



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 02:59 AM GMT

PDB ID : 2JLX  
Title : DENGUE VIRUS 4 NS3 HELICASE IN COMPLEX WITH SSRNA AND ADP-VANADATE  
Authors : Luo, D.H.; Xu, T.; Watson, R.P.; Becker, D.S.; Sampath, A.; Jahnke, W.; Yeong, S.S.; Wang, C.H.; Lim, S.P.; Vasudevan, S.G.; Lescar, J.  
Deposited on : 2008-09-15  
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](mailto: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

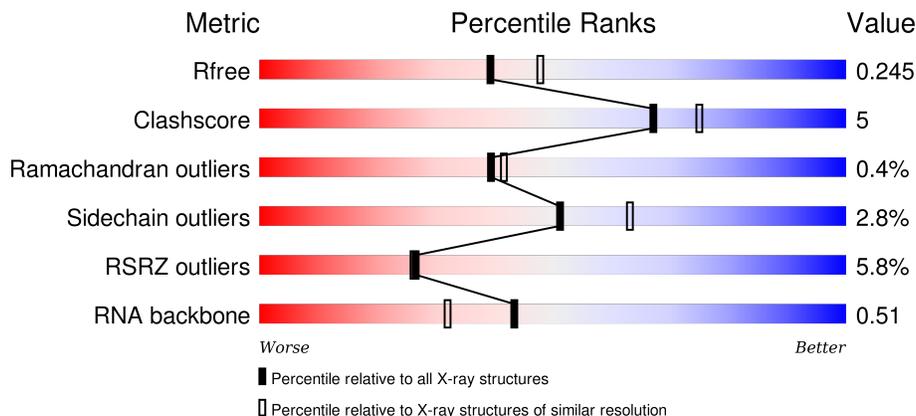
# 1 Overall quality at a glance i

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)
RNA backbone	2183	1062 (2.80-1.60)

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  $\geq 3$ , 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	451	<div style="display: flex; align-items: center;"> <div style="width: 5%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 85%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 15%; height: 10px; background-color: yellow;"></div> </div> <p style="margin-left: 5px;">5% <span style="margin-left: 100px;">85%</span> <span style="margin-left: 100px;">15%</span></p>
1	B	451	<div style="display: flex; align-items: center;"> <div style="width: 6%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 87%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 12%; height: 10px; background-color: yellow;"></div> </div> <p style="margin-left: 5px;">6% <span style="margin-left: 100px;">87%</span> <span style="margin-left: 100px;">12%</span> .</p>
2	C	12	<div style="display: flex; align-items: center;"> <div style="width: 17%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 42%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 17%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 8%; height: 10px; background-color: orange; margin-right: 5px;"></div> <div style="width: 33%; height: 10px; background-color: grey;"></div> </div> <p style="margin-left: 5px;">17% <span style="margin-left: 20px;">42%</span> <span style="margin-left: 20px;">17%</span> <span style="margin-left: 20px;">8%</span> <span style="margin-left: 20px;">33%</span></p>
2	D	12	<div style="display: flex; align-items: center;"> <div style="width: 17%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 58%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 8%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 33%; height: 10px; background-color: grey;"></div> </div> <p style="margin-left: 5px;">17% <span style="margin-left: 40px;">58%</span> <span style="margin-left: 40px;">8%</span> <span style="margin-left: 40px;">33%</span></p>

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 7975 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 SERINE PROTEASE SUBUNIT NS3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	451	Total 3609	C 2274	N 645	O 673	S 17	0	0	0
1	B	451	Total 3611	C 2276	N 643	O 675	S 17	0	1	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	250	ASP	GLU	CONFLICT	UNP Q2YHF0
A	292	CYS	SER	CONFLICT	UNP Q2YHF0
A	321	SER	THR	CONFLICT	UNP Q2YHF0
A	322	ILE	THR	CONFLICT	UNP Q2YHF0
A	381	ARG	LYS	CONFLICT	UNP Q2YHF0
A	480	LYS	ARG	CONFLICT	UNP Q2YHF0
B	250	ASP	GLU	CONFLICT	UNP Q2YHF0
B	292	CYS	SER	CONFLICT	UNP Q2YHF0
B	321	SER	THR	CONFLICT	UNP Q2YHF0
B	322	ILE	THR	CONFLICT	UNP Q2YHF0
B	381	ARG	LYS	CONFLICT	UNP Q2YHF0
B	480	LYS	ARG	CONFLICT	UNP Q2YHF0

- Molecule 2 is a RNA chain called 5'-R(\*AP\*GP\*AP\*CP\*UP\*AP\*AP\*CP\*AP\*AP\*CP\*U)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	C	8	Total 152	C 68	N 30	O 47	P 7	0	0	0
2	D	8	Total 152	C 68	N 30	O 47	P 7	0	0	0

- Molecule 3 is ADENOSINE-5'-DIPHOSPHATE (three-letter code: ADP) (formula: C<sub>10</sub>H<sub>15</sub>N<sub>5</sub>O<sub>10</sub>P<sub>2</sub>).



- Molecule 5 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	1	Total 1	Mn 1	0	0
5	A	1	Total 1	Mn 1	0	0

- Molecule 6 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	B	1	Total 1	Cl 1	0	0

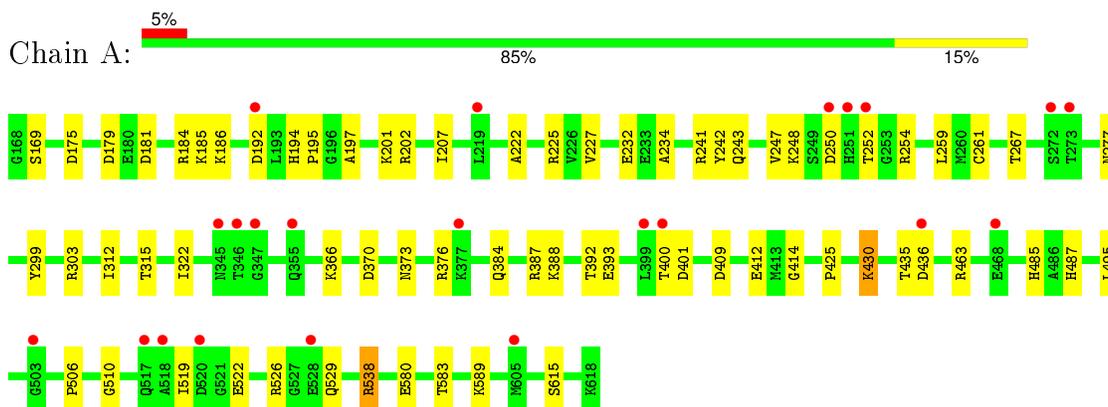
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	198	Total 198	O 198	0	0
7	B	172	Total 172	O 172	0	0
7	C	7	Total 7	O 7	0	0
7	D	7	Total 7	O 7	0	0

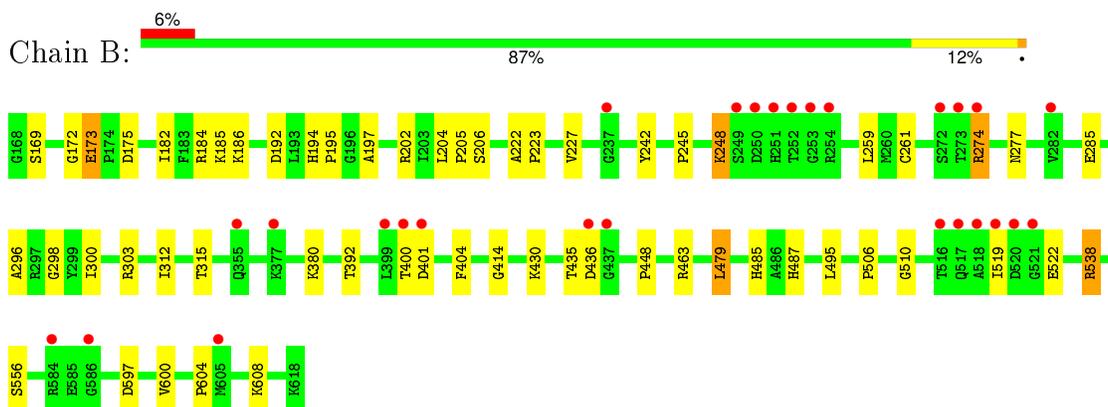
### 3 Residue-property plots [i](#)

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.

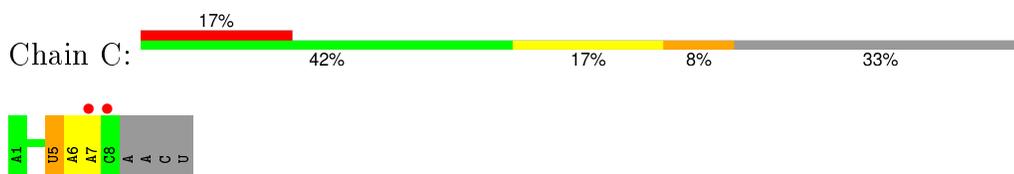
- Molecule 1: SERINE PROTEASE SUBUNIT NS3



- Molecule 1: SERINE PROTEASE SUBUNIT NS3

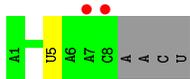


- Molecule 2: 5'-R(\*AP\*GP\*AP\*CP\*UP\*AP\*AP\*CP\*AP\*AP\*CP\*U)-3'



- Molecule 2: 5'-R(\*AP\*GP\*AP\*CP\*UP\*AP\*AP\*CP\*AP\*AP\*CP\*U)-3'





## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	132.54Å 105.21Å 72.35Å 90.00° 117.62° 90.00°	Depositor
Resolution (Å)	20.00 – 2.20 19.81 – 2.20	Depositor EDS
% Data completeness (in resolution range)	98.6 (20.00-2.20) 98.6 (19.81-2.20)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.47 (at 2.19Å)	Xtrriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.196 , 0.245 0.196 , 0.245	Depositor DCC
$R_{free}$ test set	2213 reflections (5.30%)	DCC
Wilson B-factor (Å <sup>2</sup> )	24.7	Xtrriage
Anisotropy	0.042	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.41 , 53.7	EDS
Estimated twinning fraction	No twinning to report.	Xtrriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Outliers	2 of 43955 reflections (0.005%)	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	7975	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	21.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.77% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>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: VO4, MN, ADP, CL

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	0.46	0/3690	0.59	0/4994
1	B	0.44	0/3692	0.58	0/4999
2	C	0.77	0/170	1.34	1/264 (0.4%)
2	D	0.74	0/170	1.26	1/264 (0.4%)
All	All	0.46	0/7722	0.64	2/10521 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	5	U	O4'-C1'-N1	7.01	113.81	108.20
2	C	5	U	O4'-C1'-N1	5.05	112.24	108.20

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	3609	0	3600	41	0
1	B	3611	0	3594	34	0
2	C	152	0	77	1	0
2	D	152	0	77	0	0
3	A	27	0	12	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	27	0	12	0	0
4	A	5	0	0	1	0
4	B	5	0	0	1	0
5	A	1	0	0	0	0
5	B	1	0	0	0	0
6	B	1	0	0	0	0
7	A	198	0	0	4	1
7	B	172	0	0	0	0
7	C	7	0	0	0	0
7	D	7	0	0	0	0
All	All	7975	0	7372	72	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 5.

The worst 5 of 72 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:248:LYS:HG2	1:B:248:LYS:NZ	2.02	0.74
1:B:485:HIS:HD2	1:B:487:HIS:H	1.35	0.73
1:B:169:SER:HB3	1:B:175:ASP:HA	1.69	0.73
1:A:197:ALA:O	1:A:202:ARG:NH1	2.23	0.71
1:A:485:HIS:HD2	1:A:487:HIS:H	1.41	0.69

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)
7:A:2025:HOH:O	7:A:2119:HOH:O[4_545]	0.00	2.20

## 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	449/451 (100%)	432 (96%)	15 (3%)	2 (0%)	39	42
1	B	450/451 (100%)	429 (95%)	19 (4%)	2 (0%)	39	42
All	All	899/902 (100%)	861 (96%)	34 (4%)	4 (0%)	39	42

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	510	GLY
1	B	436	ASP
1	B	510	GLY
1	A	436	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	388/388 (100%)	376 (97%)	12 (3%)	47	59
1	B	388/388 (100%)	378 (97%)	10 (3%)	54	66
All	All	776/776 (100%)	754 (97%)	22 (3%)	51	63

5 of 22 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	538	ARG
1	B	173	GLU
1	B	538	ARG
1	A	580	GLU
1	A	615	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	529	GLN
1	A	576	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	384	GLN
1	A	517	GLN
1	B	279	ASN

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	C	6/12 (50%)	1 (16%)	1 (16%)
2	D	6/12 (50%)	0	0
All	All	12/24 (50%)	1 (8%)	1 (8%)

All (1) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	C	6	A

All (1) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
2	C	5	U

## 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 7 ligands modelled in this entry, 3 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	1619	5,4	22,29,29	1.15	2 (9%)	27,45,45	1.85	3 (11%)
4	VO4	A	1620	3,5	1,4,4	4.09	1 (100%)	0,6,6	0.00	-
3	ADP	B	1619	5,4	22,29,29	1.19	3 (13%)	27,45,45	1.85	4 (14%)
4	VO4	B	1620	3,5	1,4,4	4.44	1 (100%)	0,6,6	0.00	-

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	1619	5,4	-	0/12/32/32	0/3/3/3
4	VO4	A	1620	3,5	-	0/0/0/0	0/0/0/0
3	ADP	B	1619	5,4	-	0/12/32/32	0/3/3/3
4	VO4	B	1620	3,5	-	0/0/0/0	0/0/0/0

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	1619	ADP	O4'-C1'	2.03	1.43	1.41
3	A	1619	ADP	C2-N3	2.12	1.36	1.32
3	B	1619	ADP	C2-N3	2.27	1.36	1.32
3	B	1619	ADP	C5-C4	3.43	1.48	1.40
3	A	1619	ADP	C5-C4	3.44	1.48	1.40

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1619	ADP	N3-C2-N1	-7.12	123.44	128.89
3	B	1619	ADP	N3-C2-N1	-7.02	123.52	128.89
3	B	1619	ADP	C4-C5-N7	-3.41	106.34	109.48
3	A	1619	ADP	C4-C5-N7	-2.94	106.78	109.48
3	A	1619	ADP	PA-O3A-PB	-2.57	124.03	132.67

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	1620	VO4	1	0
4	B	1620	VO4	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, 95<sup>th</sup> 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 <sup>2</sup> )	Q < 0.9
1	A	451/451 (100%)	0.14	22 (4%) 33 33	9, 20, 31, 42	3 (0%)
1	B	451/451 (100%)	0.20	27 (5%) 25 25	9, 20, 31, 42	2 (0%)
2	C	8/12 (66%)	1.00	2 (25%) 1 1	17, 20, 53, 58	0
2	D	8/12 (66%)	0.99	2 (25%) 1 1	18, 19, 51, 56	0
All	All	918/926 (99%)	0.19	53 (5%) 26 26	9, 20, 32, 58	5 (0%)

The worst 5 of 53 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	272	SER	6.8
1	B	273	THR	6.5
1	A	399	LEU	6.4
1	B	250	ASP	5.5
2	C	8	C	4.9

### 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, 95<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
3	ADP	B	1619	27/27	0.94	0.20	1.77	15,45,50,50	0
3	ADP	A	1619	27/27	0.94	0.17	0.71	15,41,46,46	0
4	VO4	A	1620	5/5	0.98	0.09	-1.81	26,26,27,27	0
4	VO4	B	1620	5/5	0.98	0.08	-1.92	25,26,27,28	0
5	MN	B	1621	1/1	0.99	0.04	-	17,17,17,17	0
5	MN	A	1621	1/1	1.00	0.03	-	17,17,17,17	0
6	CL	B	1622	1/1	0.97	0.09	-	34,34,34,34	0

## 6.5 Other polymers [\(i\)](#)

There are no such residues in this entry.