wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 03:12 PM GMT

PDB ID	:	4BUB
Title	:	CRYSTAL STRUCTURE OF MURE LIGASE FROM THERMOTOGA MARITIMA IN COMPLEX WITH ADP
Authors	:	Favini-Stabile, S.; Contreras-Martel, C.; Thielens, N.; Dessen, A.
Deposited on	:	2013-06-20
Resolution	:	2.90 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:

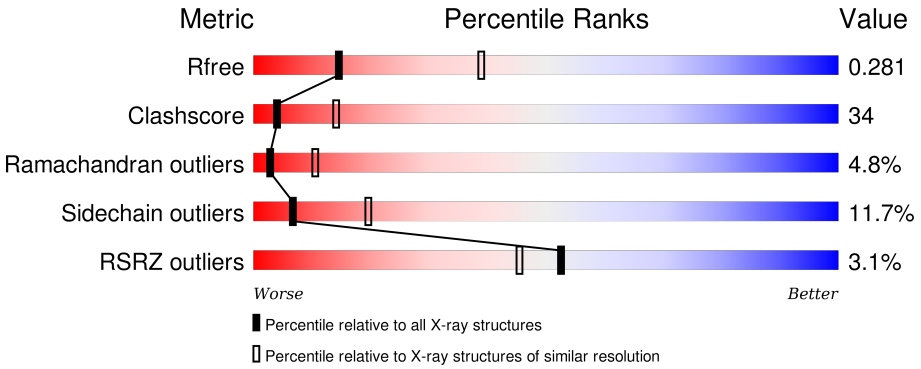
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 2.90 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	1451 (2.90-2.90)
Clashscore	102246	1668 (2.90-2.90)
Ramachandran outliers	100387	1630 (2.90-2.90)
Sidechain outliers	100360	1632 (2.90-2.90)
RSRZ outliers	91569	1456 (2.90-2.90)

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	498	2% 48% 41% 7%
1	B	498	4% 41% 47% 9%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 7620 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.

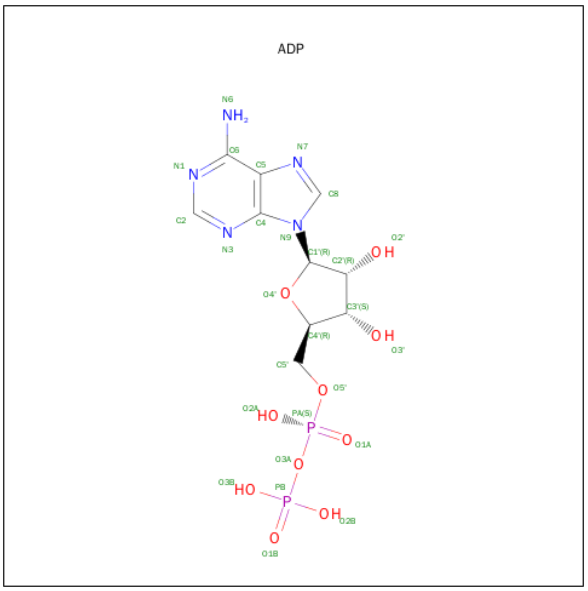
- Molecule 1 is a protein called UDP-N-ACETYLMURAMOYL-L-ALANYL-D-GLUTAMATE--LD-LYSINE LIGASE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	478	Total	C	N	O	S	Se	0	0	0
			3755	2382	648	715	1	9			
1	B	483	Total	C	N	O	S	Se	0	0	0
			3792	2402	655	725	1	9			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	491	LEU	-	EXPRESSION TAG	UNP Q9WY79
A	492	GLU	-	EXPRESSION TAG	UNP Q9WY79
A	493	HIS	-	EXPRESSION TAG	UNP Q9WY79
A	494	HIS	-	EXPRESSION TAG	UNP Q9WY79
A	495	HIS	-	EXPRESSION TAG	UNP Q9WY79
A	496	HIS	-	EXPRESSION TAG	UNP Q9WY79
A	497	HIS	-	EXPRESSION TAG	UNP Q9WY79
A	498	HIS	-	EXPRESSION TAG	UNP Q9WY79
B	491	LEU	-	EXPRESSION TAG	UNP Q9WY79
B	492	GLU	-	EXPRESSION TAG	UNP Q9WY79
B	493	HIS	-	EXPRESSION TAG	UNP Q9WY79
B	494	HIS	-	EXPRESSION TAG	UNP Q9WY79
B	495	HIS	-	EXPRESSION TAG	UNP Q9WY79
B	496	HIS	-	EXPRESSION TAG	UNP Q9WY79
B	497	HIS	-	EXPRESSION TAG	UNP Q9WY79
B	498	HIS	-	EXPRESSION TAG	UNP Q9WY79

- Molecule 2 is ADENOSINE-5'-DIPHOSPHATE (three-letter code: ADP) (formula: $C_{10}H_{15}N_5O_{10}P_2$).



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

- 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	A	1	Total	Mg	0	0
			1	1		

- Molecule 4 is water.

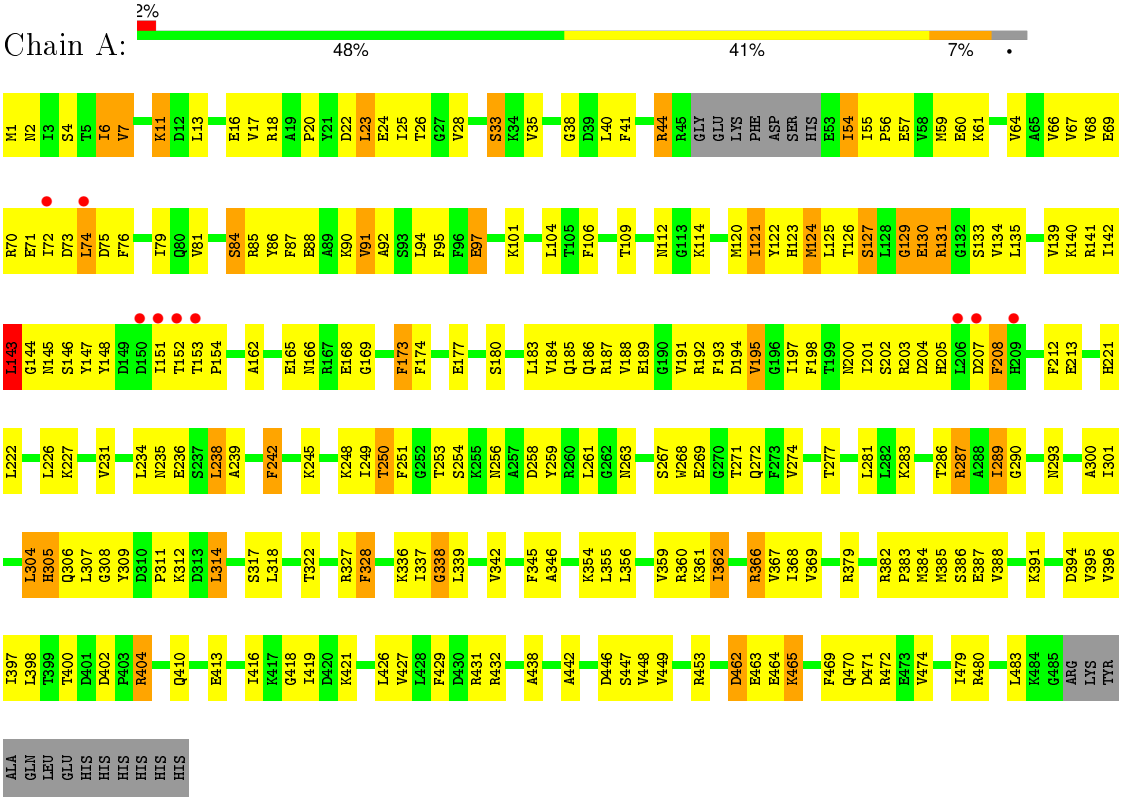
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	10	Total	O	0	0
			10	10		
4	B	7	Total	O	0	0
			7	7		

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.

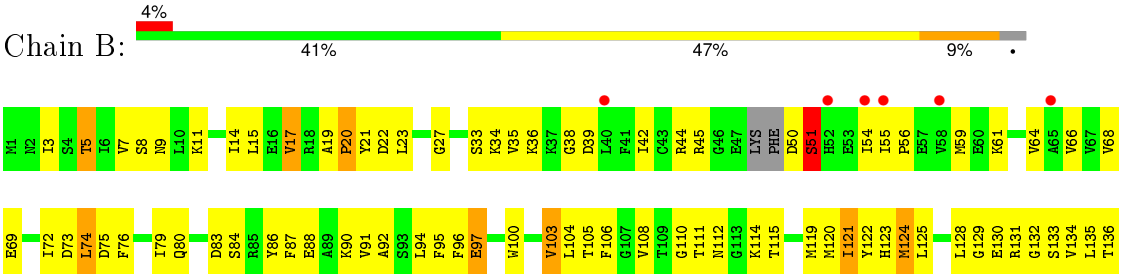
• Molecule 1: UDP-N-ACETYLMURAMOYL-L-ALANYL-D-GLUTAMATE--LD-LYSINE LIGASE

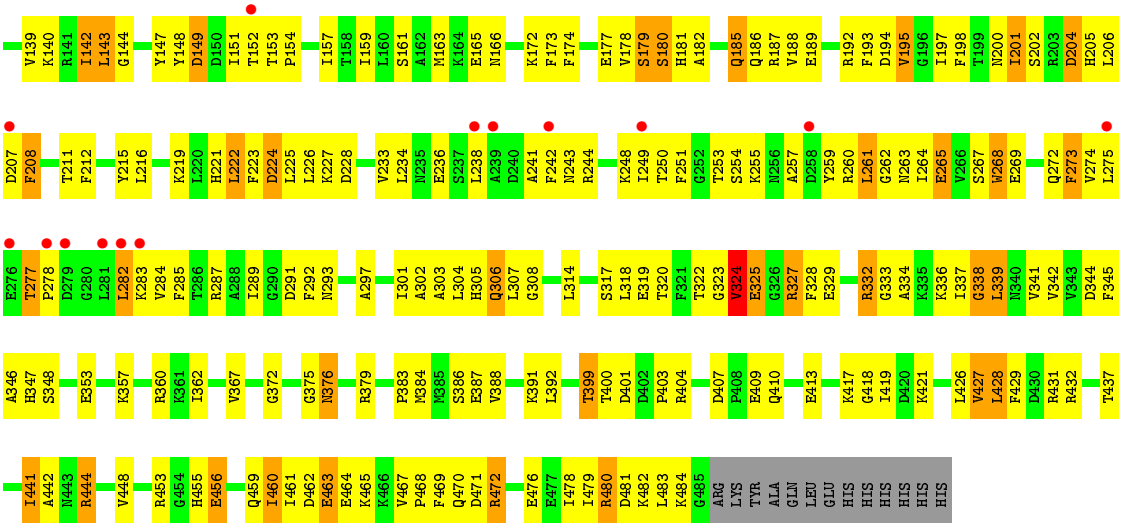
Chain A:



• Molecule 1: UDP-N-ACETYLMURAMOYL-L-ALANYL-D-GLUTAMATE--LD-LYSINE LIGASE

Chain B:




V139 K140 R141 I142 L143 G144 Y147 Y148 D149 D150 I151 T152 T153 P154 I157 T158 I159 L160 S161 A162 M163 K164 E165 N166 K172 F173 F174 E177 V178 S179 S180 H181 A182 Q185 Q186 R187 V188 E189 R192 F193 D194 V195 G196 I197 F198 T199 N200 I201 S202 R203 D204 H205 L206
D207 F208 T211 F212 Y215 L216 K219 L220 H221 L222 F223 D224 L225 L226 K227 D228 V233 L234 N235 E236 S237 L238 A239 D240 A241 F242 N243 R244 K248 I249 T250 F251 G252 T253 S254 K255 N256 A257 D258 Y259 R260 L261 G262 N263 I264 E265 V266 S267 W268 E269 Q272 F273 V274 L275
E276 T277 P278 D279 G280 L281 L282 K283 V284 F285 T286 R287 A288 I289 G290 D291 F292 N293 A297 I301 A302 A303 L304 H305 Q306 L307 G308 L314 S317 L318 E319 T320 F321 T322 G323 V324 E325 G326 R327 F328 E329 R332 G333 A334 K335 K336 I337 G338 L339 N340 V341 V342 V343 D344 F345
A346 H347 S348 E353 K357 R360 K361 I362 V367 G372 G375 N376 R379 P383 M384 M385 S386 E387 V388 K391 I392 T399 T400 D401 D402 P403 R404 D407 P408 E409 Q410 E413 K417 G418 I419 D420 K421 L426 V427 L428 F429 D430 R431 R432 T437
I441 A442 N443 R444 V448 R453 G454 H455 E456 Q459 I460 I461 D462 E463 E464 K465 K466 V467 P468 F469 Q470 D471 R472 E476 E477 I478 I479 R480 D481 K482 L483 K484 G485 ARG LYS TYR ALA GLN LEU GLU HIS HIS HIS HIS HIS HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 61	Depositor
Cell constants a, b, c, α, β, γ	74.39Å 74.39Å 441.76Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	48.48 – 2.90 48.48 – 2.90	Depositor EDS
% Data completeness (in resolution range)	96.0 (48.48-2.90) 96.0 (48.48-2.90)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.33 (at 2.91Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, R_{free}	0.218 , 0.281 0.214 , 0.281	Depositor DCC
R_{free} test set	2909 reflections (11.05%)	DCC
Wilson B-factor (Å^2)	91.7	Xtriage
Anisotropy	0.033	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 56.6	EDS
Estimated twinning fraction	0.109 for h,-h-k,-l	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.44$, $< L^2 > = 0.27$	Xtriage
Outliers	0 of 29228 reflections	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	7620	wwPDB-VP
Average B, all atoms (Å^2)	110.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.77% of the height of the origin peak. No significant pseudotranslation is detected.*

[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: MG, ADP

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	0.45	0/3808	0.62	0/5124
1	B	0.44	0/3846	0.60	0/5175
All	All	0.45	0/7654	0.61	0/10299

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

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	204	ASP	Peptide
1	B	338	GLY	Peptide

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	3755	0	3790	233	1
1	B	3792	0	3815	290	1
2	A	27	0	12	5	0
2	B	27	0	12	1	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	10	0	0	4	0
4	B	7	0	0	1	0
All	All	7620	0	7629	520	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 34.

The worst 5 of 520 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:400:THR:HG22	1:B:429:PHE:O	1.45	1.15
1:B:11:LYS:HA	1:B:14:ILE:HD12	1.22	1.13
1:A:44:ARG:HH11	1:A:44:ARG:HG3	0.93	1.08
1:A:54:ILE:H	1:A:54:ILE:HD12	0.95	1.08
1:B:135:LEU:CD1	1:B:140:LYS:HG3	1.85	1.06

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:A:186:GLN:NE2	1:B:211:THR:OG1[5_664]	2.17	0.03

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	474/498 (95%)	395 (83%)	60 (13%)	19 (4%)	4 15
1	B	479/498 (96%)	377 (79%)	75 (16%)	27 (6%)	2 7
All	All	953/996 (96%)	772 (81%)	135 (14%)	46 (5%)	3 10

5 of 46 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	23	LEU
1	A	147	TYR
1	A	404	ARG
1	B	75	ASP
1	B	149	ASP

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	407/416 (98%)	363 (89%)	44 (11%)	8 24
1	B	411/416 (99%)	359 (87%)	52 (13%)	5 16
All	All	818/832 (98%)	722 (88%)	96 (12%)	7 19

5 of 96 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	471	ASP
1	B	74	LEU
1	B	399	THR
1	B	5	THR
1	B	36	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	112	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	123	HIS
1	B	305	HIS
1	A	470	GLN
1	B	296	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

Of 4 ligands modelled in this entry, 2 are monoatomic - leaving 2 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	$\#\|Z\| > 2$
2	ADP	A	1486	3	22,29,29	1.05	2 (9%)	27,45,45	2.20	6 (22%)
2	ADP	B	1486	3	22,29,29	1.22	2 (9%)	27,45,45	2.29	10 (37%)

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	ADP	A	1486	3	-	0/12/32/32	0/3/3/3
2	ADP	B	1486	3	-	0/12/32/32	0/3/3/3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1486	ADP	O4'-C1'	2.46	1.44	1.41
2	B	1486	ADP	O4'-C1'	2.88	1.44	1.41
2	A	1486	ADP	C5-C4	2.91	1.47	1.40
2	B	1486	ADP	C5-C4	3.53	1.48	1.40

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	1486	ADP	N3-C2-N1	-7.18	123.40	128.89
2	B	1486	ADP	N3-C2-N1	-7.18	123.40	128.89
2	A	1486	ADP	PA-O3A-PB	-5.23	115.14	132.67
2	B	1486	ADP	PA-O3A-PB	-4.33	118.15	132.67
2	B	1486	ADP	C2'-C1'-N9	-3.81	108.47	114.29

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1486	ADP	5	0
2	B	1486	ADP	1	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	469/498 (94%)	0.08	9 (1%) 70 66	67, 103, 151, 210	1 (0%)
1	B	474/498 (95%)	0.15	20 (4%) 40 33	72, 109, 151, 197	1 (0%)
All	All	943/996 (94%)	0.11	29 (3%) 52 45	67, 105, 151, 210	2 (0%)

The worst 5 of 29 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	151	ILE	9.1
1	A	152	THR	9.0
1	B	207	ASP	6.7
1	A	207	ASP	5.5
1	A	153	THR	5.2

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	ADP	A	1486	27/27	0.96	0.25	1.85	92,116,141,168	0
2	ADP	B	1486	27/27	0.90	0.27	1.14	92,128,155,157	0
3	MG	B	1487	1/1	0.90	0.35	-	75,75,75,75	0
3	MG	A	1487	1/1	0.98	0.39	-	70,70,70,70	0

6.5 Other polymers (i)

There are no such residues in this entry.