



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 12:36 am GMT

PDB ID : 4FI3
Title : Structure of vitamin B12 transporter BtuCD-F in a nucleotide-bound state
Authors : Korkhov, V.M.; Mireku, S.A.; Locher, K.P.
Deposited on : 2012-06-07
Resolution : 3.47 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 : 1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix) : 1.9-1692
EDS : trunk28620
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
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) : recalc28949

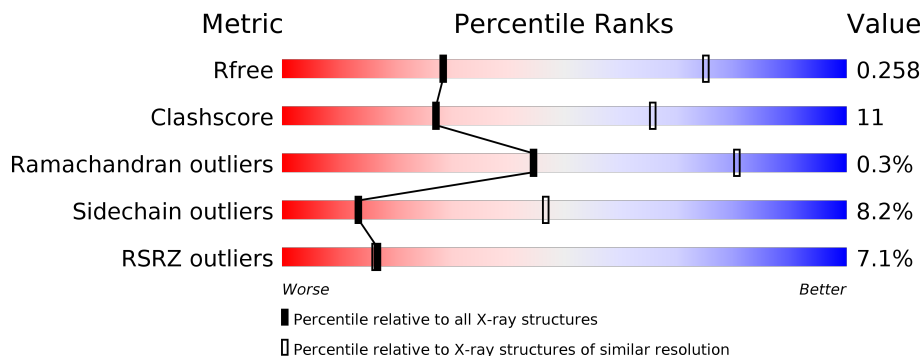
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.47 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}	100719	1135 (3.56-3.36)
Clashscore	112137	1040 (3.52-3.40)
Ramachandran outliers	110173	1009 (3.52-3.40)
Sidechain outliers	110143	1010 (3.52-3.40)
RSRZ outliers	101464	1017 (3.54-3.38)

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	349	<div> <div>3%</div> <div> <div></div> <div>69%</div> <div>21%</div> <div>7%</div> </div> </div>
1	B	349	<div> <div>5%</div> <div> <div></div> <div>67%</div> <div>23%</div> <div>7%</div> </div> </div>
2	C	249	<div> <div></div> <div> <div></div> <div>68%</div> <div>27%</div> <div></div> </div> </div>
2	D	249	<div> <div>%</div> <div> <div></div> <div>69%</div> <div>28%</div> <div></div> </div> </div>
3	F	255	<div> <div>27%</div> <div> <div></div> <div>72%</div> <div>23%</div> <div></div> </div> </div>

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 10635 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 Vitamin B12 import system permease protein BtuC.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	324	Total	C	N	O	S	0	0	0
			2441	1611	418	400	12			
1	B	324	Total	C	N	O	S	0	0	0
			2441	1611	418	400	12			

There are 60 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-22	MET	-	EXPRESSION TAG	UNP P06609
A	-21	GLY	-	EXPRESSION TAG	UNP P06609
A	-20	HIS	-	EXPRESSION TAG	UNP P06609
A	-19	HIS	-	EXPRESSION TAG	UNP P06609
A	-18	HIS	-	EXPRESSION TAG	UNP P06609
A	-17	HIS	-	EXPRESSION TAG	UNP P06609
A	-16	HIS	-	EXPRESSION TAG	UNP P06609
A	-15	HIS	-	EXPRESSION TAG	UNP P06609
A	-14	HIS	-	EXPRESSION TAG	UNP P06609
A	-13	HIS	-	EXPRESSION TAG	UNP P06609
A	-12	HIS	-	EXPRESSION TAG	UNP P06609
A	-11	HIS	-	EXPRESSION TAG	UNP P06609
A	-10	SER	-	EXPRESSION TAG	UNP P06609
A	-9	SER	-	EXPRESSION TAG	UNP P06609
A	-8	GLY	-	EXPRESSION TAG	UNP P06609
A	-7	GLU	-	EXPRESSION TAG	UNP P06609
A	-6	ASN	-	EXPRESSION TAG	UNP P06609
A	-5	LEU	-	EXPRESSION TAG	UNP P06609
A	-4	TYR	-	EXPRESSION TAG	UNP P06609
A	-3	PHE	-	EXPRESSION TAG	UNP P06609
A	-2	GLN	-	EXPRESSION TAG	UNP P06609
A	-1	GLY	-	EXPRESSION TAG	UNP P06609
A	0	HIS	-	EXPRESSION TAG	UNP P06609
A	18	SER	CYS	ENGINEERED MUTATION	UNP P06609
A	32	SER	CYS	ENGINEERED MUTATION	UNP P06609

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	120	SER	CYS	ENGINEERED MUTATION	UNP P06609
A	156	SER	CYS	ENGINEERED MUTATION	UNP P06609
A	205	SER	CYS	ENGINEERED MUTATION	UNP P06609
A	206	SER	CYS	ENGINEERED MUTATION	UNP P06609
A	267	SER	CYS	ENGINEERED MUTATION	UNP P06609
B	-22	MET	-	EXPRESSION TAG	UNP P06609
B	-21	GLY	-	EXPRESSION TAG	UNP P06609
B	-20	HIS	-	EXPRESSION TAG	UNP P06609
B	-19	HIS	-	EXPRESSION TAG	UNP P06609
B	-18	HIS	-	EXPRESSION TAG	UNP P06609
B	-17	HIS	-	EXPRESSION TAG	UNP P06609
B	-16	HIS	-	EXPRESSION TAG	UNP P06609
B	-15	HIS	-	EXPRESSION TAG	UNP P06609
B	-14	HIS	-	EXPRESSION TAG	UNP P06609
B	-13	HIS	-	EXPRESSION TAG	UNP P06609
B	-12	HIS	-	EXPRESSION TAG	UNP P06609
B	-11	HIS	-	EXPRESSION TAG	UNP P06609
B	-10	SER	-	EXPRESSION TAG	UNP P06609
B	-9	SER	-	EXPRESSION TAG	UNP P06609
B	-8	GLY	-	EXPRESSION TAG	UNP P06609
B	-7	GLU	-	EXPRESSION TAG	UNP P06609
B	-6	ASN	-	EXPRESSION TAG	UNP P06609
B	-5	LEU	-	EXPRESSION TAG	UNP P06609
B	-4	TYR	-	EXPRESSION TAG	UNP P06609
B	-3	PHE	-	EXPRESSION TAG	UNP P06609
B	-2	GLN	-	EXPRESSION TAG	UNP P06609
B	-1	GLY	-	EXPRESSION TAG	UNP P06609
B	0	HIS	-	EXPRESSION TAG	UNP P06609
B	18	SER	CYS	ENGINEERED MUTATION	UNP P06609
B	32	SER	CYS	ENGINEERED MUTATION	UNP P06609
B	120	SER	CYS	ENGINEERED MUTATION	UNP P06609
B	156	SER	CYS	ENGINEERED MUTATION	UNP P06609
B	205	SER	CYS	ENGINEERED MUTATION	UNP P06609
B	206	SER	CYS	ENGINEERED MUTATION	UNP P06609
B	267	SER	CYS	ENGINEERED MUTATION	UNP P06609

- Molecule 2 is a protein called Vitamin B12 import ATP-binding protein BtuD.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	248	Total	C	N	O	S	0	0	0
			1891	1183	351	348	9			
2	D	248	Total	C	N	O	S	0	0	0
			1891	1183	351	348	9			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	159	GLN	GLU	ENGINEERED MUTATION	UNP P06611
C	162	CYS	ASN	ENGINEERED MUTATION	UNP P06611
C	180	SER	CYS	ENGINEERED MUTATION	UNP P06611
D	159	GLN	GLU	ENGINEERED MUTATION	UNP P06611
D	162	CYS	ASN	ENGINEERED MUTATION	UNP P06611
D	180	SER	CYS	ENGINEERED MUTATION	UNP P06611

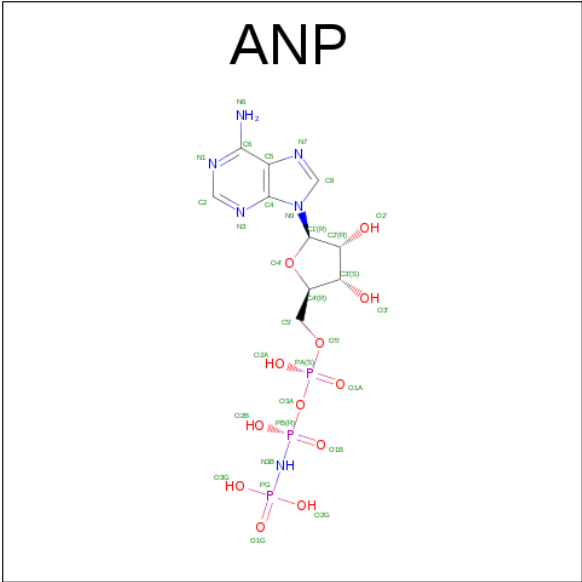
- Molecule 3 is a protein called Vitamin B12-binding protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	F	245	Total	C	N	O	S	0	0	0
			1907	1216	332	355	4			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	21	MET	-	EXPRESSION TAG	UNP P37028
F	267	SER	-	EXPRESSION TAG	UNP P37028
F	268	GLY	-	EXPRESSION TAG	UNP P37028
F	269	SER	-	EXPRESSION TAG	UNP P37028
F	270	HIS	-	EXPRESSION TAG	UNP P37028
F	271	HIS	-	EXPRESSION TAG	UNP P37028
F	272	HIS	-	EXPRESSION TAG	UNP P37028
F	273	HIS	-	EXPRESSION TAG	UNP P37028
F	274	HIS	-	EXPRESSION TAG	UNP P37028
F	275	HIS	-	EXPRESSION TAG	UNP P37028

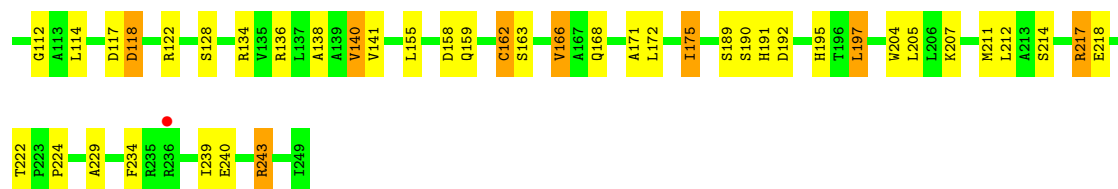
- Molecule 4 is PHOSPHOAMINOPHOSPHONIC ACID-ADENYLATE ESTER (three-letter code: ANP) (formula: C₁₀H₁₇N₆O₁₂P₃).



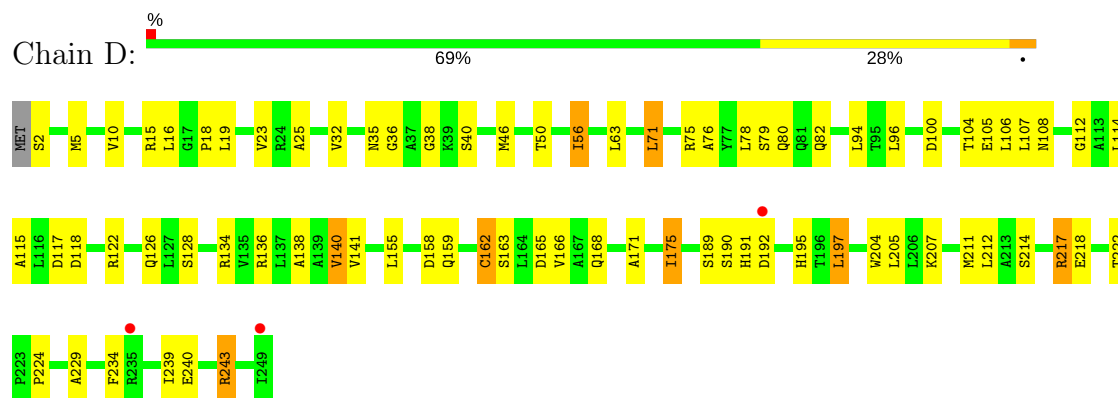
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	C	1	Total	C	N	O	P	0	0
			31	10	6	12	3		
4	D	1	Total	C	N	O	P	0	0
			31	10	6	12	3		

- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

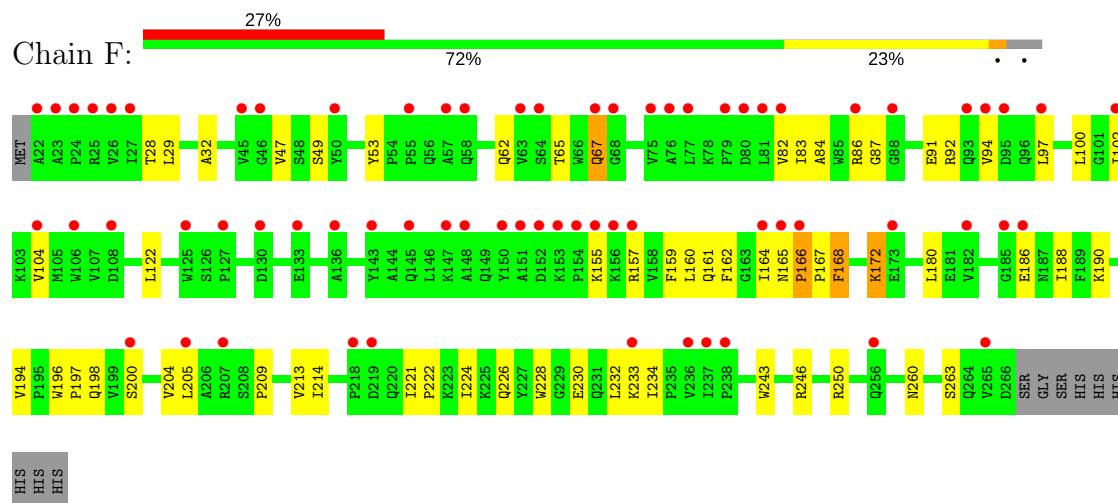
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	D	1	Total	Mg	0	0
			1	1		
5	C	1	Total	Mg	0	0
			1	1		



• Molecule 2: Vitamin B12 import ATP-binding protein BtuD



• Molecule 3: Vitamin B12-binding protein



4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	127.73Å 211.99Å 179.10Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 3.47 30.14 – 3.47	Depositor EDS
% Data completeness (in resolution range)	96.3 (30.00-3.47) 96.4 (30.14-3.47)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.13	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.29 (at 3.47Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.2_869)	Depositor
R, R_{free}	0.229 , 0.274 0.214 , 0.258	Depositor DCC
R_{free} test set	2007 reflections (6.54%)	DCC
Wilson B-factor (Å ²)	118.0	Xtriage
Anisotropy	0.008	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.27 , 94.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.002 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.020 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	10635	wwPDB-VP
Average B, all atoms (Å ²)	169.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.78% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 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: MG, ANP

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	0.33	0/2495	0.49	0/3408
1	B	0.31	0/2495	0.49	0/3408
2	C	0.37	0/1925	0.57	0/2608
2	D	0.34	0/1925	0.56	0/2608
3	F	0.27	0/1951	0.44	0/2661
All	All	0.32	0/10791	0.51	0/14693

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts [i](#)

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	2441	0	2606	62	0
1	B	2441	0	2606	61	0
2	C	1891	0	1925	51	0
2	D	1891	0	1925	45	0
3	F	1907	0	1924	51	0
4	C	31	0	13	6	0
4	D	31	0	13	5	0
5	C	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	D	1	0	0	0	0
All	All	10635	0	11012	247	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 247 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)
3:F:168:PHE:HB2	3:F:198:GLN:HE22	1.42	0.84
1:B:111:GLN:HB3	1:B:113:PRO:HD2	1.59	0.84
1:A:167:SER:HB3	1:A:171:ASP:HB2	1.61	0.82
2:C:5:MET:HB3	2:C:23:VAL:HB	1.59	0.82
1:A:111:GLN:HB3	1:A:113:PRO:HD2	1.60	0.82

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	322/349 (92%)	303 (94%)	19 (6%)	0	100	100
1	B	322/349 (92%)	303 (94%)	19 (6%)	0	100	100
2	C	246/249 (99%)	238 (97%)	7 (3%)	1 (0%)	38	77
2	D	246/249 (99%)	238 (97%)	8 (3%)	0	100	100
3	F	243/255 (95%)	229 (94%)	11 (4%)	3 (1%)	15	56
All	All	1379/1451 (95%)	1311 (95%)	64 (5%)	4 (0%)	44	80

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	F	230	GLU
3	F	233	LYS
2	C	166	VAL
3	F	166	PRO

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	249/270 (92%)	229 (92%)	20 (8%)	14	48
1	B	249/270 (92%)	229 (92%)	20 (8%)	14	48
2	C	200/201 (100%)	178 (89%)	22 (11%)	7	32
2	D	200/201 (100%)	178 (89%)	22 (11%)	7	32
3	F	205/214 (96%)	199 (97%)	6 (3%)	48	79
All	All	1103/1156 (95%)	1013 (92%)	90 (8%)	13	47

5 of 90 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	15	ARG
2	C	106	LEU
2	D	243	ARG
2	C	16	LEU
2	C	71	LEU

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

Mol	Chain	Res	Type
2	D	81	GLN
3	F	198	GLN

5.3.3 RNA ⓘ

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](#)

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$
4	ANP	C	301	5	29,33,33	2.16	9 (31%)	28,52,52	2.63	10 (35%)
4	ANP	D	301	5	29,33,33	2.04	8 (27%)	28,52,52	2.53	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
4	ANP	C	301	5	-	0/13/38/38	0/3/3/3
4	ANP	D	301	5	-	0/13/38/38	0/3/3/3

The worst 5 of 17 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	D	301	ANP	PB-O2B	-2.73	1.49	1.56
4	C	301	ANP	PB-O2B	-2.69	1.49	1.56
4	D	301	ANP	PG-O3G	-2.50	1.49	1.56
4	C	301	ANP	PG-O3G	-2.39	1.50	1.56
4	D	301	ANP	PG-O2G	-2.23	1.50	1.56

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	C	301	ANP	O1G-PG-N3B	-8.98	98.37	111.79
4	D	301	ANP	O1G-PG-N3B	-8.95	98.40	111.79
4	C	301	ANP	N3-C2-N1	-6.10	123.54	128.86
4	D	301	ANP	N3-C2-N1	-5.57	124.01	128.86
4	C	301	ANP	O1B-PB-N3B	-4.40	105.22	111.79

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	C	301	ANP	6	0
4	D	301	ANP	5	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	324/349 (92%)	-0.13	10 (3%) 49 43	85, 144, 237, 373	0
1	B	324/349 (92%)	-0.02	16 (4%) 30 26	103, 151, 239, 405	0
2	C	248/249 (99%)	-0.22	1 (0%) 92 89	68, 113, 195, 311	0
2	D	248/249 (99%)	-0.13	3 (1%) 79 72	79, 134, 200, 262	0
3	F	245/255 (96%)	1.29	68 (27%) 1 1	173, 255, 350, 525	0
All	All	1389/1451 (95%)	0.13	98 (7%) 17 16	68, 148, 286, 525	0

The worst 5 of 98 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	1	MET	7.3
3	F	186	GLU	5.9
3	F	45	VAL	5.7
1	A	46	PRO	5.7
3	F	67	GLN	5.7

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
4	ANP	C	301	31/31	0.90	0.19	-0.14	101,120,141,142	0
4	ANP	D	301	31/31	0.90	0.17	-0.28	101,123,145,146	0
5	MG	C	302	1/1	0.94	0.19	-	57,57,57,57	0
5	MG	D	302	1/1	0.90	0.16	-	57,57,57,57	0

6.5 Other polymers

There are no such residues in this entry.