



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 19, 2016 – 10:52 PM GMT

PDB ID : 5EAW
Title : Crystal structure of Dna2 nuclease-helicase
Authors : Zhou, C.; Pourmal, S.; Pavletich, N.P.
Deposited on : 2015-10-17
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 ⓘ symbol.

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

MolProbity : 4.02b-467
Mogul : 1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026982
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) : rb-20026982

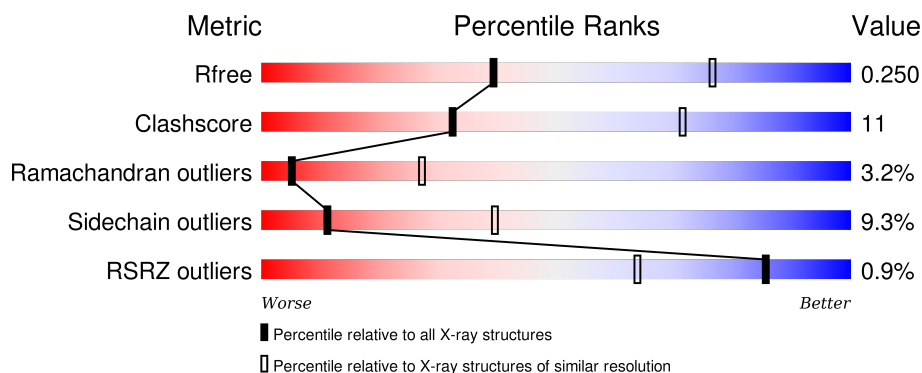
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

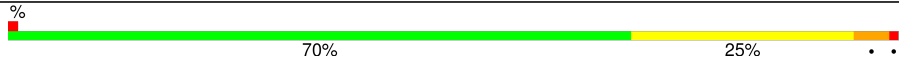
The reported resolution of this entry is 3.00 Å.

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 $\leq 5\%$. 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	1056	 70% 24% . .
1	B	1056	 70% 25% . .

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
3	ADP	B	1102	-	-	-	X

2 Entry composition [i](#)

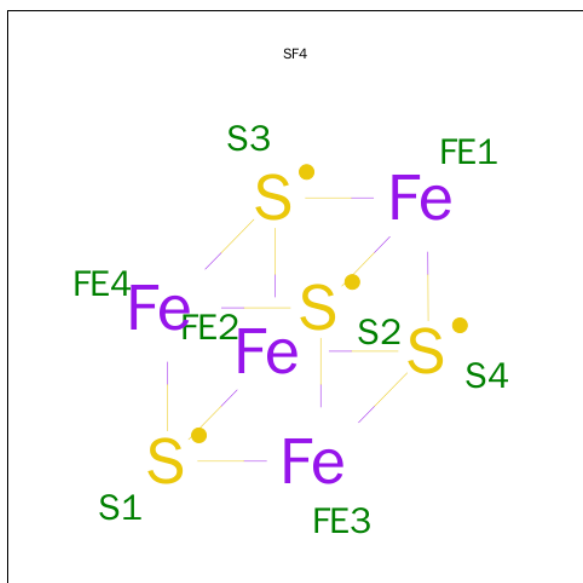
There are 3 unique types of molecules in this entry. The entry contains 16612 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 DNA replication ATP-dependent helicase/nuclease DNA2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1046	Total	C	N	O	S	0	0	0
			8271	5222	1456	1548	45			
1	B	1046	Total	C	N	O	S	0	0	0
			8271	5222	1456	1548	45			

- Molecule 2 is IRON/SULFUR CLUSTER (three-letter code: SF4) (formula: Fe₄S₄).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	Fe	S	0	0
			8	4	4		
2	B	1	Total	Fe	S	0	0
			8	4	4		

- Molecule 3 is ADENOSINE-5'-DIPHOSPHATE (three-letter code: ADP) (formula: C₁₀H₁₅N₅O₁₀P₂).

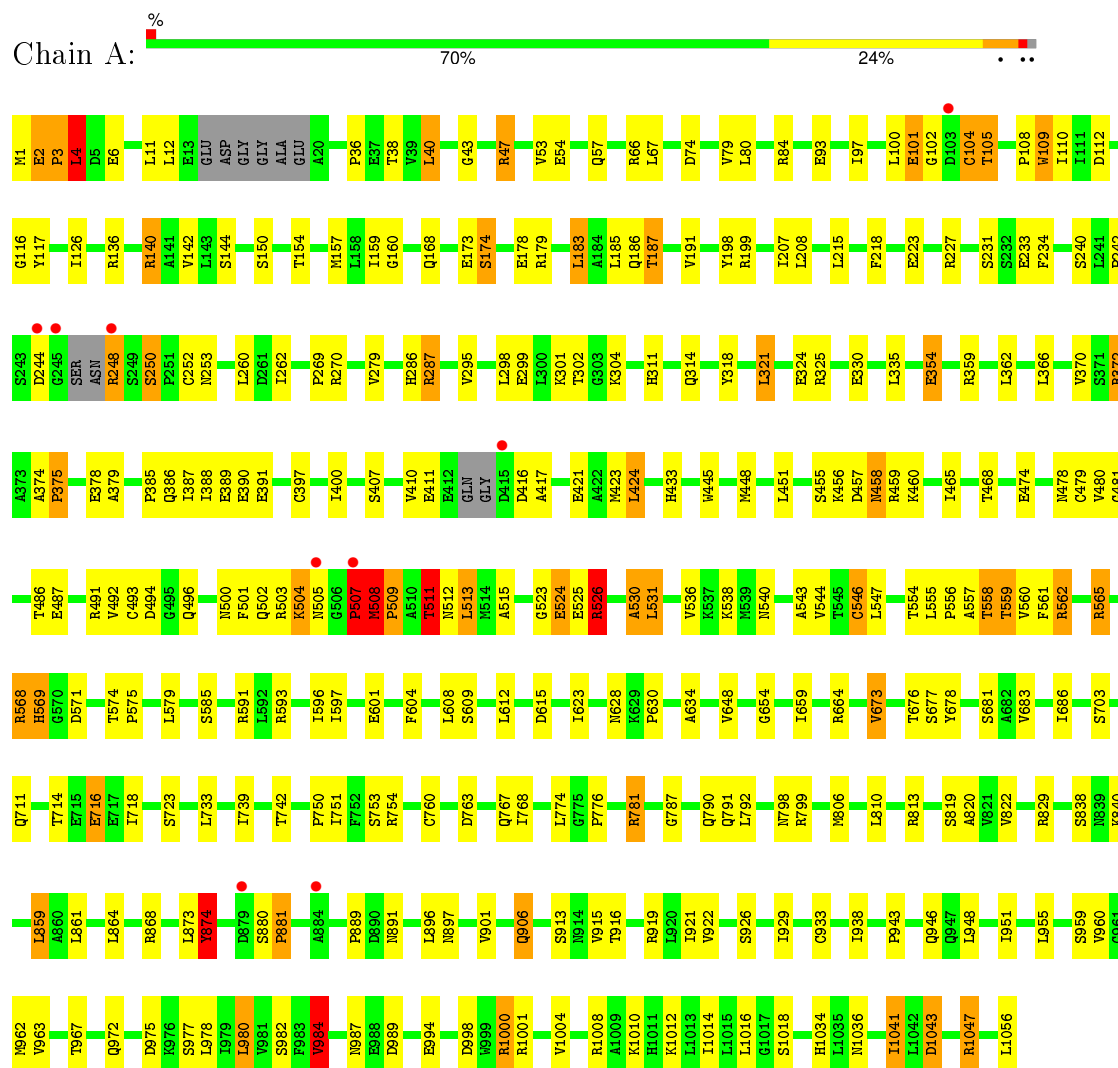


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total 27	C 10	N 5	O 10	P 2	0	0
3	B	1	Total 27	C 10	N 5	O 10	P 2	0	0

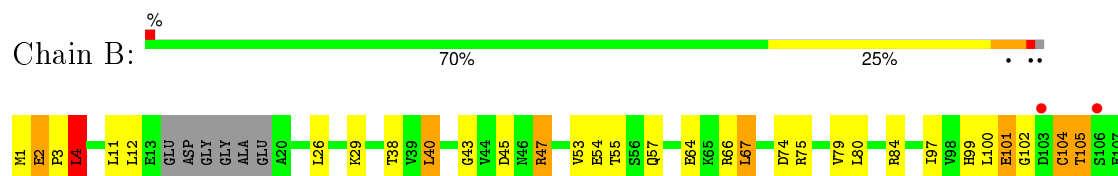
3 Residue-property plots

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 ($\text{RSRZ} > 2$). 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: DNA replication ATP-dependent helicase/nuclease DNA2



- Molecule 1: DNA replication ATP-dependent helicase/nuclease DNA2



L1056	M962	V963	G845	S703	V560	C479	S371	E233	P108
	P963				F561	G480	R372	F234	W109
	T967	L859	L859	T714	L563	T466	P375	S240	I111
	V968	A860	A860	E715	D564	E487	E378	L241	D112
	D969	L861	L861	E716	R665		A379	P242	G116
	Q972		R868	E717	E566	R491		S243	Y117
				I718	R567	V492		D244	
	S977	L871	L871	S723	R568	C493	I387	G245	I126
	L978	Q872	Q872	L733	H569	C493	I388	SER	
	I979	L873	L873	L733	G570	D494	E389	ASN	R136
	L980	Y874	Y874		D571	G496	E390	R248	
	V981			I739	I572	Q496	E391	S249	
	S982	Y877	Y877		L579	N500	C397	P251	R140
	F983	S878	S878	T742	S585	F501	I400	C252	A141
	V884	D879	D879			Q502		N253	V142
	N987	S880	S880	T747	R591	R503	L405	L260	S150
		P881	P881		R592	K504	Y406	L260	T154
	E988	P889	P889	F750	R593	N505	S407	I262	
	D989	D890	D890	F752	P507	P507	V410	P269	M157
	G990	N891	N891	S763	I596	X508	E411	R270	L158
	T991				I597	P509	E412		I159
	E994	L896	L896	F759	E601	A510	GLN	V279	G160
	L995	N897	N897	D763	F604	N512	GLY	H286	Q168
	L996	V901	V901			L513	D415	R287	E173
	D998	Q906	Q906	Q767	L608	M514	A417		S174
	W999			I768		A515		V295	
	R1000	V915	V915	R781	L612	I520	E421	L298	E178
	R1001	T916	T916	R782	D615	L521	A422	E299	R179
	V1004	R919	R919	V766	I623	S522	M423	E300	
	R1008	L920	L920			G523	L424	I300	L183
	A1009	I921	I921	Q790	R623	E524		K301	A184
	K1010	V922	V922	Q791	A634	E525	H433	T302	L185
	H1011			L792		R526		G303	Q186
	K1012	S926	S926	P793	V638		M445	K304	T187
	L1013					A530			
	I1014	I929	I929	N798	I647	L531	M448	H311	V191
	S1018	Q933	Q933	R799	G654	V536	S453	L321	R192
	L1028	I938	I938	M806	I659	K538	Q454	E324	H193
	H1034	P943	P943	L810	L662	M539	K456	R325	R199
	L1035	Q946	Q946	R813	V663	N540	D457	E330	L208
	A1037	Q947	Q947	S819	R664	A543	R459	E300	C209
	E1038	L948	L948	A820		V544	K460	L335	E210
	Q1039			W821	V673	T545	I465	E354	L215
	L1040	I951	I951	V822	T676	L547	T468	R359	F218
	I1041	L955	L955	N831	S677				E223
	D1043				Y678	T554	S471	A363	
	R1047	S959	S959	S838	S681	P556	E474	L366	R227
		V960	V960	N839	I866	A557			S231
		G961	G961	T940		T558	N478	V270	S232

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	120.90 Å 148.60 Å 170.50 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 3.00 49.31 – 2.98	Depositor EDS
% Data completeness (in resolution range)	89.2 (50.00-3.00) 88.4 (49.31-2.98)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	0.96 (at 2.96 Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.209 , 0.247 0.215 , 0.250	Depositor DCC
R_{free} test set	1545 reflections (2.86%)	DCC
Wilson B-factor (Å ²)	48.3	Xtriage
Anisotropy	0.455	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 33.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	1 of 61724 reflections (0.002%)	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	16612	wwPDB-VP
Average B, all atoms (Å ²)	63.0	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 34.13 % 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 7.1691e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹ Intensities estimated from amplitudes.

² Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: SF4, 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	0.57	0/8421	0.83	5/11387 (0.0%)
1	B	0.56	0/8421	0.83	4/11387 (0.0%)
All	All	0.56	0/16842	0.83	9/22774 (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	3
1	B	0	3
All	All	0	6

There are no bond length outliers.

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	A	4	LEU	CA-CB-CG	7.61	132.81	115.30
1	B	4	LEU	CA-CB-CG	7.15	131.75	115.30
1	A	12	LEU	CA-CB-CG	5.94	128.97	115.30
1	B	12	LEU	CA-CB-CG	5.91	128.90	115.30
1	A	774	LEU	CA-CB-CG	5.34	127.58	115.30

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	104	CYS	Peptide

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	A	507	PRO	Peptide
1	A	508	MET	Peptide
1	B	104	CYS	Peptide
1	B	507	PRO	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	8271	0	8411	177	0
1	B	8271	0	8411	177	0
2	A	8	0	0	0	0
2	B	8	0	0	1	0
3	A	27	0	12	1	0
3	B	27	0	12	1	0
All	All	16612	0	16846	352	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 11.

The worst 5 of 352 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:458:ASN:HB2	1:B:569:HIS:HB3	1.53	0.90
1:A:511:THR:H	1:A:513:LEU:HD21	1.37	0.89
1:A:43:GLY:H	1:A:105:THR:HG21	1.41	0.86
1:B:511:THR:H	1:B:513:LEU:HD21	1.38	0.85
1:B:183:LEU:O	1:B:187:THR:HG23	1.77	0.85

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	1038/1056 (98%)	916 (88%)	90 (9%)	32 (3%)	5	28
1	B	1038/1056 (98%)	919 (88%)	85 (8%)	34 (3%)	5	26
All	All	2076/2112 (98%)	1835 (88%)	175 (8%)	66 (3%)	5	27

5 of 66 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	2	GLU
1	A	105	THR
1	A	375	PRO
1	A	508	MET
1	A	524	GLU

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	931/937 (99%)	845 (91%)	86 (9%)	11	40
1	B	931/937 (99%)	844 (91%)	87 (9%)	11	39
All	All	1862/1874 (99%)	1689 (91%)	173 (9%)	11	39

5 of 173 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	980	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	183	LEU
1	B	962	MET
1	A	987	ASN
1	B	47	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	946	GLN
1	A	972	GLN
1	B	496	GLN
1	A	770	GLN
1	B	505	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](#)

4 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	SF4	A	1101	1	0,12,12	0.00	-	0,24,24	0.00	-
3	ADP	A	1102	-	24,29,29	1.18	2 (8%)	23,45,45	1.85	1 (4%)
2	SF4	B	1101	1	0,12,12	0.00	-	0,24,24	0.00	-
3	ADP	B	1102	-	24,29,29	1.16	2 (8%)	23,45,45	1.76	1 (4%)

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	SF4	A	1101	1	-	0/0/48/48	0/6/5/5
3	ADP	A	1102	-	-	0/12/32/32	0/3/3/3
2	SF4	B	1101	1	-	0/0/48/48	0/6/5/5
3	ADP	B	1102	-	-	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)
3	A	1102	ADP	O4'-C1'	2.06	1.44	1.41
3	B	1102	ADP	C2-N3	2.25	1.36	1.32
3	A	1102	ADP	C5-C4	3.57	1.48	1.40
3	B	1102	ADP	C5-C4	3.60	1.48	1.40

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1102	ADP	N3-C2-N1	-7.15	123.25	128.87
3	B	1102	ADP	N3-C2-N1	-6.52	123.75	128.87

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1102	ADP	1	0
2	B	1101	SF4	1	0
3	B	1102	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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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	1046/1056 (99%)	-0.38	9 (0%) 85 64	24, 53, 119, 169	0
1	B	1046/1056 (99%)	-0.34	10 (0%) 84 60	23, 56, 126, 193	0
All	All	2092/2112 (99%)	-0.36	19 (0%) 85 64	23, 55, 123, 193	0

The worst 5 of 19 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	415	ASP	4.8
1	B	103	ASP	4.2
1	B	106	SER	3.6
1	B	478	ASN	3.3
1	B	877	TYR	2.9

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [i](#)

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	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
3	ADP	B	1102	27/27	0.86	0.20	2.12	62,99,131,140	0
3	ADP	A	1102	27/27	0.85	0.19	0.51	63,86,136,141	0
2	SF4	A	1101	8/8	0.99	0.10	-1.15	51,58,63,64	0
2	SF4	B	1101	8/8	0.99	0.11	-1.41	57,63,71,72	0

6.5 Other polymers [i](#)

There are no such residues in this entry.