VP
Average B, all atoms (Å^2)	59.0	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: ANP

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	0.30	0/3255	0.58	3/4385 (0.1%)
1	B	0.28	0/3255	0.50	0/4385
1	C	0.29	0/3255	0.51	0/4385
1	D	0.28	0/3255	0.52	0/4385
All	All	0.29	0/13020	0.53	3/17540 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	156	ARG	CB-CG-CD	12.03	142.87	111.60
1	A	223	LYS	CB-CG-CD	8.81	134.51	111.60
1	A	232	LYS	N-CA-CB	5.85	121.13	110.60

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	3216	0	3284	386	0
1	B	3216	0	3284	289	0
1	C	3216	0	3284	434	0
1	D	3216	0	3284	316	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	31	0	13	9	0
2	C	31	0	13	11	0
All	All	12926	0	13162	1411	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 60.

The worst 5 of 1411 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:356:LYS:HA	1:A:366:ILE:HD11	1.38	1.01
1:C:54:LEU:HD13	1:C:56:ILE:HD11	1.39	1.01
1:A:366:ILE:HD13	1:A:366:ILE:H	1.28	0.98
1:A:54:LEU:HD13	1:A:56:ILE:HD11	1.43	0.97
1:A:282:LEU:HD21	1:A:321:GLU:HB3	1.46	0.97

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	403/442 (91%)	262 (65%)	102 (25%)	39 (10%)	0 10
1	B	403/442 (91%)	251 (62%)	111 (28%)	41 (10%)	0 9
1	C	403/442 (91%)	258 (64%)	104 (26%)	41 (10%)	0 9
1	D	403/442 (91%)	251 (62%)	115 (28%)	37 (9%)	1 11
All	All	1612/1768 (91%)	1022 (63%)	432 (27%)	158 (10%)	0 10

5 of 158 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	45	ARG
1	A	63	LYS
1	A	110	MET
1	A	268	SER
1	A	398	VAL

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	349/376 (93%)	226 (65%)	123 (35%)	0	1
1	B	349/376 (93%)	227 (65%)	122 (35%)	0	1
1	C	349/376 (93%)	220 (63%)	129 (37%)	0	0
1	D	349/376 (93%)	231 (66%)	118 (34%)	0	2
All	All	1396/1504 (93%)	904 (65%)	492 (35%)	0	1

5 of 492 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	374	LYS
1	C	94	LYS
1	D	316	SER
1	B	389	ASN
1	C	7	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 36 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	323	GLN
1	C	39	GLN
1	D	235	ASN
1	B	384	ASN
1	C	75	ASN

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

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	ANP	A	900	-	29,33,33	1.21	3 (10%)	31,52,52	1.53	2 (6%)
2	ANP	C	905	-	29,33,33	1.17	4 (13%)	31,52,52	1.52	3 (9%)

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	ANP	A	900	-	-	6/14/38/38	0/3/3/3
2	ANP	C	905	-	-	6/14/38/38	0/3/3/3

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	900	ANP	PG-O1G	2.81	1.50	1.46

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	900	ANP	PG-O2G	-2.59	1.49	1.56
2	C	905	ANP	PG-O1G	2.58	1.50	1.46
2	C	905	ANP	PG-O2G	-2.58	1.49	1.56
2	A	900	ANP	PB-O2B	-2.35	1.50	1.56

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	900	ANP	O2B-PB-O1B	4.88	120.14	109.92
2	C	905	ANP	O2B-PB-O1B	4.87	120.14	109.92
2	A	900	ANP	O1G-PG-N3B	-4.23	105.55	111.77
2	C	905	ANP	O1G-PG-N3B	-4.19	105.61	111.77
2	C	905	ANP	C5-C6-N6	2.05	123.46	120.35

There are no chirality outliers.

5 of 12 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	900	ANP	PB-N3B-PG-O1G
2	A	900	ANP	PG-N3B-PB-O1B
2	A	900	ANP	PG-N3B-PB-O3A
2	C	905	ANP	PB-N3B-PG-O1G
2	C	905	ANP	PG-N3B-PB-O1B

There are no ring outliers.

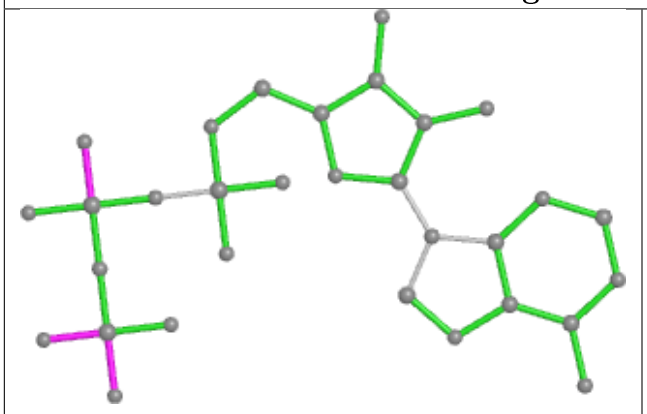
2 monomers are involved in 20 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	900	ANP	9	0
2	C	905	ANP	11	0

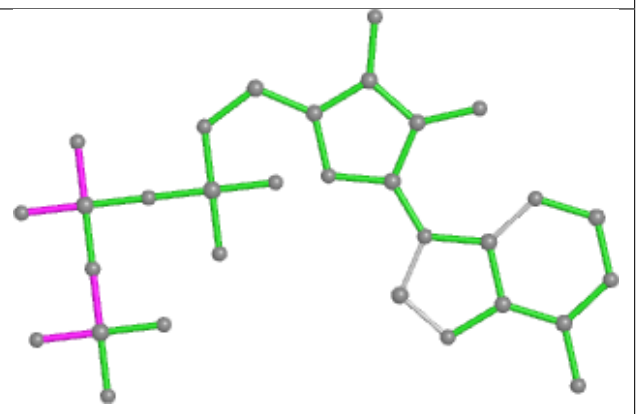
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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

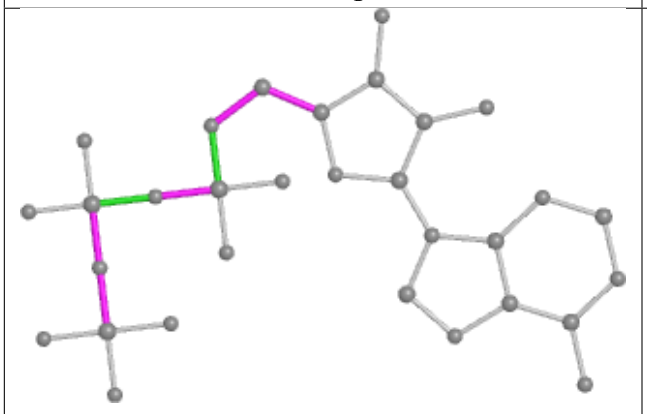
Ligand ANP A 900



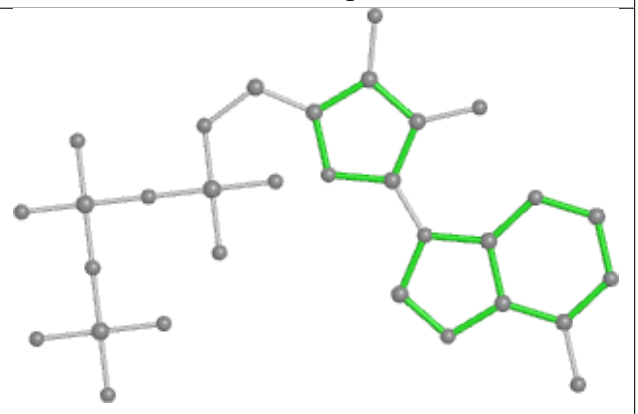
Bond lengths



Bond angles

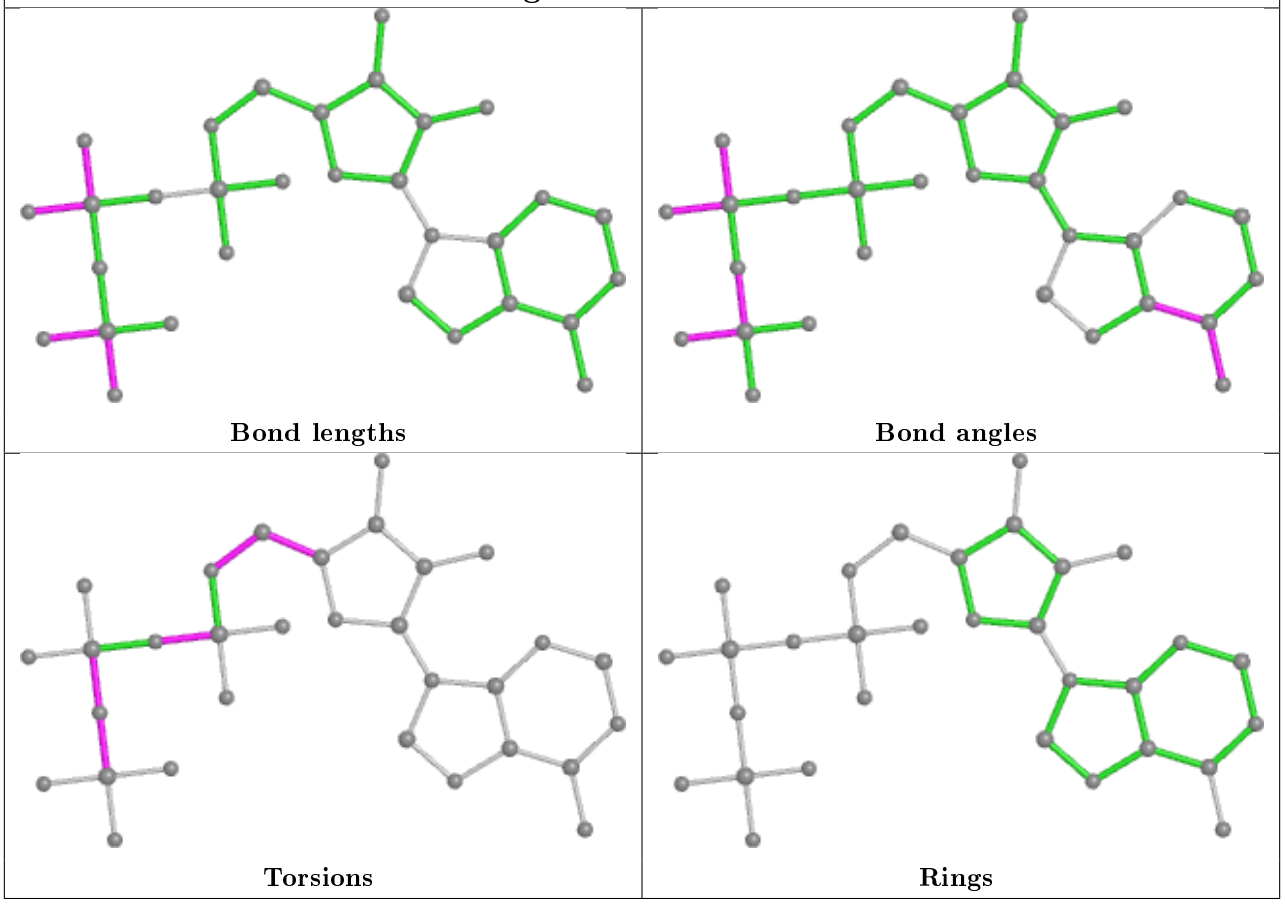


Torsions



Rings

Ligand ANP C 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.