



wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 26, 2017 – 07:12 PM EDT

PDB ID : 1UF2
Title : The Atomic Structure of Rice dwarf Virus (RDV)
Authors : Nakagawa, A.; Miyazaki, N.; Taka, J.; Naitow, H.; Ogawa, A.; Fujimoto, Z.; Mizuno, H.; Higashi, T.; Watanabe, Y.; Omura, T.; Cheng, R.H.; Tsukihara, T.
Deposited on : unknown
Resolution : 3.50 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
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20030345
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)	:	rb-20030345

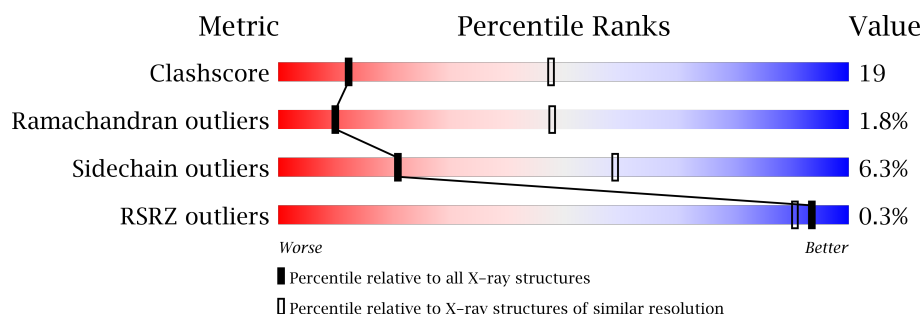
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.50 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))
Clashscore	112137	1322 (3.60-3.40)
Ramachandran outliers	110173	1283 (3.60-3.40)
Sidechain outliers	110143	1284 (3.60-3.40)
RSRZ outliers	101464	1226 (3.60-3.40)

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	1019	
1	B	1019	
2	C	421	
2	D	421	
2	E	421	
2	F	421	
2	G	421	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	H	421	<div><div></div><div>65%33%</div><div></div></div>
2	I	421	<div><div></div><div>59%36%5%</div><div></div></div>
2	J	421	<div><div></div><div>56%38%</div><div></div></div>
2	P	421	<div><div></div><div>%62%33%</div><div></div></div>
2	Q	421	<div><div></div><div>59%37%</div><div></div></div>
2	R	421	<div><div></div><div>69%28%</div><div></div></div>
2	S	421	<div><div></div><div>62%33%5%</div><div></div></div>
2	T	421	<div><div></div><div>65%32%</div><div></div></div>
3	K	506	<div><div></div><div>98%</div><div></div></div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 58130 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 Core protein P3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	967	Total	C	N	O	S	0	0	0
			7653	4885	1309	1427	32			
1	B	1019	Total	C	N	O	S	0	0	0
			8053	5138	1370	1512	33			

- Molecule 2 is a protein called Outer capsid protein P8.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	P	414	Total	C	N	O	S	0	0	0
			3227	2060	542	608	17			
2	C	415	Total	C	N	O	S	0	0	0
			3234	2065	543	609	17			
2	D	417	Total	C	N	O	S	0	0	0
			3247	2073	545	612	17			
2	Q	421	Total	C	N	O	S	0	0	0
			3274	2090	549	618	17			
2	E	421	Total	C	N	O	S	0	0	0
			3274	2090	549	618	17			
2	F	415	Total	C	N	O	S	0	0	0
			3234	2065	543	609	17			
2	R	421	Total	C	N	O	S	0	0	0
			3274	2090	549	618	17			
2	G	417	Total	C	N	O	S	0	0	0
			3253	2079	545	612	17			
2	H	421	Total	C	N	O	S	0	0	0
			3274	2090	549	618	17			
2	S	421	Total	C	N	O	S	0	0	0
			3274	2090	549	618	17			
2	I	419	Total	C	N	O	S	0	0	0
			3265	2085	547	616	17			
2	J	413	Total	C	N	O	S	0	0	0
			3221	2057	541	606	17			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	T	421	Total	C	N	O	S	0	0	0
			3274	2090	549	618	17			

There are 13 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
P	178	ALA	-	SEE REMARK 999	UNP P17379
C	178	ALA	-	SEE REMARK 999	UNP P17379
D	178	ALA	-	SEE REMARK 999	UNP P17379
Q	178	ALA	-	SEE REMARK 999	UNP P17379
E	178	ALA	-	SEE REMARK 999	UNP P17379
F	178	ALA	-	SEE REMARK 999	UNP P17379
R	178	ALA	-	SEE REMARK 999	UNP P17379
G	178	ALA	-	SEE REMARK 999	UNP P17379
H	178	ALA	-	SEE REMARK 999	UNP P17379
S	178	ALA	-	SEE REMARK 999	UNP P17379
I	178	ALA	-	SEE REMARK 999	UNP P17379
J	178	ALA	-	SEE REMARK 999	UNP P17379
T	178	ALA	-	SEE REMARK 999	UNP P17379

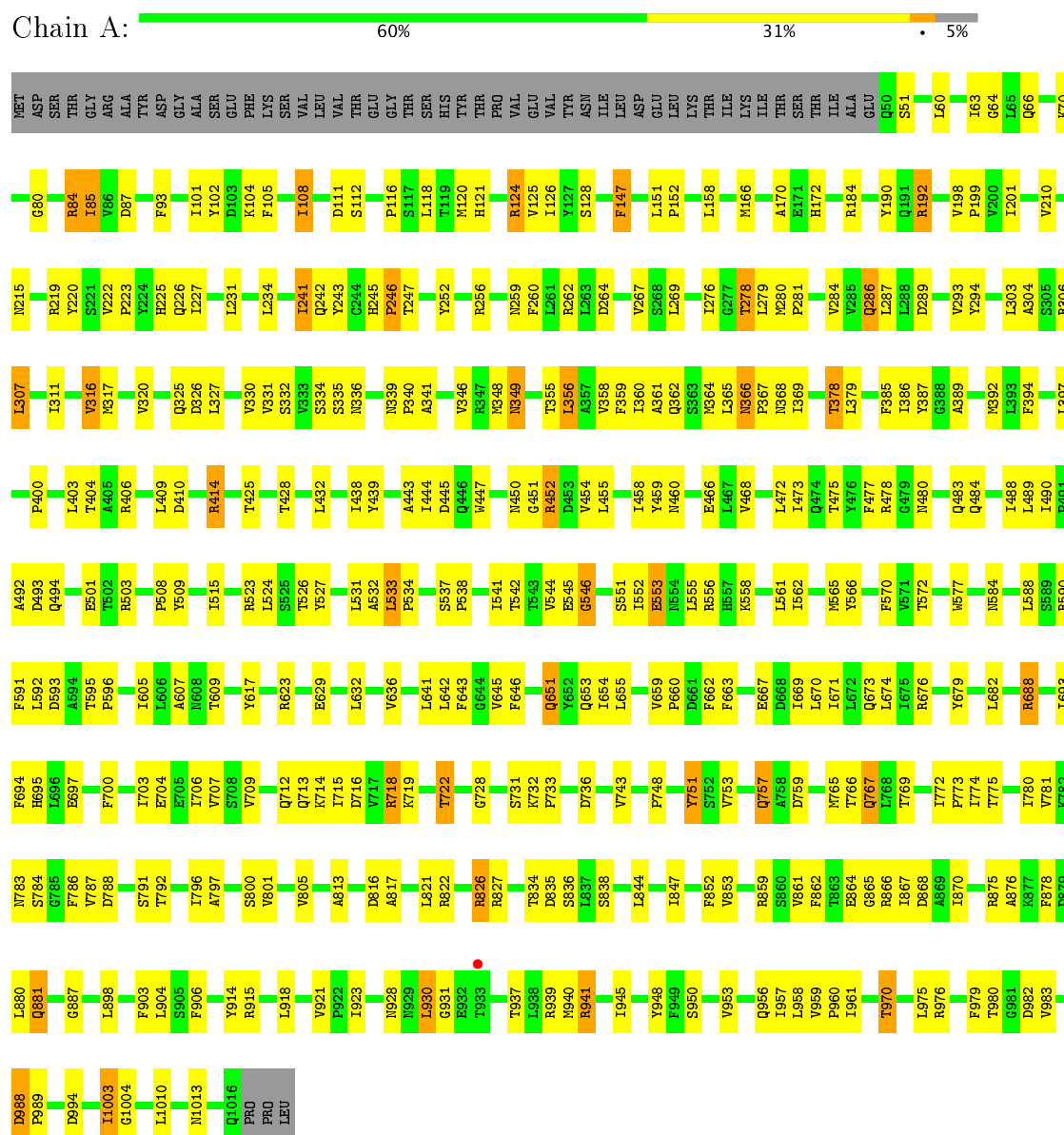
- Molecule 3 is a protein called Structural protein P7.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	K	12	Total	C	N	O	0	0	0
			99	60	16	23			

3 Residue-property plots

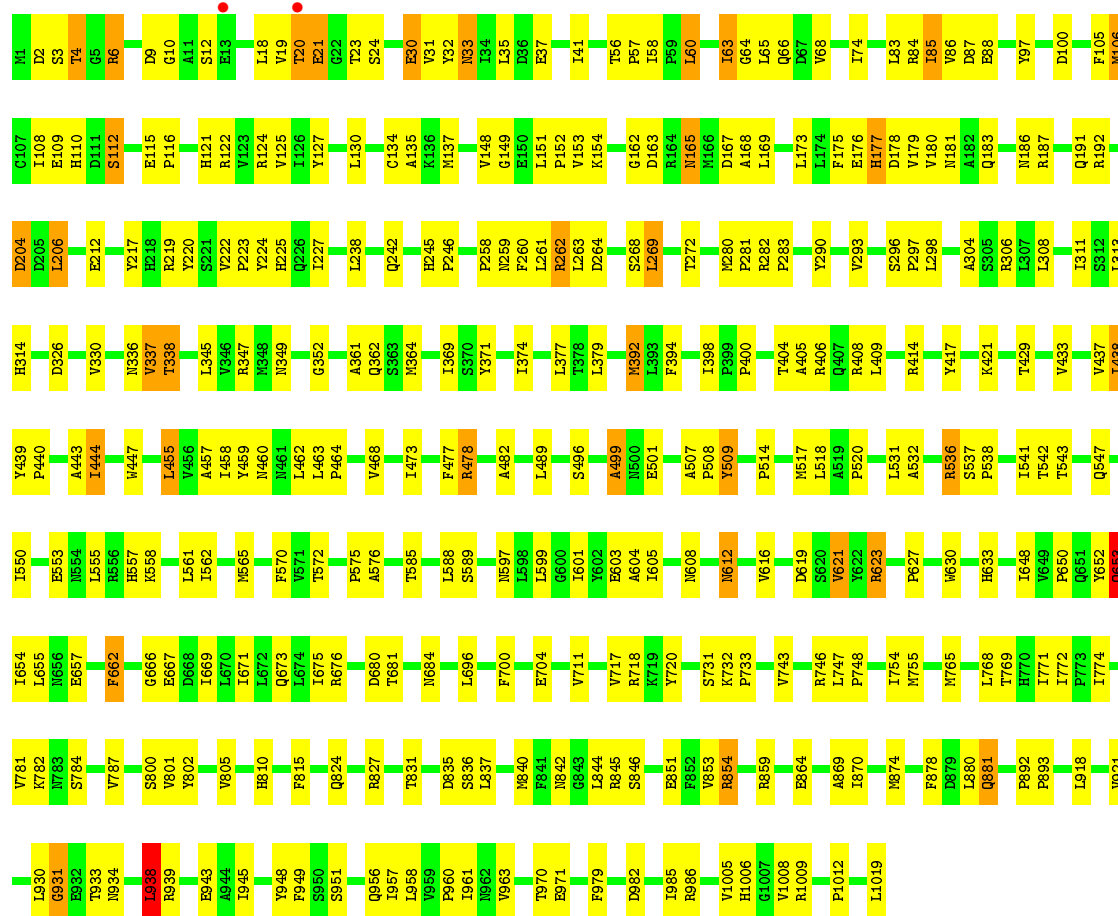
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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: Core protein P3



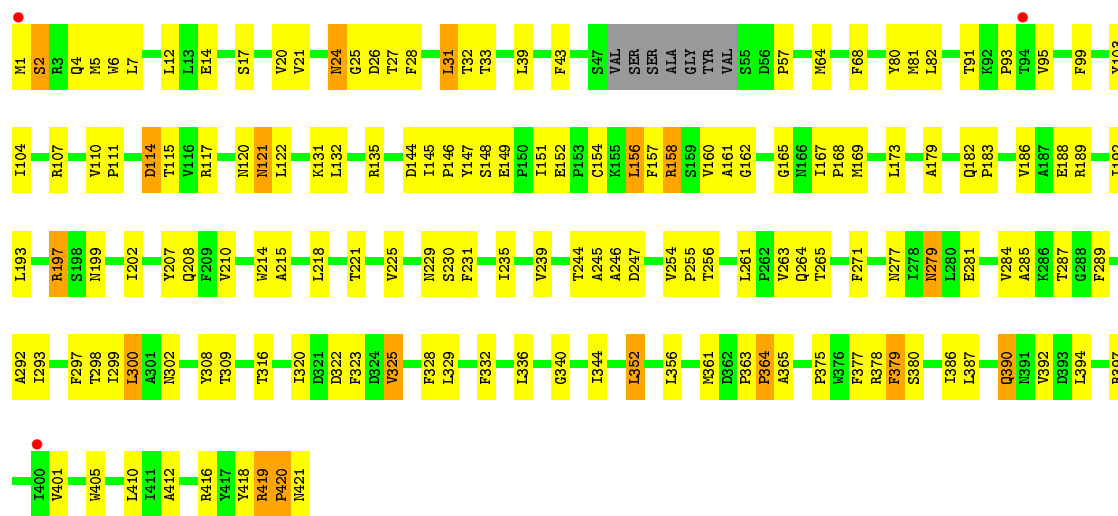
- Molecule 1: Core protein P3

Chain B:  68% 28%



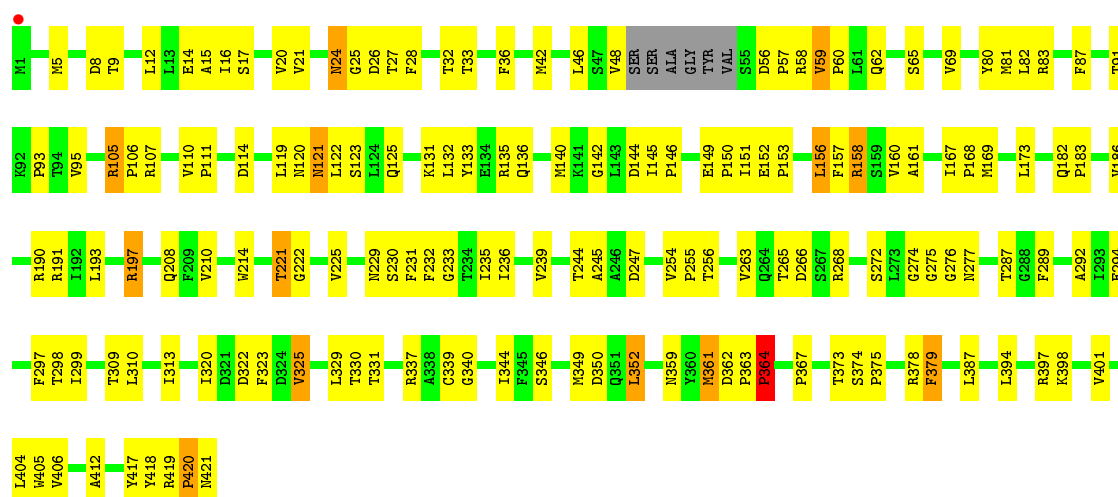
• Molecule 2: Outer capsid protein P8

Chain P:  62% 33%



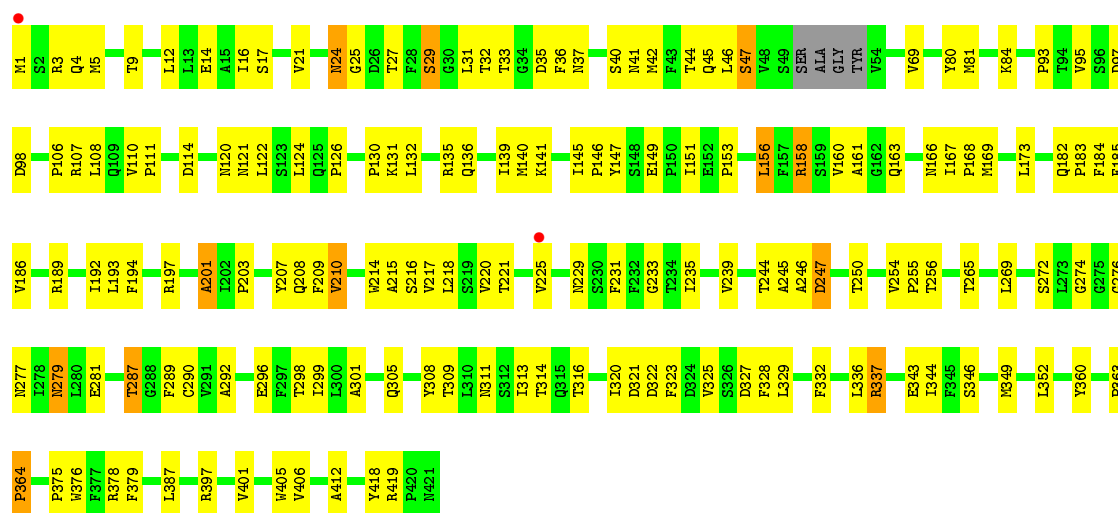
• Molecule 2: Outer capsid protein P8

Chain C: 



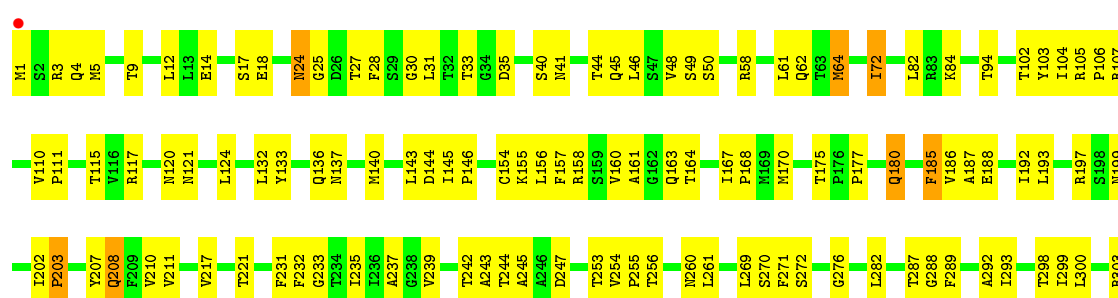
• Molecule 2: Outer capsid protein P8

Chain D: 

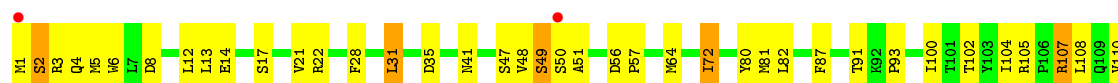


• Molecule 2: Outer capsid protein P8

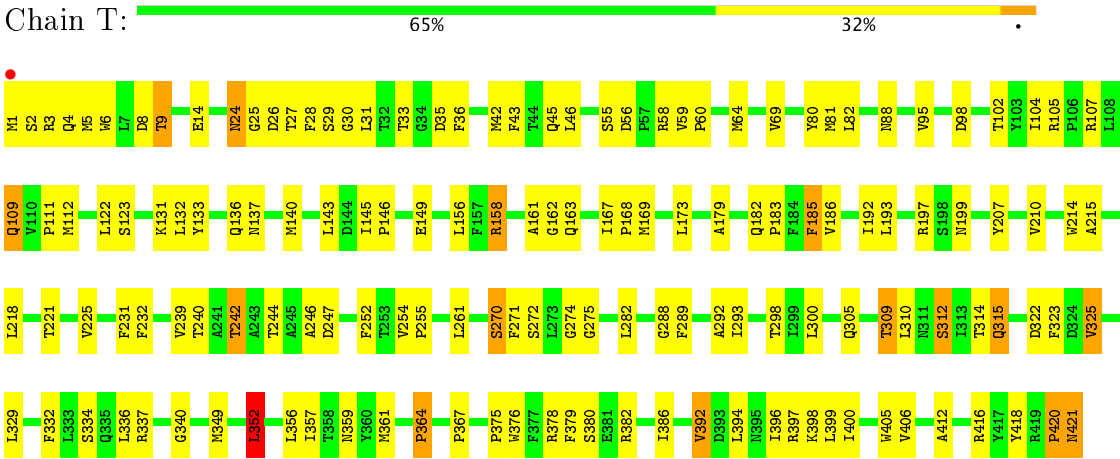
Chain Q: 



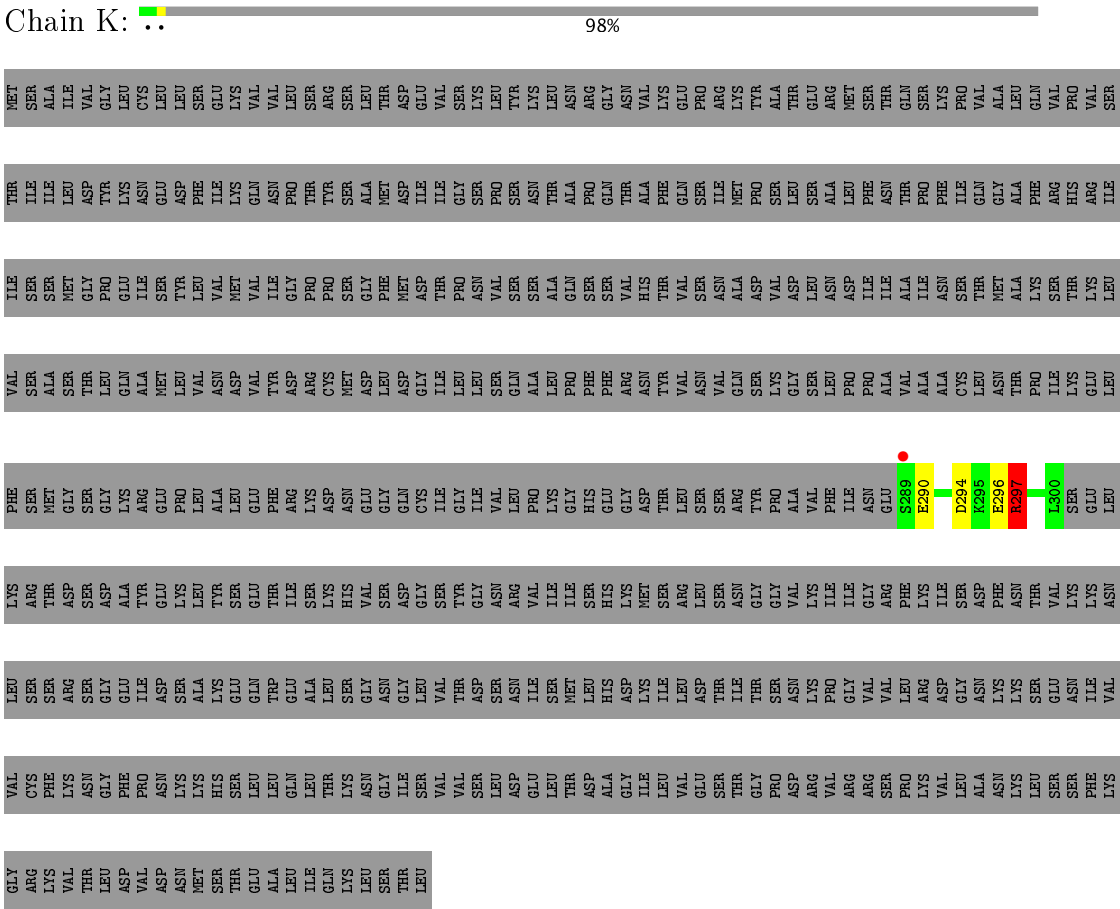








● Molecule 3: Structural protein P7



4 Data and refinement statistics

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α , β , γ	770.00Å 795.00Å 814.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	228.74 – 3.50 327.81 – 3.50	Depositor EDS
% Data completeness (in resolution range)	90.1 (228.74-3.50) 90.2 (327.81-3.50)	Depositor EDS
R_{merge}	0.19	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.50 (at 3.49Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.303 , 0.306 0.269 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	DCC
Wilson B-factor (Å ²)	59.9	Xtriage
Anisotropy	0.048	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.24 , 24.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.009 for -h,-l,-k	Xtriage
F_o, F_c correlation	0.82	EDS
Total number of atoms	58130	wwPDB-VP
Average B, all atoms (Å ²)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.10% 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 ⓘ

5.1 Standard geometry ⓘ

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	0.49	0/7816	0.72	1/10641 (0.0%)
1	B	0.52	0/8224	0.74	6/11197 (0.1%)
2	C	0.41	0/3308	0.65	0/4508
2	D	0.42	0/3321	0.69	1/4526 (0.0%)
2	E	0.45	0/3350	0.69	0/4567
2	F	0.44	0/3308	0.70	0/4508
2	G	0.48	0/3328	0.70	0/4536
2	H	0.52	3/3350 (0.1%)	0.72	0/4567
2	I	0.47	0/3340	0.70	0/4552
2	J	0.43	0/3295	0.68	0/4490
2	P	0.41	0/3301	0.65	0/4498
2	Q	0.45	0/3350	0.69	0/4567
2	R	0.46	0/3350	0.69	0/4567
2	S	0.43	0/3350	0.68	0/4567
2	T	0.51	1/3350 (0.0%)	0.73	1/4567 (0.0%)
3	K	1.00	0/100	1.28	0/132
All	All	0.47	4/59441 (0.0%)	0.70	9/80990 (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	1
1	B	0	1
All	All	0	2

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	H	421	ASN	C-O	6.87	1.36	1.23
2	H	421	ASN	CB-CG	6.18	1.65	1.51

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	T	421	ASN	C-O	5.30	1.33	1.23
2	H	421	ASN	CA-CB	5.19	1.66	1.53

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	1019	LEU	CA-C-O	-8.56	102.13	120.10
1	B	1019	LEU	CB-CG-CD1	-8.01	97.38	111.00
1	A	651	GLN	N-CA-C	-7.82	89.89	111.00
1	B	149	GLY	N-CA-C	-5.66	98.96	113.10
1	B	1019	LEU	CB-CG-CD2	5.58	120.48	111.00

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	751	TYR	Sidechain
1	B	509	TYR	Sidechain

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	7653	0	7664	275	0
1	B	8053	0	8062	248	0
2	C	3234	0	3208	145	0
2	D	3247	0	3222	140	0
2	E	3274	0	3245	147	0
2	F	3234	0	3208	172	0
2	G	3253	0	3226	157	0
2	H	3274	0	3245	132	0
2	I	3265	0	3236	152	0
2	J	3221	0	3194	144	0
2	P	3227	0	3199	123	0
2	Q	3274	0	3245	158	0
2	R	3274	0	3245	131	0
2	S	3274	0	3245	147	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	T	3274	0	3245	107	0
3	K	99	0	89	2	0
All	All	58130	0	57778	2194	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 19.

The worst 5 of 2194 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)
2:S:235:ILE:HD11	2:J:235:ILE:HD12	1.19	1.14
2:S:361:MET:HE1	2:S:367:PRO:HD3	1.24	1.12
2:P:235:ILE:HD13	2:D:229:ASN:HD21	1.06	1.09
2:S:344:ILE:HD12	2:S:344:ILE:H	1.20	1.07
1:B:225:HIS:ND1	1:B:227:ILE:HG22	1.71	1.05

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	965/1019 (95%)	845 (88%)	105 (11%)	15 (2%)	11	50
1	B	1017/1019 (100%)	901 (89%)	97 (10%)	19 (2%)	9	47
2	C	411/421 (98%)	367 (89%)	38 (9%)	6 (2%)	12	52
2	D	413/421 (98%)	361 (87%)	46 (11%)	6 (2%)	12	52
2	E	419/421 (100%)	368 (88%)	41 (10%)	10 (2%)	7	42
2	F	411/421 (98%)	370 (90%)	34 (8%)	7 (2%)	11	49
2	G	413/421 (98%)	364 (88%)	43 (10%)	6 (2%)	12	52
2	H	419/421 (100%)	379 (90%)	34 (8%)	6 (1%)	13	53

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	I	415/421 (99%)	375 (90%)	34 (8%)	6 (1%)	13	53
2	J	409/421 (97%)	348 (85%)	47 (12%)	14 (3%)	4	35
2	P	410/421 (97%)	361 (88%)	40 (10%)	9 (2%)	8	44
2	Q	419/421 (100%)	365 (87%)	47 (11%)	7 (2%)	11	49
2	R	419/421 (100%)	376 (90%)	37 (9%)	6 (1%)	13	53
2	S	419/421 (100%)	366 (87%)	44 (10%)	9 (2%)	8	45
2	T	419/421 (100%)	369 (88%)	41 (10%)	9 (2%)	8	45
3	K	10/506 (2%)	8 (80%)	1 (10%)	1 (10%)	1	8
All	All	7388/8017 (92%)	6523 (88%)	729 (10%)	136 (2%)	10	48

5 of 136 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	101	ILE
1	A	451	GLY
1	A	546	GLY
1	B	24	SER
1	B	112	SER

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	854/900 (95%)	793 (93%)	61 (7%)	17	54
1	B	900/900 (100%)	844 (94%)	56 (6%)	21	59
2	C	356/360 (99%)	338 (95%)	18 (5%)	28	64
2	D	358/360 (99%)	339 (95%)	19 (5%)	26	63
2	E	360/360 (100%)	340 (94%)	20 (6%)	25	62
2	F	356/360 (99%)	336 (94%)	20 (6%)	25	62
2	G	358/360 (99%)	337 (94%)	21 (6%)	23	61
2	H	360/360 (100%)	341 (95%)	19 (5%)	26	63

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	I	360/360 (100%)	333 (92%)	27 (8%)	16	51
2	J	354/360 (98%)	328 (93%)	26 (7%)	16	53
2	P	355/360 (99%)	335 (94%)	20 (6%)	25	62
2	Q	360/360 (100%)	334 (93%)	26 (7%)	17	53
2	R	360/360 (100%)	344 (96%)	16 (4%)	33	69
2	S	360/360 (100%)	334 (93%)	26 (7%)	17	53
2	T	360/360 (100%)	330 (92%)	30 (8%)	13	46
3	K	12/450 (3%)	10 (83%)	2 (17%)	2	14
All	All	6423/6930 (93%)	6016 (94%)	407 (6%)	21	59

5 of 407 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	Q	164	THR
2	F	185	PHE
2	T	9	THR
2	Q	287	THR
2	E	158	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 160 such sidechains are listed below:

Mol	Chain	Res	Type
2	E	163	GLN
2	F	315	GLN
2	J	208	GLN
2	E	208	GLN
2	F	125	GLN

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 [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

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 ⓘ

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	967/1019 (94%)	-0.35	1 (0%)	95 94	1, 19, 50, 125	0
1	B	1019/1019 (100%)	-0.27	2 (0%)	94 93	1, 15, 49, 115	0
2	C	415/421 (98%)	-0.30	1 (0%)	94 93	44, 75, 99, 104	0
2	D	417/421 (99%)	-0.24	2 (0%)	90 86	18, 68, 106, 112	0
2	E	421/421 (100%)	-0.34	1 (0%)	94 93	3, 38, 71, 86	0
2	F	415/421 (98%)	-0.45	1 (0%)	94 93	6, 42, 76, 82	0
2	G	417/421 (99%)	-0.49	1 (0%)	94 93	1, 26, 59, 112	0
2	H	421/421 (100%)	-0.45	1 (0%)	94 93	1, 26, 59, 89	0
2	I	419/421 (99%)	-0.28	1 (0%)	94 93	12, 49, 82, 87	0
2	J	413/421 (98%)	-0.46	1 (0%)	94 93	18, 51, 77, 93	0
2	P	414/421 (98%)	0.07	3 (0%)	87 82	49, 93, 113, 118	0
2	Q	421/421 (100%)	-0.51	1 (0%)	94 93	11, 48, 76, 96	0
2	R	421/421 (100%)	-0.50	1 (0%)	94 93	1, 27, 58, 70	0
2	S	421/421 (100%)	-0.32	2 (0%)	90 86	6, 41, 72, 88	0
2	T	421/421 (100%)	-0.37	1 (0%)	94 93	1, 15, 44, 74	0
3	K	12/506 (2%)	0.64	1 (8%)	12 12	61, 67, 74, 80	0
All	All	7434/8017 (92%)	-0.34	21 (0%)	93 90	1, 35, 92, 125	0

The worst 5 of 21 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	P	1	MET	3.2
2	G	1	MET	3.2
2	T	1	MET	3.1
1	A	933	THR	3.0
2	D	1	MET	2.8

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](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.