wwPDB X-ray Structure Validation Summary Report (i)

Jan 31, 2016 – 08:51 PM GMT

PDB ID	:	1MAW
Title	:	Crystal Structure of Tryptophanyl-tRNA Synthetase Complexed with ATP in an Open Conformation
Authors	:	Retailleau, P.; Huang, X.; Yin, Y.; Hu, M.; Weinreb, V.; Vachette, P.; Vonrhein, C.; Bricogne, G.; Roversi, P.; Ilyin, V.; Carter Jr., C.W.
Deposited on	:	2002-08-02
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
http://wwpdb.org/validation/2016/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:

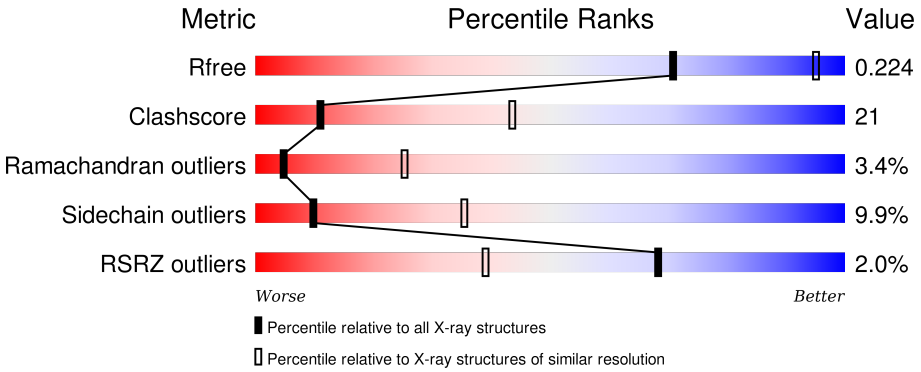
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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.



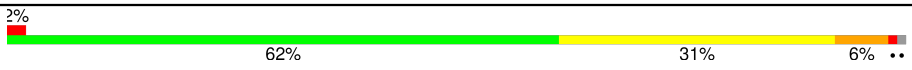
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	1578 (3.00-3.00)
Clashscore	102246	1912 (3.00-3.00)
Ramachandran outliers	100387	1853 (3.00-3.00)
Sidechain outliers	100360	1856 (3.00-3.00)
RSRZ outliers	91569	1592 (3.00-3.00)

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. 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	328	2% 56% 38% 5% •
1	B	328	2% 57% 37% 5% ••
1	C	328	2% 55% 38% 5% ••
1	D	328	2% 58% 36% 5% ••
1	E	328	2% 58% 35% 5% ••

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	328	

2 Entry composition

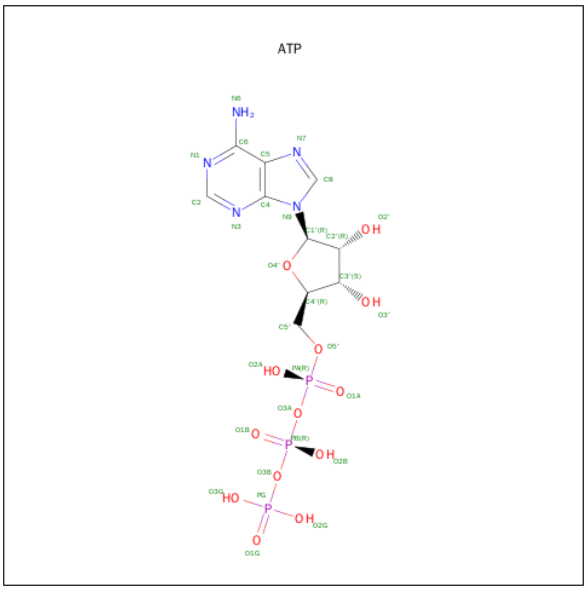
There are 2 unique types of molecules in this entry. The entry contains 15564 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 TRYPTOPHAN-TRNA LIGASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	326	Total	C	N	O	S	0	0	0
			2563	1624	442	484	13			
1	B	326	Total	C	N	O	S	0	0	0
			2563	1624	442	484	13			
1	C	326	Total	C	N	O	S	0	0	0
			2563	1624	442	484	13			
1	D	326	Total	C	N	O	S	0	0	0
			2563	1624	442	484	13			
1	E	326	Total	C	N	O	S	0	0	0
			2563	1624	442	484	13			
1	F	326	Total	C	N	O	S	0	0	0
			2563	1624	442	484	13			

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



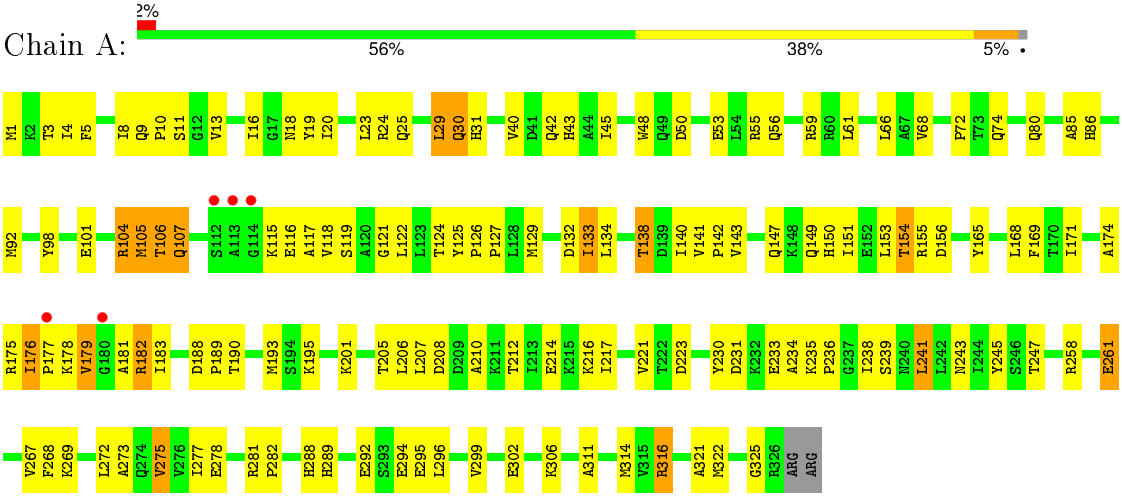
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
2	B	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
2	D	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
2	E	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
2	C	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
2	F	1	Total	C	N	O	P	0	0
			31	10	5	13	3		

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 errors 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.

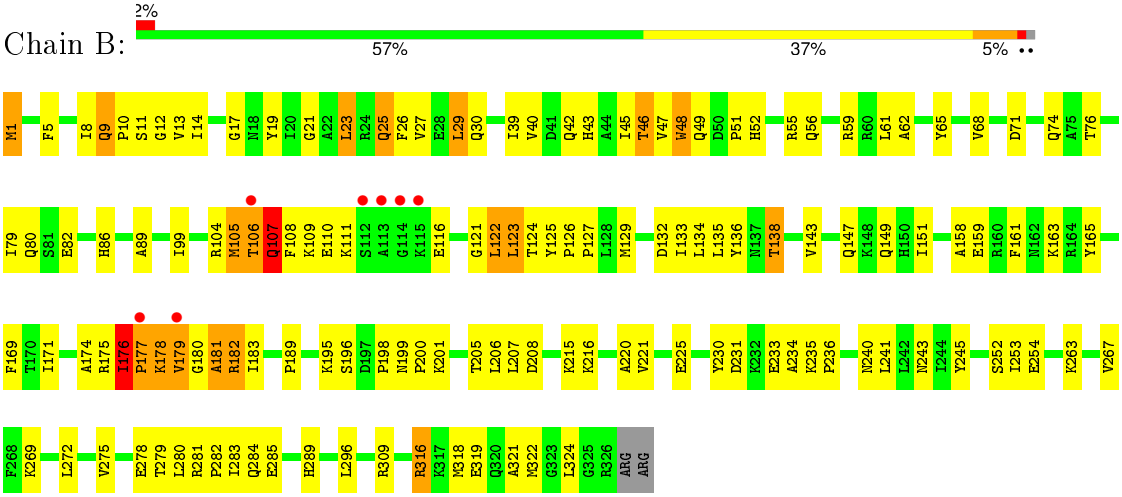
- Molecule 1: TRYPTOPHAN-TRNA LIGASE

Chain A:



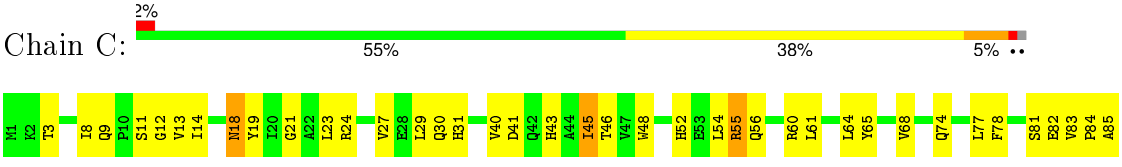
- Molecule 1: TRYPTOPHAN-TRNA LIGASE

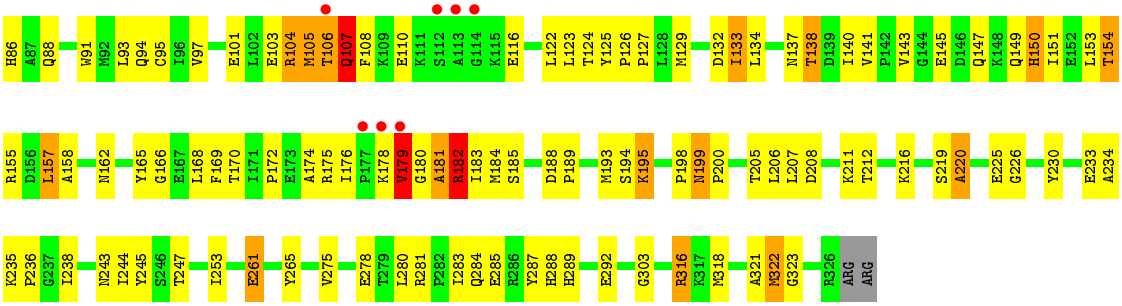
Chain B:



- Molecule 1: TRYPTOPHAN-TRNA LIGASE

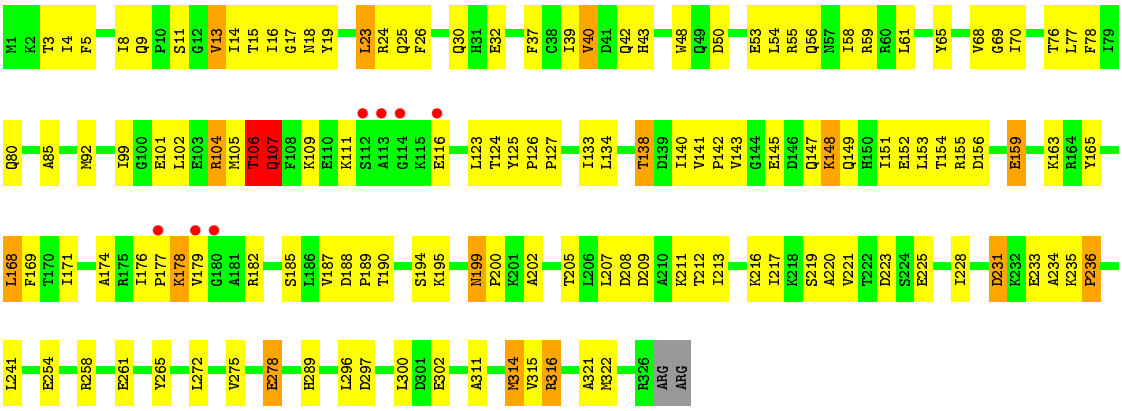
Chain C:





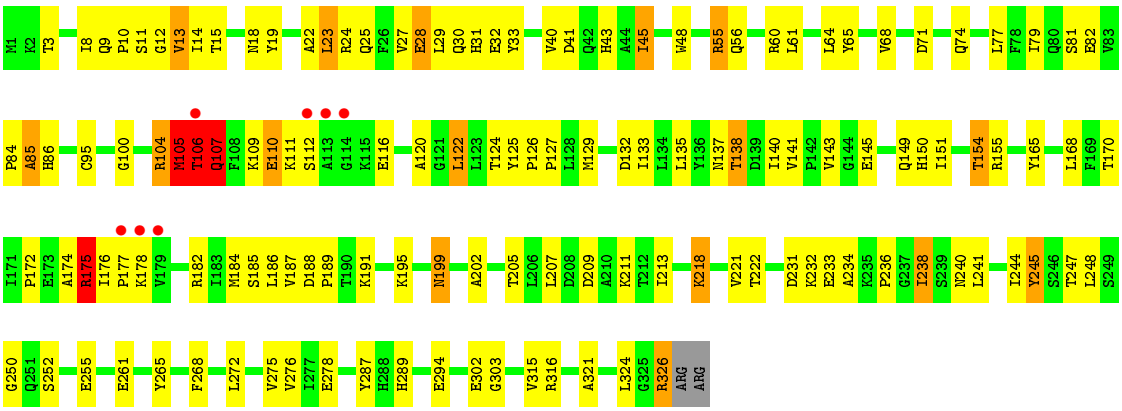
• Molecule 1: TRYPTOPHAN-TRNA LIGASE

Chain D:



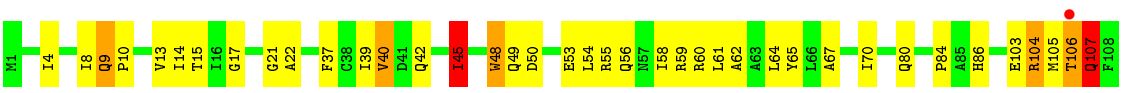
• Molecule 1: TRYPTOPHAN-TRNA LIGASE

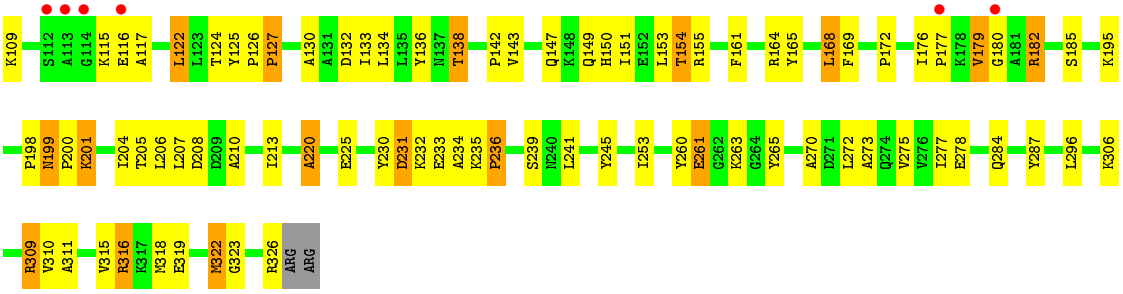
Chain E:



• Molecule 1: TRYPTOPHAN-TRNA LIGASE

Chain F:




K109
S112
A113
G114
K115
E116
A117
L122
L123
T124
Y125
P126
P127
A130
A131
D132
I133
L134
L135
Y136
N137
T138
P142
V143
Q147
K148
Q149
H150
I151
E152
L153
T154
R155
F161
R164
Y165
L168
F169
P172
I176
P177
K178
V179
G180
A181
R182
S185
K195
P198
N199
P200
K201
I204
T205
L206
L207
D208
D209
A210
I213
A220
E225
Y230
D231
K232
E233
A234
K235
P236
S239
N240
L241
Y245
I253
Y260
E261
G262
K263
G264
Y265
A270
D271
L272
A273
Q274
V275
V276
I277
E278
Q284
Y287
L296
K306
R309
V310
A311
V315
R316
K317
N318
E319
N322
G323
R326
ARG
ARG

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	228.55Å 91.98Å 156.89Å 90.00° 132.34° 90.00°	Depositor
Resolution (Å)	15.00 – 3.00 15.05 – 2.99	Depositor EDS
% Data completeness (in resolution range)	88.4 (15.00-3.00) 87.6 (15.05-2.99)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.08	Depositor
$< I/\sigma(I) >$ [1]	2.24 (at 3.01Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.218 , 0.253 0.193 , 0.224	Depositor DCC
R_{free} test set	4285 reflections (11.22%)	DCC
Wilson B-factor (Å^2)	45.6	Xtriage
Anisotropy	0.316	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 71.3	EDS
Estimated twinning fraction	0.022 for -h-2*l,-k,l	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.28$	Xtriage
Outliers	0 of 42607 reflections	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	15564	wwPDB-VP
Average B, all atoms (Å^2)	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 27.27 % of the origin peak, indicating pseudo translational symmetry.The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 2.2682e-03.The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: 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	0.49	0/2611	0.73	0/3532
1	B	0.47	0/2611	0.73	0/3532
1	C	0.49	0/2611	0.76	0/3532
1	D	0.50	0/2611	0.75	0/3532
1	E	0.50	0/2611	0.75	0/3532
1	F	0.45	0/2611	0.73	1/3532 (0.0%)
All	All	0.48	0/15666	0.74	1/21192 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	F	180	GLY	N-CA-C	-5.67	98.93	113.10

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	2563	0	2554	104	0
1	B	2563	0	2554	131	0
1	C	2563	0	2554	111	0
1	D	2563	0	2554	117	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	E	2563	0	2554	102	0
1	F	2563	0	2554	104	0
2	A	31	0	12	2	0
2	B	31	0	12	2	0
2	C	31	0	12	3	0
2	D	31	0	12	3	0
2	E	31	0	12	5	0
2	F	31	0	12	4	0
All	All	15564	0	15396	647	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 21.

The worst 5 of 647 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:F:205:THR:HG22	1:F:207:LEU:H	1.17	1.05
1:F:8:ILE:HG22	1:F:61:LEU:HD21	1.48	0.93
1:F:309:ARG:HD3	1:F:310:VAL:HG23	1.51	0.92
1:F:106:THR:H	1:F:149:GLN:HE22	1.18	0.92
1:B:205:THR:HG22	1:B:207:LEU:H	1.34	0.89

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	324/328 (99%)	280 (86%)	33 (10%)	11 (3%)	5 25
1	B	324/328 (99%)	276 (85%)	36 (11%)	12 (4%)	4 23
1	C	324/328 (99%)	275 (85%)	36 (11%)	13 (4%)	4 21

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	D	324/328 (99%)	276 (85%)	41 (13%)	7 (2%)	8 38
1	E	324/328 (99%)	282 (87%)	30 (9%)	12 (4%)	4 23
1	F	324/328 (99%)	285 (88%)	27 (8%)	12 (4%)	4 23
All	All	1944/1968 (99%)	1674 (86%)	203 (10%)	67 (3%)	5 25

5 of 67 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	107	GLN
1	B	107	GLN
1	B	176	ILE
1	B	179	VAL
1	C	105	MET

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	270/280 (96%)	244 (90%)	26 (10%)	10 38
1	B	270/280 (96%)	246 (91%)	24 (9%)	12 42
1	C	270/280 (96%)	239 (88%)	31 (12%)	7 28
1	D	270/280 (96%)	242 (90%)	28 (10%)	9 32
1	E	270/280 (96%)	239 (88%)	31 (12%)	7 28
1	F	270/280 (96%)	249 (92%)	21 (8%)	16 49
All	All	1620/1680 (96%)	1459 (90%)	161 (10%)	10 35

5 of 161 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	195	LYS
1	D	107	GLN
1	F	138	THR
1	C	245	TYR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	9	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 52 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	251	GLN
1	D	149	GLN
1	F	56	GLN
1	C	259	GLN
1	D	43	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	ATP	A	400	-	24,33,33	2.17	5 (20%)	31,52,52	3.07	12 (38%)
2	ATP	B	401	-	24,33,33	2.06	7 (29%)	31,52,52	3.51	15 (48%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
2	ATP	C	404	-	24,33,33	2.16	7 (29%)	31,52,52	3.50	12 (38%)
2	ATP	D	402	-	24,33,33	2.18	5 (20%)	31,52,52	3.37	11 (35%)
2	ATP	E	403	-	24,33,33	2.19	7 (29%)	31,52,52	3.51	13 (41%)
2	ATP	F	405	-	24,33,33	2.11	7 (29%)	31,52,52	3.12	8 (25%)

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	A	400	-	-	0/18/38/38	0/3/3/3
2	ATP	B	401	-	-	0/18/38/38	0/3/3/3
2	ATP	C	404	-	-	0/18/38/38	0/3/3/3
2	ATP	D	402	-	-	0/18/38/38	0/3/3/3
2	ATP	E	403	-	-	0/18/38/38	0/3/3/3
2	ATP	F	405	-	-	0/18/38/38	0/3/3/3

The worst 5 of 38 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	402	ATP	O5'-C5'	-4.47	1.26	1.44
2	A	400	ATP	O5'-C5'	-4.42	1.26	1.44
2	F	405	ATP	O5'-C5'	-4.35	1.27	1.44
2	E	403	ATP	O5'-C5'	-4.18	1.27	1.44
2	C	404	ATP	O5'-C5'	-4.07	1.28	1.44

The worst 5 of 71 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	400	ATP	O5'-PA-O1A	-8.10	78.19	109.62
2	B	401	ATP	O5'-PA-O1A	-8.08	78.27	109.62
2	F	405	ATP	O5'-PA-O1A	-7.85	79.15	109.62
2	C	404	ATP	O5'-PA-O1A	-7.77	79.47	109.62
2	E	403	ATP	O5'-PA-O1A	-7.58	80.21	109.62

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6 monomers are involved in 19 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	400	ATP	2	0
2	B	401	ATP	2	0
2	C	404	ATP	3	0
2	D	402	ATP	3	0
2	E	403	ATP	5	0
2	F	405	ATP	4	0

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

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	326/328 (99%)	-0.63	5 (1%)	76	49	17, 37, 80, 102	0
1	B	326/328 (99%)	-0.61	7 (2%)	67	36	11, 38, 78, 99	0
1	C	326/328 (99%)	-0.64	7 (2%)	67	36	13, 35, 77, 96	0
1	D	326/328 (99%)	-0.58	7 (2%)	67	36	14, 37, 79, 102	0
1	E	326/328 (99%)	-0.59	7 (2%)	67	36	13, 34, 79, 103	0
1	F	326/328 (99%)	-0.56	7 (2%)	67	36	12, 41, 79, 100	0
All	All	1956/1968 (99%)	-0.60	40 (2%)	68	39	11, 37, 81, 103	0

The worst 5 of 40 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	179	VAL	5.5
1	F	114	GLY	5.1
1	F	180	GLY	5.0
1	A	177	PRO	5.0
1	F	113	ALA	4.9

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

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

6.3 Carbohydrates

There are no carbohydrates in this entry.

6.4 Ligands

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
2	ATP	C	404	31/31	0.90	0.23	0.56	53,63,78,78	31
2	ATP	F	405	31/31	0.89	0.25	0.43	57,61,79,80	31
2	ATP	B	401	31/31	0.89	0.23	0.35	51,56,74,75	31
2	ATP	D	402	31/31	0.91	0.23	0.32	49,60,75,76	31
2	ATP	E	403	31/31	0.91	0.25	0.32	50,58,78,80	31
2	ATP	A	400	31/31	0.86	0.26	0.26	46,59,81,83	31

6.5 Other polymers

There are no such residues in this entry.