



wwPDB X-ray Structure Validation Summary Report

Mar 31, 2014 – 03:48 PM BST

PDB ID : 4BEB
Title : MUTANT (K220E) OF THE HSDR SUBUNIT OF THE ECOR124I RESTRICTION ENZYME IN COMPLEX WITH ATP
Authors : Csefalvay, E.; Lapkouski, M.; Guzanova, A.; Csefalvay, L.; Baikova, T.; Shevelev, I.; Janscak, P.; Smatanova, I.K.; Panjikar, S.; Carey, J.; Weiserova, M.; Ettrich, R.
Deposited on : 2013-03-07
Resolution : 2.99 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

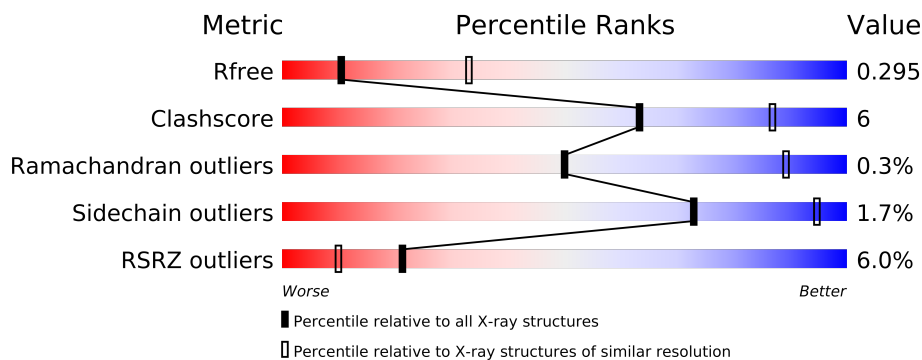
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.15 2013
Xtriage (Phenix) : dev-1323
EDS : stable23004
Percentile statistics : 21963
Refmac : 5.8.0049
CCP4 : 6.3.0 (Settle)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : stable23004

1 Overall quality at a glance





The reported resolution of this entry is 2.99 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}	66092	1468 (3.00-2.96)
Clashscore	79885	1894 (3.00-2.96)
Ramachandran outliers	78287	1826 (3.00-2.96)
Sidechain outliers	78261	1829 (3.00-2.96)
RSRZ outliers	66119	1469 (3.00-2.96)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	1038	
1	B	1038	
1	C	1038	
1	D	1038	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 27369 atoms, of which 0 are hydrogen and 0 are deuterium.

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 TYPE I RESTRICTION ENZYME HSDR.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	839	Total	C	N	O	S	0	0	0
			6821	4340	1158	1307	16			
1	B	842	Total	C	N	O	S	0	0	0
			6846	4353	1158	1319	16			
1	C	838	Total	C	N	O	S	0	0	0
			6802	4326	1151	1309	16			
1	D	833	Total	C	N	O	S	0	0	0
			6772	4309	1146	1301	16			

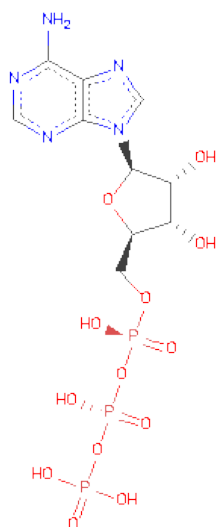
There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	220	GLU	LYS	ENGINEERED MUTATION	UNP Q304R3
B	220	GLU	LYS	ENGINEERED MUTATION	UNP Q304R3
C	220	GLU	LYS	ENGINEERED MUTATION	UNP Q304R3
D	220	GLU	LYS	ENGINEERED MUTATION	UNP Q304R3

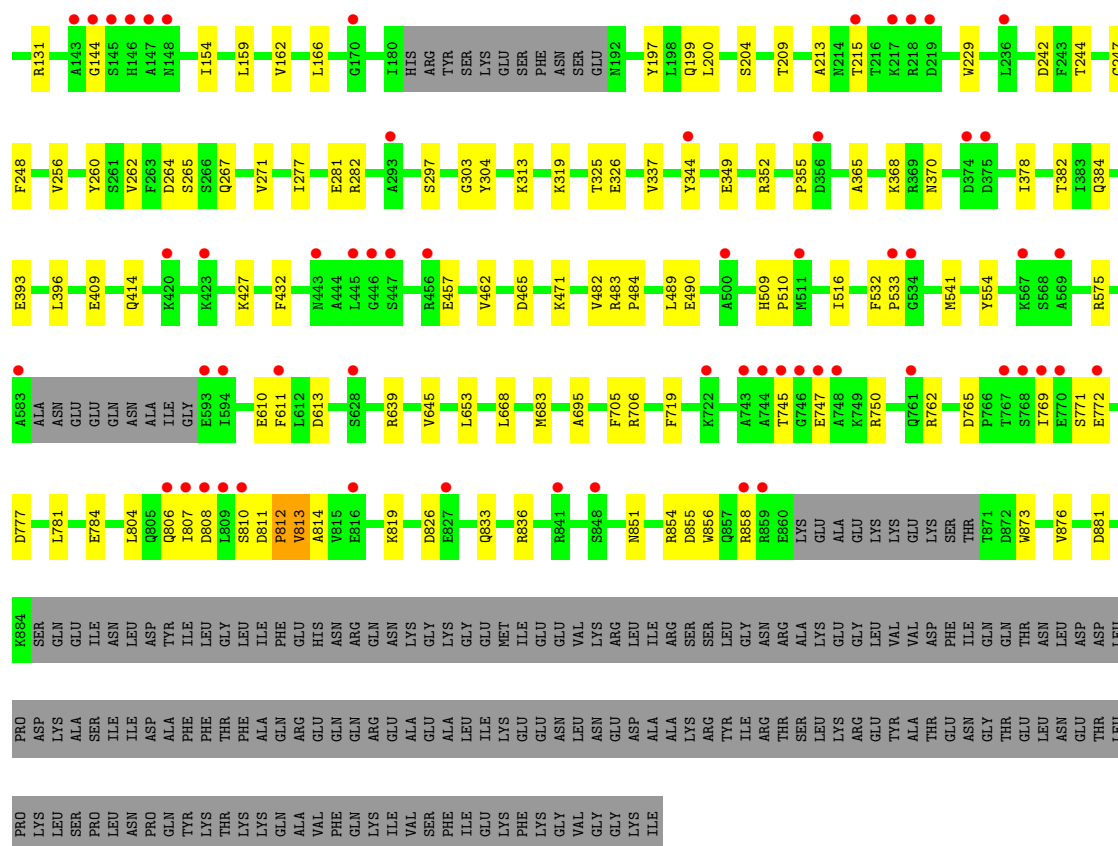
- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Mg	0	0
			1	1		
2	A	1	Total	Mg	0	0
			1	1		
2	D	1	Total	Mg	0	0
			1	1		
2	C	1	Total	Mg	0	0
			1	1		

- Molecule 3 is ADENOSINE-5'-TRIPHOSPHATE (three-letter code: ATP) (formula: C₁₀H₁₆N₅O₁₃P₃).

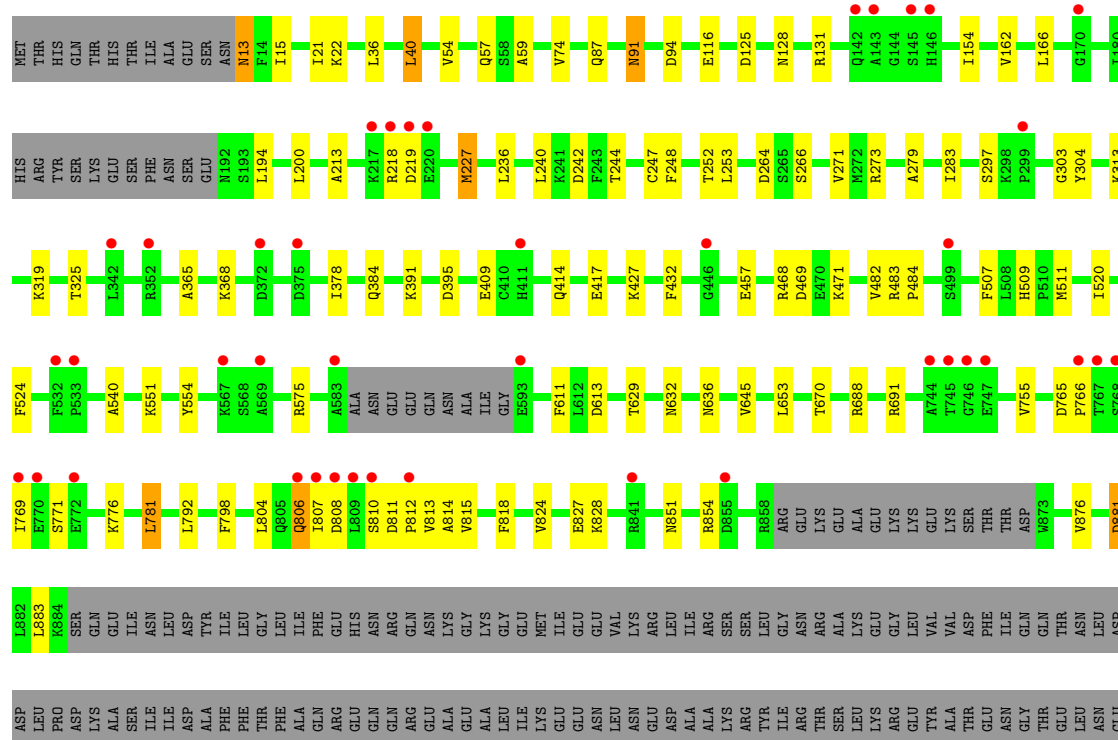


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
3	B	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
3	C	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
3	D	1	Total	C	N	O	P	0	0
			31	10	5	13	3		



• Molecule 1: TYPE I RESTRICTION ENZYME HSDR

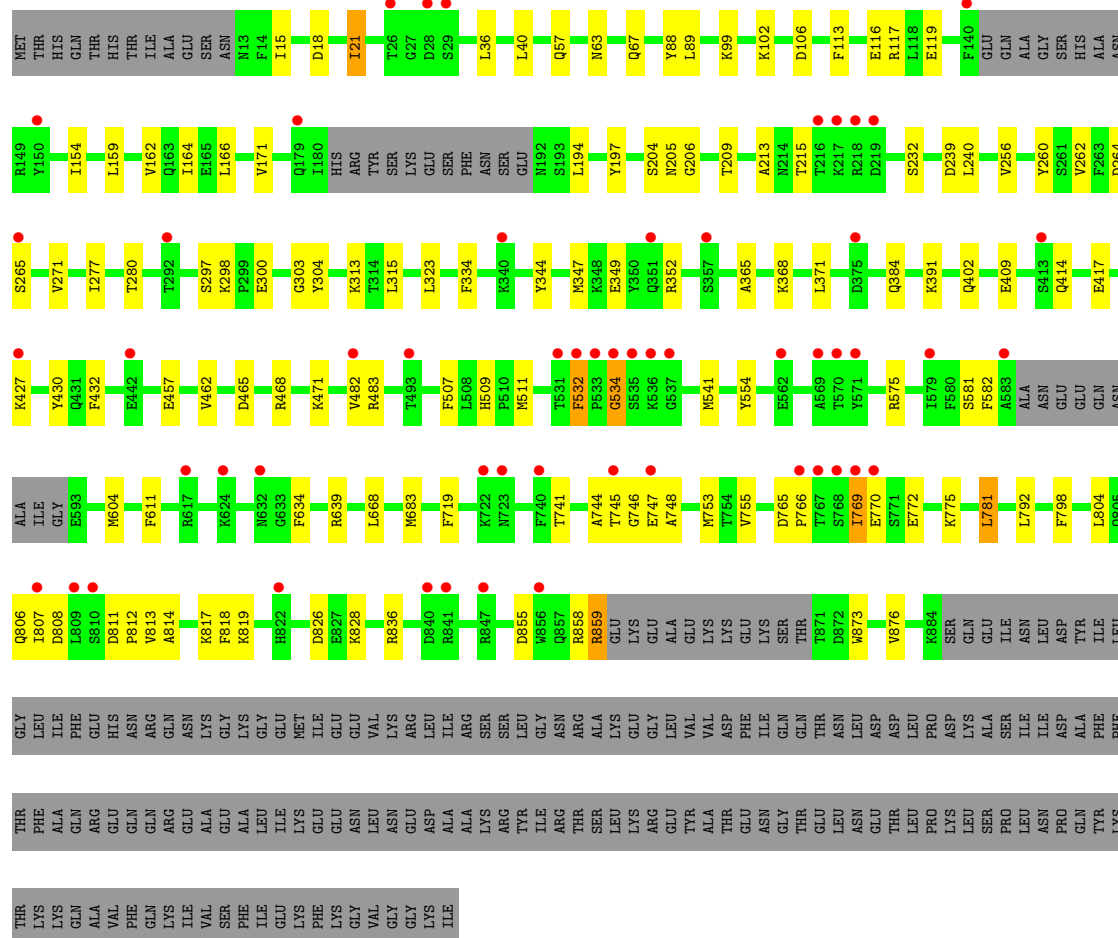
Chain C:



THR
LEU
PRO
GLN
LEU
SER
PRO
LEU
ASN
PRO
GLN
TYR
LYS
THR
LYS
LYS
GLN
ALA
VAL
PHE
GLN
LYS
ILE
ILE
VAL
SER
PHE
ILE
GLU
LYS
PHE
LYS
GLY
VAL
GLY
GLY
LYS
ILE

• Molecule 1: TYPE I RESTRICTION ENZYME HSDR

Chain D: 



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	127.11Å 123.11Å 160.11Å 90.00° 111.48° 90.00°	Depositor
Resolution (Å)	19.89 – 2.99 19.88 – 2.99	Depositor EDS
% Data completeness (in resolution range)	98.3 (19.89-2.99) 98.4 (19.88-2.99)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.38 (at 2.98Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, R_{free}	0.255 , 0.296 0.255 , 0.295	Depositor DCC
R_{free} test set	966 reflections (1.06%)	DCC
Wilson B-factor (Å ²)	41.1	Xtriage
Anisotropy	1.042	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 22.6	EDS
Estimated twinning fraction	0.025 for h,-k,-h-l	Xtriage
L-test for twinning	$\langle L \rangle = 0.45$, $\langle L^2 \rangle = 0.28$	Xtriage
Outliers	4 of 91281 reflections (0.004%)	Xtriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	27369	wwPDB-VP
Average B, all atoms (Å ²)	54.0	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 39.36 % 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 3.2132e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹Intensities estimated from amplitudes.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: MG, 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	0.27	0/6954	0.51	1/9383 (0.0%)
1	B	0.25	0/6979	0.50	1/9418 (0.0%)
1	C	0.27	0/6933	0.49	0/9355
1	D	0.25	0/6901	0.48	1/9309 (0.0%)
All	All	0.26	0/27767	0.50	3/37465 (0.0%)

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	A	0	4
1	B	0	4
1	C	0	2
1	D	0	3
All	All	0	13

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	813	VAL	N-CA-C	7.00	129.89	111.00
1	B	813	VAL	N-CA-C	6.66	128.98	111.00
1	D	534	GLY	N-CA-C	5.49	126.81	113.10

There are no chirality outliers.

5 of 13 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	147	ALA	Peptide
1	A	769	ILE	Peptide
1	A	806	GLN	Peptide
1	A	812	PRO	Peptide
1	B	144	GLY	Peptide

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	6821	0	6660	83	0
1	B	6846	0	6672	78	0
1	C	6802	0	6636	67	0
1	D	6772	0	6607	75	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	A	31	0	12	0	0
3	B	31	0	12	0	0
3	C	31	0	12	1	0
3	D	31	0	12	0	0
All	All	27369	0	26623	297	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 6.

The worst 5 of 297 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:147:ALA:HB1	1:A:148:ASN:HA	1.49	0.93
1:A:807:ILE:H	1:A:808:ASP:HB2	1.39	0.86
1:B:807:ILE:H	1:B:808:ASP:HB2	1.44	0.81
1:C:807:ILE:H	1:C:808:ASP:HB2	1.46	0.80
1:D:859:ARG:H	1:D:859:ARG:HH11	1.27	0.79

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone ⓘ

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	831/1038 (80%)	793 (95%)	36 (4%)	2 (0%)	56	91
1	B	834/1038 (80%)	800 (96%)	32 (4%)	2 (0%)	56	91
1	C	830/1038 (80%)	796 (96%)	33 (4%)	1 (0%)	59	93
1	D	821/1038 (79%)	782 (95%)	35 (4%)	4 (0%)	38	83
All	All	3316/4152 (80%)	3171 (96%)	136 (4%)	9 (0%)	50	89

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	746	GLY
1	D	813	VAL
1	A	765	ASP
1	C	813	VAL
1	B	813	VAL

5.3.2 Protein sidechains ⓘ

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	737/927 (80%)	724 (98%)	13 (2%)	71	94
1	B	741/927 (80%)	727 (98%)	14 (2%)	69	94
1	C	736/927 (79%)	722 (98%)	14 (2%)	69	94
1	D	733/927 (79%)	725 (99%)	8 (1%)	84	97
All	All	2947/3708 (80%)	2898 (98%)	49 (2%)	73	95

5 of 49 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	610	GLU
1	C	13	ASN
1	D	753	MET
1	B	771	SER
1	C	21	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	57	GLN
1	D	57	GLN

5.3.3 RNA ⓘ

There are no RNA chains 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 8 ligands modelled in this entry, 4 are monoatomic - leaving 4 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	$\# Z > 2$
3	ATP	A	1886	2	33,33,33	1.04	2 (6%)	52,52,52	1.65	7 (13%)
3	ATP	B	1886	2	33,33,33	1.07	2 (6%)	52,52,52	1.66	9 (17%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
3	ATP	C	1886	2	33,33,33	1.01	2 (6%)	52,52,52	1.72	10 (19%)
3	ATP	D	1886	2	33,33,33	1.04	2 (6%)	52,52,52	1.69	10 (19%)

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
3	ATP	A	1886	2	-	0/22/38/38	0/1/3/3
3	ATP	B	1886	2	-	0/22/38/38	0/1/3/3
3	ATP	C	1886	2	-	0/22/38/38	0/1/3/3
3	ATP	D	1886	2	-	0/22/38/38	0/1/3/3

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	1886	ATP	C5-C4	3.17	1.47	1.40
3	D	1886	ATP	C5-C4	3.12	1.47	1.40
3	A	1886	ATP	C5-C4	3.04	1.47	1.40
3	C	1886	ATP	C5-C4	3.00	1.47	1.40
3	C	1886	ATP	C4-N9	-2.83	1.33	1.37

The worst 5 of 36 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1886	ATP	N3-C2-N1	-6.05	123.65	128.71
3	B	1886	ATP	N3-C2-N1	-5.88	123.79	128.71
3	D	1886	ATP	N3-C2-N1	-5.84	123.83	128.71
3	C	1886	ATP	N3-C2-N1	-5.83	123.83	128.71
3	B	1886	ATP	N3-C4-N9	5.33	135.05	125.43

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues

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, 95th 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	839/1038 (80%)	0.01	43 (5%) 27 12	20, 42, 79, 109	0
1	B	842/1038 (81%)	0.28	63 (7%) 14 7	38, 58, 86, 115	0
1	C	838/1038 (80%)	0.01	41 (4%) 28 13	21, 42, 79, 105	0
1	D	833/1038 (80%)	0.37	55 (6%) 18 9	41, 62, 91, 117	0
All	All	3352/4152 (80%)	0.17	202 (6%) 21 10	20, 54, 85, 117	0

The worst 5 of 202 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	767	THR	10.8
1	A	767	THR	8.6
1	C	145	SER	7.2
1	D	583	ALA	6.6
1	C	767	THR	6.4

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, 95th 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	RSR	LLDF	B-factors(\AA^2)	Q<0.9
2	MG	B	1885	1/1	0.21	1.17	52,52,52,52	0
2	MG	A	1885	1/1	0.19	1.03	45,45,45,45	0
2	MG	D	1885	1/1	0.27	0.85	67,67,67,67	0
2	MG	C	1885	1/1	0.15	-0.26	37,37,37,37	0
3	ATP	C	1886	31/31	0.12	-1.22	21,29,39,40	0
3	ATP	D	1886	31/31	0.13	-1.23	45,51,60,62	0
3	ATP	A	1886	31/31	0.12	-1.32	27,34,43,51	0
3	ATP	B	1886	31/31	0.12	-1.39	42,50,57,61	0

6.5 Other polymers ⓘ

There are no such residues in this entry.