



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 06:45 AM GMT

PDB ID : 2Y65
Title : Crystal structure of Drosophila melanogaster kinesin-1 motor domain dimer-tail complex
Authors : Kaan, H.Y.K.; Hackney, D.D.; Kozielski, F.
Deposited on : 2011-01-19
Resolution : 2.20 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 (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

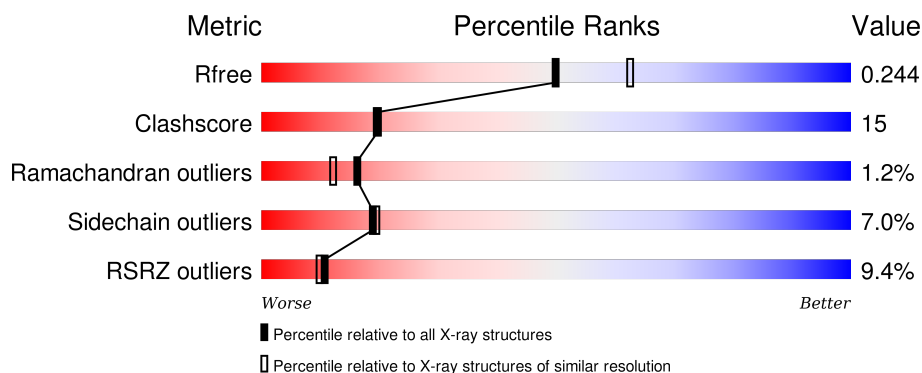
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.20 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	3774 (2.20-2.20)
Clashscore	102246	4477 (2.20-2.20)
Ramachandran outliers	100387	4404 (2.20-2.20)
Sidechain outliers	100360	4405 (2.20-2.20)
RSRZ outliers	91569	3781 (2.20-2.20)

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	365	<div> <div>6%</div> <div> <div></div> <div>67%</div> <div>19%</div> <div>•</div> <div>10%</div> </div> </div>
1	B	365	<div> <div>7%</div> <div> <div></div> <div>65%</div> <div>25%</div> <div>•</div> <div>7%</div> </div> </div>
1	C	365	<div> <div>10%</div> <div> <div></div> <div>73%</div> <div>18%</div> <div>•</div> <div>6%</div> </div> </div>
1	D	365	<div> <div>11%</div> <div> <div></div> <div>69%</div> <div>21%</div> <div>•</div> <div>7%</div> </div> </div>
2	W	20	<div> <div>10%</div> <div> <div></div> <div>10%</div> <div>35%</div> <div>20%</div> <div>35%</div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	X	20	<div><div></div><div></div><div></div><div></div><div></div></div> <div>15%40%20%5%35%</div>
2	Y	20	<div><div></div><div></div><div></div><div></div><div></div></div> <div>10%45%5%10%40%</div>

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 11750 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 KINESIN HEAVY CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	328	Total	C	N	O	S	0	10	0
			2598	1640	450	499	9			
1	B	341	Total	C	N	O	S	0	6	0
			2681	1692	460	519	10			
1	C	343	Total	C	N	O	S	0	3	0
			2694	1693	471	521	9			
1	D	338	Total	C	N	O	S	0	6	0
			2654	1672	457	516	9			

- Molecule 2 is a protein called KINESIN HEAVY CHAIN.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	W	13	Total	C	N	O	0	13	0
			174	108	36	30			
2	X	13	Total	C	N	O	0	0	0
			93	58	19	16			
2	Y	12	Total	C	N	O	0	0	0
			88	55	18	15			

There are 12 discrepancies between the modelled and reference sequences:

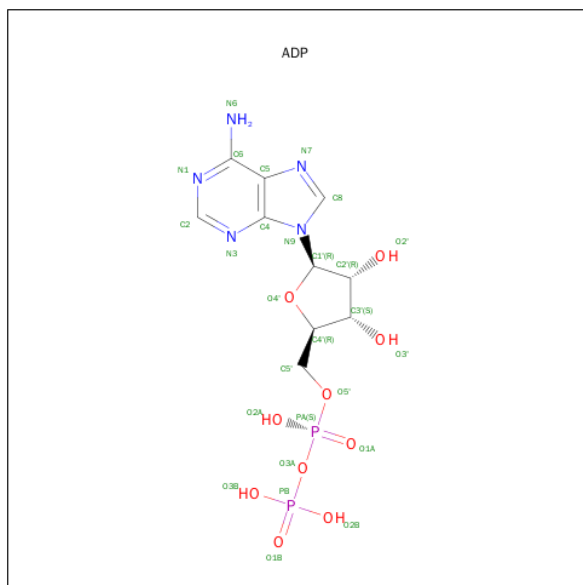
Chain	Residue	Modelled	Actual	Comment	Reference
W	935	GLY	-	EXPRESSION TAG	UNP P17210
W	936	SER	-	EXPRESSION TAG	UNP P17210
W	953	THR	-	EXPRESSION TAG	UNP P17210
W	954	SER	-	EXPRESSION TAG	UNP P17210
X	935	GLY	-	EXPRESSION TAG	UNP P17210
X	936	SER	-	EXPRESSION TAG	UNP P17210
X	953	THR	-	EXPRESSION TAG	UNP P17210
X	954	SER	-	EXPRESSION TAG	UNP P17210
Y	935	GLY	-	EXPRESSION TAG	UNP P17210
Y	936	SER	-	EXPRESSION TAG	UNP P17210

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
Y	953	THR	-	EXPRESSION TAG	UNP P17210
Y	954	SER	-	EXPRESSION TAG	UNP P17210

- Molecule 3 is ADENOSINE-5'-DIPHOSPHATE (three-letter code: ADP) (formula: $C_{10}H_{15}N_5O_{10}P_2$).



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

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

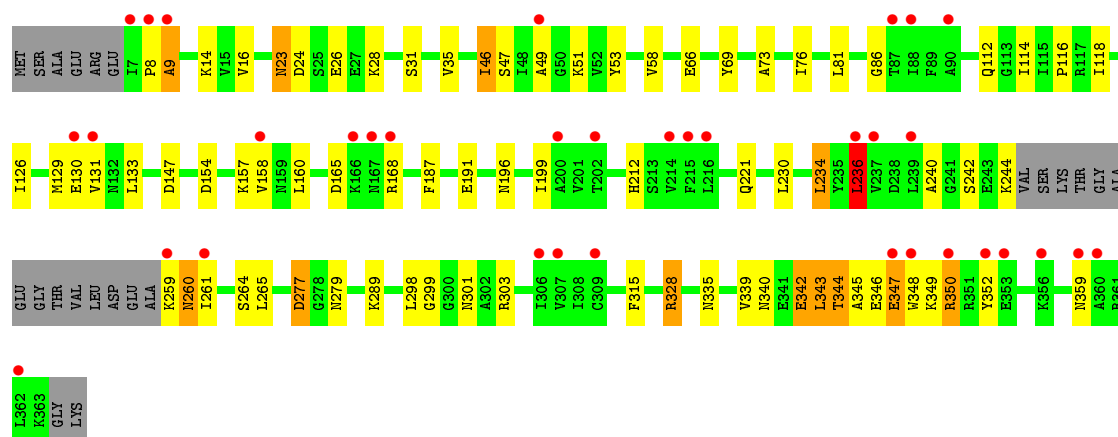
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total	Mg	0	0
			1	1		
4	A	1	Total	Mg	0	0
			1	1		
4	D	1	Total	Mg	0	0
			1	1		
4	C	1	Total	Mg	0	0
			1	1		

- Molecule 5 is water.

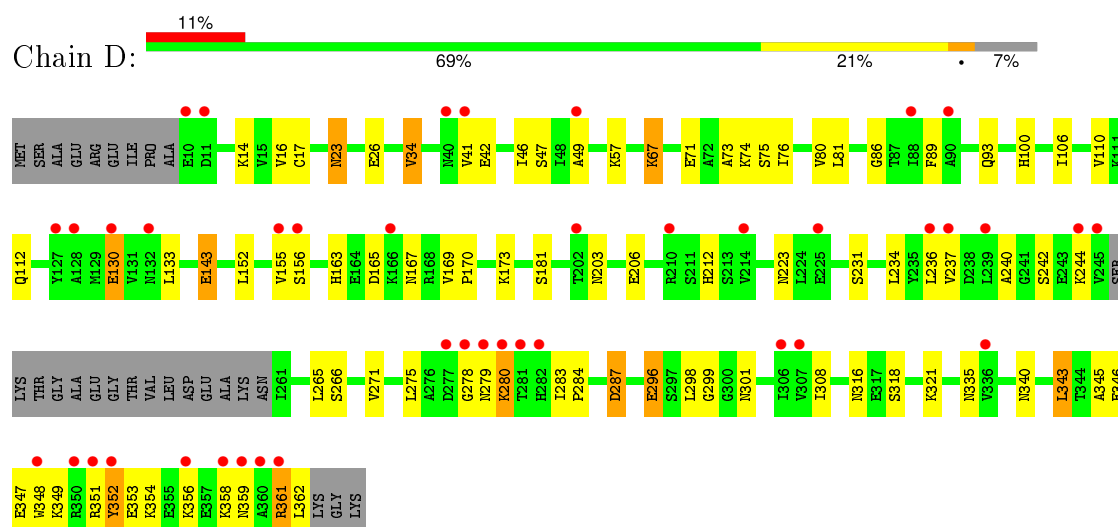
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	196	Total 196	O 196	0	0
5	B	175	Total 175	O 175	0	0
5	C	180	Total 180	O 180	0	0
5	D	96	Total 96	O 96	0	0
5	W	2	Total 2	O 2	0	0
5	X	6	Total 6	O 6	0	0
5	Y	1	Total 1	O 1	0	0

- Molecule 1: KINESIN HEAVY CHAIN

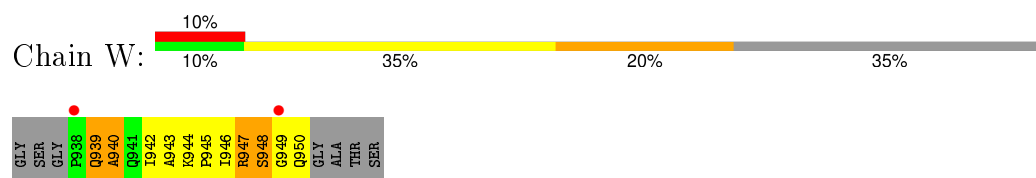




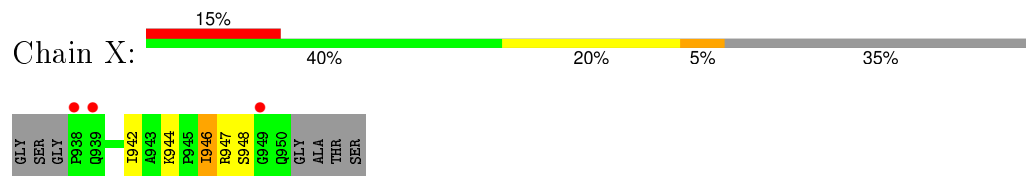
• Molecule 1: KINESIN HEAVY CHAIN



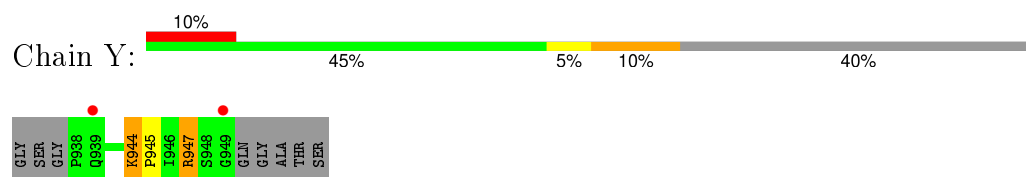
• Molecule 2: KINESIN HEAVY CHAIN



• Molecule 2: KINESIN HEAVY CHAIN



• Molecule 2: KINESIN HEAVY CHAIN



4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	113.77Å 190.71Å 145.98Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 2.20 29.74 – 2.20	Depositor EDS
% Data completeness (in resolution range)	99.8 (50.00-2.20) 99.9 (29.74-2.20)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.20 (at 2.20Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.195 , 0.252 0.188 , 0.244	Depositor DCC
R_{free} test set	4044 reflections (5.29%)	DCC
Wilson B-factor (Å ²)	40.8	Xtriage
Anisotropy	0.066	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 46.4	EDS
Estimated twinning fraction	0.000 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.011 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 80446 reflections	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	11750	wwPDB-VP
Average B, all atoms (Å ²)	42.0	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 22.29 % 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 5.8708e-03. 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

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	0.63	0/2669	0.90	1/3604 (0.0%)
1	B	0.57	0/2741	0.84	2/3699 (0.1%)
1	C	0.59	0/2745	0.89	5/3702 (0.1%)
1	D	0.48	0/2713	0.77	0/3661
2	W	0.34	0/174	0.76	0/229
2	X	0.37	0/94	0.60	0/125
2	Y	0.47	0/89	0.80	0/118
All	All	0.56	0/11225	0.85	8/15138 (0.1%)

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	D	0	1

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	328	ARG	NE-CZ-NH2	-6.42	117.09	120.30
1	C	236	LEU	CB-CG-CD1	6.36	121.81	111.00
1	A	16	VAL	CG1-CB-CG2	5.93	120.39	110.90
1	C	16	VAL	CG1-CB-CG2	5.90	120.34	110.90
1	B	133	LEU	CA-CB-CG	5.40	127.73	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	D	49	ALA	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	2598	0	2613	77	0
1	B	2681	0	2663	79	0
1	C	2694	0	2679	76	0
1	D	2654	0	2637	79	0
2	W	174	0	188	29	0
2	X	93	0	98	7	0
2	Y	88	0	96	4	0
3	A	27	0	12	0	0
3	B	27	0	12	0	0
3	C	27	0	12	0	0
3	D	27	0	12	0	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
4	C	1	0	0	0	0
4	D	1	0	0	0	0
5	A	196	0	0	9	0
5	B	175	0	0	8	0
5	C	180	0	0	6	0
5	D	96	0	0	9	0
5	W	2	0	0	0	0
5	X	6	0	0	4	0
5	Y	1	0	0	0	0
All	All	11750	0	11022	330	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 15.

The worst 5 of 330 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:168:ARG:HH11	1:A:168:ARG:HG3	1.04	1.18

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:W:943[A]:ALA:HB3	2:W:944[A]:LYS:CG	1.79	1.12
2:W:945[B]:PRO:C	2:W:946[B]:ILE:HD13	1.74	1.07
2:W:943[A]:ALA:HB1	2:W:944[A]:LYS:HA	1.30	1.07
2:W:943[A]:ALA:CB	2:W:944[A]:LYS:HG2	1.88	1.03

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	334/365 (92%)	320 (96%)	11 (3%)	3 (1%)	21	19
1	B	343/365 (94%)	327 (95%)	15 (4%)	1 (0%)	46	50
1	C	342/365 (94%)	329 (96%)	7 (2%)	6 (2%)	11	7
1	D	340/365 (93%)	316 (93%)	20 (6%)	4 (1%)	16	12
2	W	21/20 (105%)	11 (52%)	6 (29%)	4 (19%)	0	0
2	X	11/20 (55%)	9 (82%)	2 (18%)	0	100	100
2	Y	10/20 (50%)	9 (90%)	1 (10%)	0	100	100
All	All	1401/1520 (92%)	1321 (94%)	62 (4%)	18 (1%)	16	11

5 of 18 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	9	ALA
1	C	260	ASN
1	C	344	THR
1	D	280	LYS
1	D	346	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	293/321 (91%)	273 (93%)	20 (7%)	20	21
1	B	297/321 (92%)	276 (93%)	21 (7%)	18	19
1	C	298/321 (93%)	281 (94%)	17 (6%)	25	29
1	D	295/321 (92%)	271 (92%)	24 (8%)	15	14
2	W	18/13 (138%)	14 (78%)	4 (22%)	1	1
2	X	9/13 (69%)	8 (89%)	1 (11%)	8	6
2	Y	9/13 (69%)	7 (78%)	2 (22%)	1	1
All	All	1219/1323 (92%)	1130 (93%)	89 (7%)	19	18

5 of 89 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	339	VAL
1	C	261	ILE
2	W	939[A]	GLN
1	C	23	ASN
1	C	133	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 57 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	316	ASN
1	C	100	HIS
1	D	212	HIS
1	B	335	ASN
1	C	23	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 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	ADP	A	603	4	22,29,29	1.13	2 (9%)	27,45,45	1.85	6 (22%)
3	ADP	B	603	4	22,29,29	1.04	2 (9%)	27,45,45	2.21	7 (25%)
3	ADP	C	603	4	22,29,29	1.15	1 (4%)	27,45,45	2.18	7 (25%)
3	ADP	D	603	4	22,29,29	1.11	1 (4%)	27,45,45	2.16	7 (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
3	ADP	A	603	4	-	0/12/32/32	0/3/3/3
3	ADP	B	603	4	-	0/12/32/32	0/3/3/3
3	ADP	C	603	4	-	0/12/32/32	0/3/3/3
3	ADP	D	603	4	-	0/12/32/32	0/3/3/3

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	603	ADP	O4'-C1'	2.09	1.43	1.41
3	A	603	ADP	O4'-C1'	2.57	1.44	1.41

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	603	ADP	C5-C4	2.90	1.47	1.40
3	B	603	ADP	C5-C4	3.01	1.47	1.40
3	D	603	ADP	C5-C4	3.72	1.48	1.40

The worst 5 of 27 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	603	ADP	N3-C2-N1	-7.85	122.89	128.89
3	C	603	ADP	N3-C2-N1	-7.02	123.52	128.89
3	D	603	ADP	C2'-C1'-N9	-5.84	105.36	114.29
3	A	603	ADP	N3-C2-N1	-5.75	124.50	128.89
3	D	603	ADP	N3-C2-N1	-5.50	124.68	128.89

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	328/365 (89%)	0.10	21 (6%) 23 22	19, 32, 67, 99	0
1	B	341/365 (93%)	0.22	27 (7%) 15 15	21, 39, 70, 96	0
1	C	343/365 (93%)	0.19	35 (10%) 9 8	20, 36, 70, 91	0
1	D	338/365 (92%)	0.54	41 (12%) 6 5	31, 50, 83, 95	0
2	W	13/20 (65%)	0.77	2 (15%) 3 2	44, 49, 68, 74	3 (23%)
2	X	13/20 (65%)	1.36	3 (23%) 1 1	39, 44, 68, 69	13 (100%)
2	Y	12/20 (60%)	0.50	2 (16%) 2 2	28, 34, 46, 47	12 (100%)
All	All	1388/1520 (91%)	0.28	131 (9%) 11 10	19, 39, 75, 99	28 (2%)

The worst 5 of 131 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	X	938	PRO	9.2
1	D	352	TYR	8.9
1	C	8	PRO	8.9
1	D	281	THR	6.7
1	B	360	ALA	6.6

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(Å ²)	Q<0.9
3	ADP	B	603	27/27	0.98	0.10	-0.63	18,26,31,32	0
3	ADP	A	603	27/27	0.98	0.12	-0.86	14,23,28,30	0
3	ADP	C	603	27/27	0.98	0.10	-0.93	20,29,34,37	0
3	ADP	D	603	27/27	0.97	0.09	-1.09	29,35,39,44	0
4	MG	D	800	1/1	0.88	0.10	-	36,36,36,36	0
4	MG	A	800	1/1	0.99	0.22	-	20,20,20,20	0
4	MG	B	800	1/1	0.99	0.15	-	20,20,20,20	0
4	MG	C	800	1/1	0.91	0.10	-	28,28,28,28	0

6.5 Other polymers [i](#)

There are no such residues in this entry.