wwPDB X-ray Structure Validation Summary Report (i)

Jun 11, 2024 – 10:40 PM EDT

PDB ID	:	1W8X
Title	:	Structural analysis of PRD1
Authors	:	Abrescia, N.G.A.; Cockburn, J.J.B.; Grimes, J.M.; Sutton, G.C.; Diprose, J.M.; Butcher, S.J.; Fuller, S.D.; San Martin, C.; Burnett, R.M.; Stuart, D.I.; Bamford, D.H.; Bamford, J.K.H.
Deposited on	:	2004-10-01
Resolution	:	4.20 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

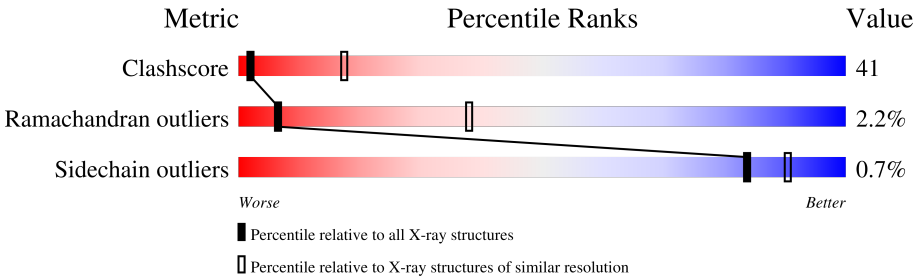
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 4.20 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))
Clashscore	141614	1044 (4.60-3.80)
Ramachandran outliers	138981	1000 (4.60-3.80)
Sidechain outliers	138945	1007 (4.62-3.78)

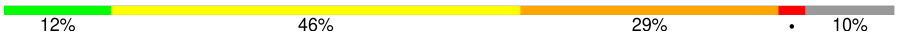
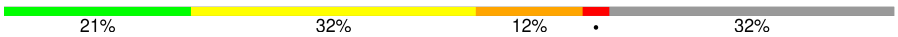
The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. 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 <=5%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	395	78% 13%
1	B	395	72% 21%
1	C	395	79% 15%
1	D	395	73% 20%
1	E	395	77% 16%
1	F	395	79% 17%
1	G	395	75% 16% 5%
1	H	395	76% 19%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	I	395	
1	J	395	
1	K	395	
1	L	395	
2	M	83	
3	N	126	
4	P	117	

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 38116 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 MAJOR CAPSID PROTEIN (PROTEIN P3).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	380	Total	C	N	O	S	0	0	0
			2956	1873	503	573	7			
1	B	392	Total	C	N	O	S	0	0	0
			3045	1929	517	592	7			
1	C	385	Total	C	N	O	S	0	0	0
			2999	1901	510	581	7			
1	D	392	Total	C	N	O	S	0	0	0
			3045	1929	517	592	7			
1	E	378	Total	C	N	O	S	0	0	0
			2944	1868	500	569	7			
1	F	388	Total	C	N	O	S	0	0	0
			3009	1908	510	584	7			
1	G	375	Total	C	N	O	S	0	0	0
			2926	1857	496	566	7			
1	H	392	Total	C	N	O	S	0	0	0
			3045	1929	517	592	7			
1	I	392	Total	C	N	O	S	0	0	0
			3045	1929	517	592	7			
1	J	390	Total	C	N	O	S	0	0	0
			3027	1919	514	587	7			
1	K	384	Total	C	N	O	S	0	0	0
			2992	1897	509	579	7			
1	L	379	Total	C	N	O	S	0	0	0
			2953	1873	502	571	7			

- Molecule 2 is a protein called PROTEIN P30.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	M	83	Total	C	N	O	S	0	0	0
			638	408	113	114	3			

- Molecule 3 is a protein called PROTEIN P31.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	N	114	Total	C	N	O	S	0	0	0
			872	552	147	169	4			

- Molecule 4 is a protein called PROTEIN P16.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	P	80	Total	C	N	O	S	0	0	1
			620	401	103	114	2			

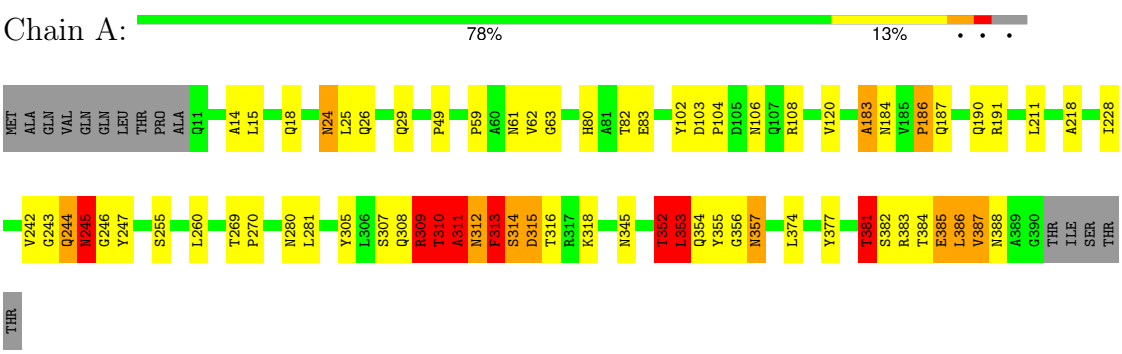
3 Residue-property plots

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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. 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. 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.

Note EDS was not executed.

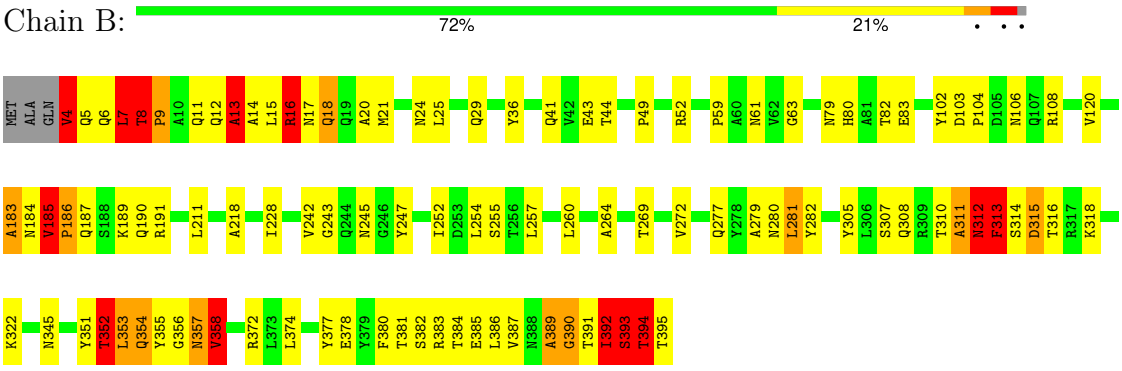
- Molecule 1: MAJOR CAPSID PROTEIN (PROTEIN P3)

Chain A:



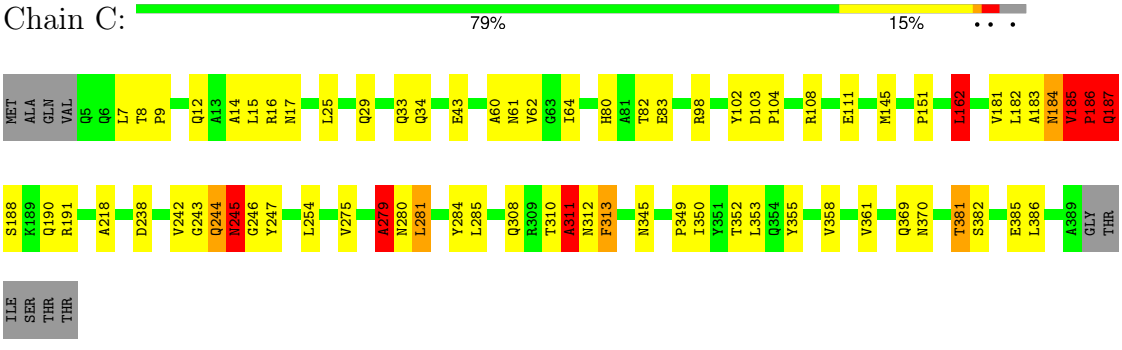
- Molecule 1: MAJOR CAPSID PROTEIN (PROTEIN P3)

Chain B:



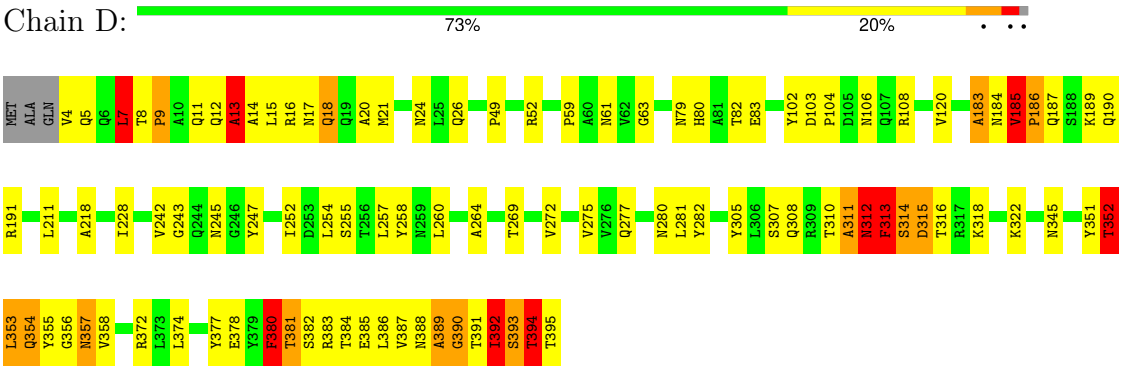
- Molecule 1: MAJOR CAPSID PROTEIN (PROTEIN P3)

Chain C:



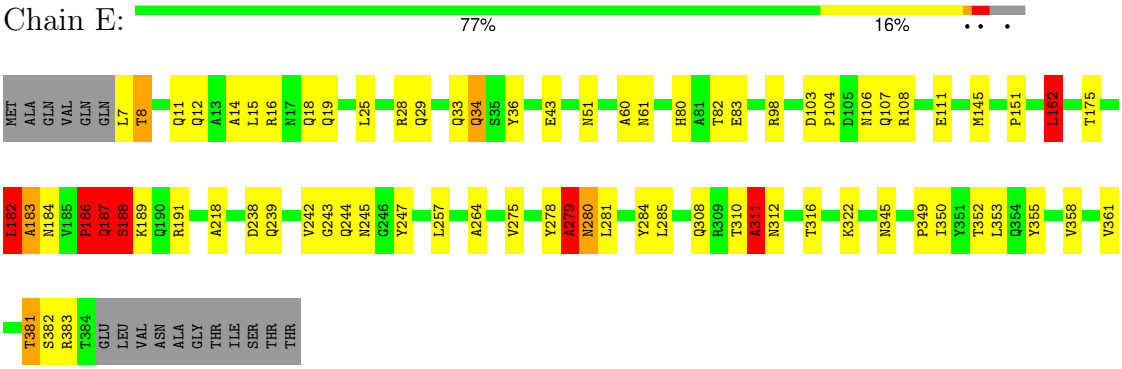
• Molecule 1: MAJOR CAPSID PROTEIN (PROTEIN P3)

Chain D:



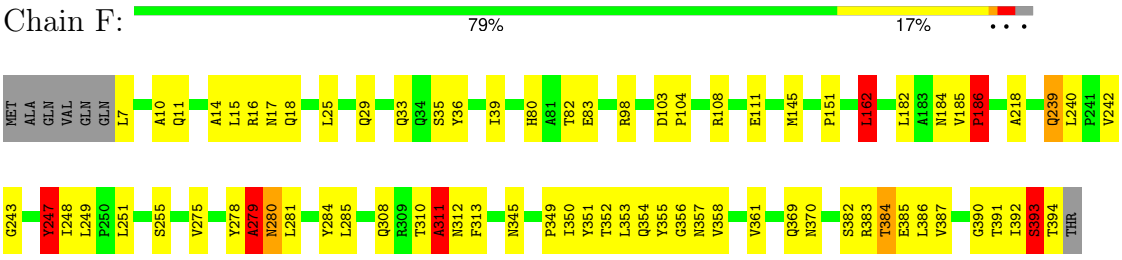
• Molecule 1: MAJOR CAPSID PROTEIN (PROTEIN P3)

Chain E:



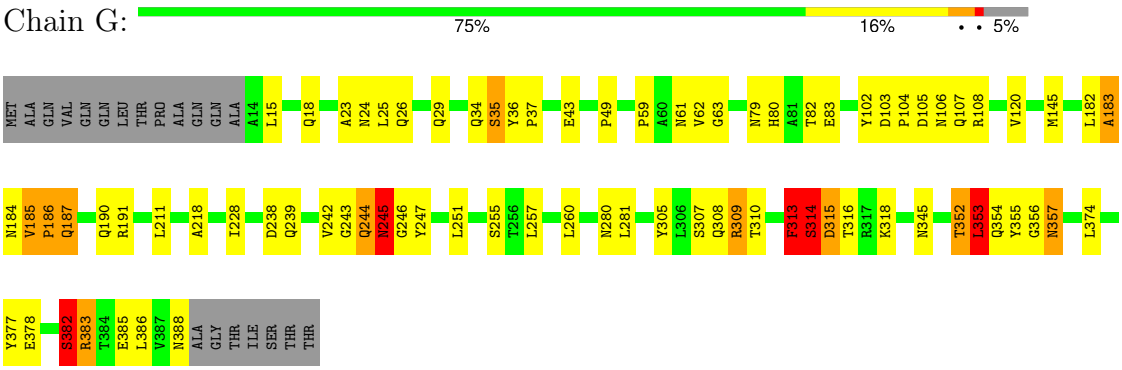
• Molecule 1: MAJOR CAPSID PROTEIN (PROTEIN P3)

Chain F:



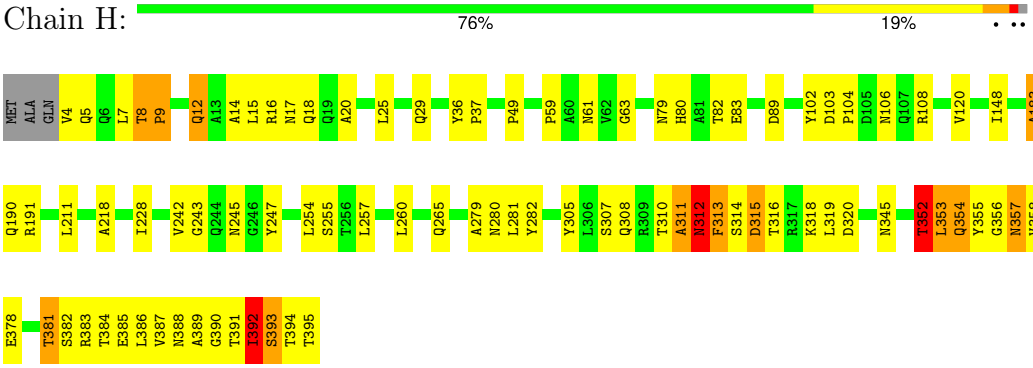
• Molecule 1: MAJOR CAPSID PROTEIN (PROTEIN P3)

Chain G:



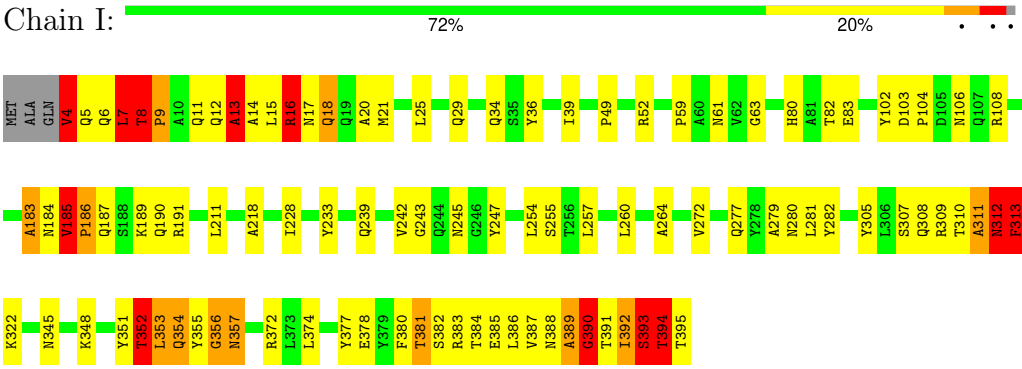
• Molecule 1: MAJOR CAPSID PROTEIN (PROTEIN P3)

Chain H:



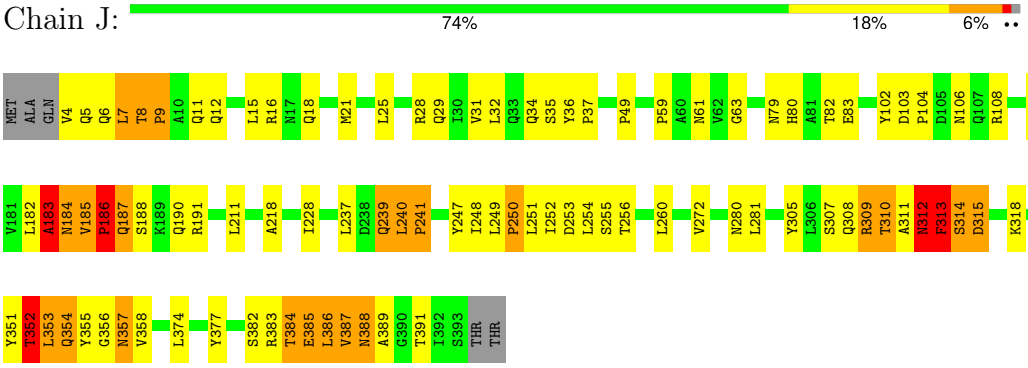
• Molecule 1: MAJOR CAPSID PROTEIN (PROTEIN P3)

Chain I:



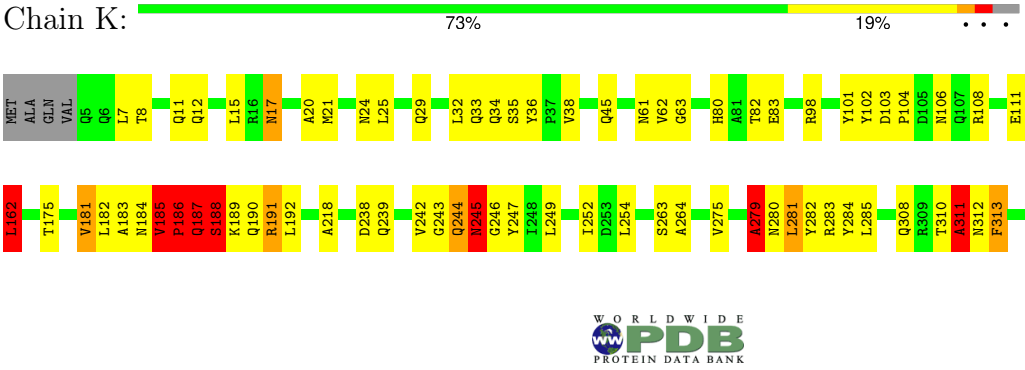
• Molecule 1: MAJOR CAPSID PROTEIN (PROTEIN P3)

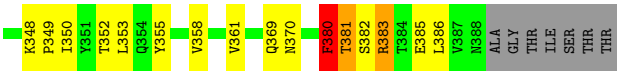
Chain J:



• Molecule 1: MAJOR CAPSID PROTEIN (PROTEIN P3)

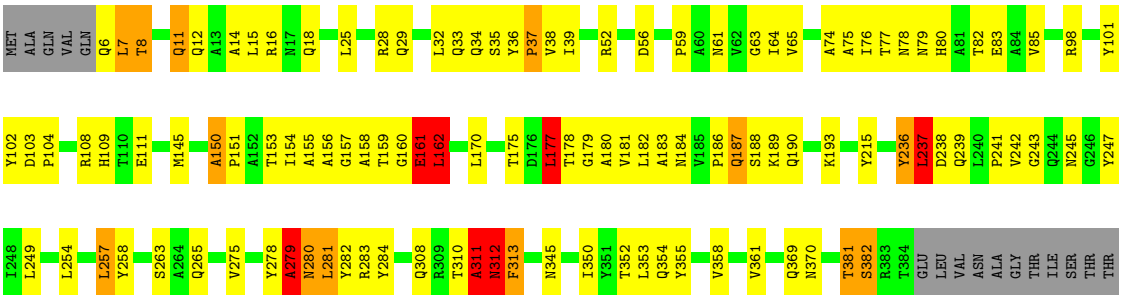
Chain K:





- Molecule 1: MAJOR CAPSID PROTEIN (PROTEIN P3)

Chain L:

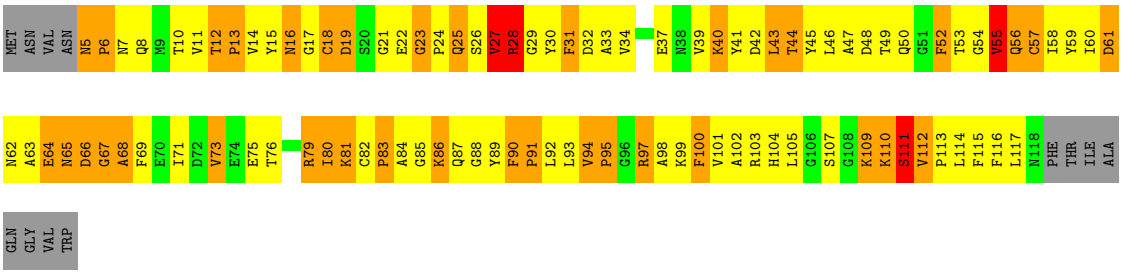


- Molecule 2: PROTEIN P30

Chain M:

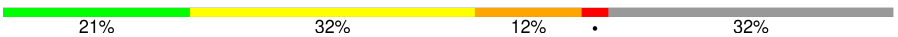
- Molecule 3: PROTEIN P31

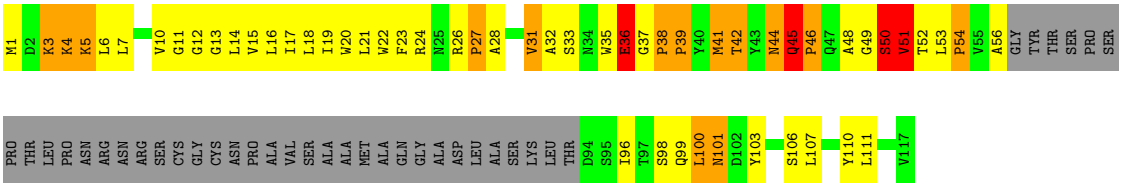
Chain N:



- Molecule 4: PROTEIN P16

Chain P:





4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	903.00Å 920.60Å 926.20Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	100.00 – 4.20	Depositor
% Data completeness (in resolution range)	(Not available) (100.00-4.20)	Depositor
R_{merge}	0.33	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	X-PLOR	Depositor
R, R_{free}	(Not available) , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	38116	wwPDB-VP
Average B, all atoms (Å^2)	20.0	wwPDB-VP

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	1.32	17/3021 (0.6%)	1.47	28/4128 (0.7%)
1	B	1.71	32/3110 (1.0%)	1.89	50/4250 (1.2%)
1	C	1.08	12/3064 (0.4%)	1.15	26/4186 (0.6%)
1	D	1.38	26/3111 (0.8%)	1.63	37/4253 (0.9%)
1	E	0.77	9/3011 (0.3%)	1.18	20/4117 (0.5%)
1	F	1.34	15/3074 (0.5%)	1.56	22/4201 (0.5%)
1	G	1.16	15/2991 (0.5%)	1.33	23/4087 (0.6%)
1	H	1.32	21/3112 (0.7%)	1.23	25/4256 (0.6%)
1	I	1.55	30/3110 (1.0%)	1.87	48/4250 (1.1%)
1	J	1.25	32/3092 (1.0%)	1.13	20/4225 (0.5%)
1	K	1.19	15/3058 (0.5%)	1.49	37/4179 (0.9%)
1	L	1.25	10/3018 (0.3%)	1.17	28/4123 (0.7%)
2	M	2.88	42/657 (6.4%)	2.16	27/898 (3.0%)
3	N	3.68	54/892 (6.1%)	2.82	46/1209 (3.8%)
4	P	2.89	37/637 (5.8%)	1.83	19/871 (2.2%)
All	All	1.48	367/38958 (0.9%)	1.52	456/53233 (0.9%)

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	7
1	B	0	13
1	C	0	7
1	D	0	10
1	E	0	5
1	F	0	6
1	G	0	5
1	H	0	5
1	I	0	10
1	J	0	5

Continued on next page...

Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
1	K	0	10
1	L	0	5
3	N	0	3
All	All	0	91

The worst 5 of 367 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	N	111	SER	C-O	-43.65	0.40	1.23
1	F	391	THR	C-O	-43.51	0.40	1.23
3	N	27	VAL	C-O	-41.51	0.44	1.23
3	N	26	SER	C-O	-36.70	0.53	1.23
1	A	313	PHE	C-N	-35.91	0.51	1.34

The worst 5 of 456 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	24	ASN	O-C-N	-53.44	37.20	122.70
1	F	247	TYR	O-C-N	-45.14	50.48	122.70
1	G	313	PHE	O-C-N	-44.62	51.31	122.70
1	B	16	ARG	O-C-N	-42.15	55.26	122.70
1	I	16	ARG	O-C-N	-42.14	55.28	122.70

There are no chirality outliers.

5 of 91 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	245	ASN	Mainchain
1	A	309	ARG	Mainchain
1	A	311	ALA	Mainchain
1	A	313	PHE	Mainchain
1	A	352	THR	Peptide

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	2956	0	2851	116	0
1	B	3045	0	2937	307	0
1	C	2999	0	2903	129	0
1	D	3045	0	2939	334	0
1	E	2944	0	2846	212	0
1	F	3009	0	2912	268	0
1	G	2926	0	2832	196	0
1	H	3045	0	2951	197	0
1	I	3045	0	2931	337	0
1	J	3027	0	2927	210	0
1	K	2992	0	2880	234	0
1	L	2953	0	2854	329	0
2	M	638	0	620	405	0
3	N	872	0	814	459	0
4	P	620	0	598	161	0
All	All	38116	0	36795	3092	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 41.

The worst 5 of 3092 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:J:16:ARG:HD3	1:J:21:MET:SD	1.29	1.69
1:L:108:ARG:HH11	3:N:89:TYR:CB	0.99	1.62
1:A:255:SER:HA	1:A:383:ARG:CD	1.30	1.60
1:E:189:LYS:HZ2	1:K:61:ASN:CB	1.05	1.60
3:N:60:ILE:HD12	3:N:97:ARG:CD	1.23	1.59

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	378/395 (96%)	358 (95%)	14 (4%)	6 (2%)	9 45
1	B	386/395 (98%)	366 (95%)	11 (3%)	9 (2%)	6 37
1	C	379/395 (96%)	356 (94%)	18 (5%)	5 (1%)	12 48
1	D	388/395 (98%)	368 (95%)	13 (3%)	7 (2%)	8 42
1	E	376/395 (95%)	354 (94%)	17 (4%)	5 (1%)	12 48
1	F	382/395 (97%)	359 (94%)	18 (5%)	5 (1%)	12 48
1	G	371/395 (94%)	354 (95%)	12 (3%)	5 (1%)	12 48
1	H	390/395 (99%)	368 (94%)	15 (4%)	7 (2%)	8 42
1	I	386/395 (98%)	367 (95%)	11 (3%)	8 (2%)	7 39
1	J	384/395 (97%)	364 (95%)	13 (3%)	7 (2%)	8 42
1	K	380/395 (96%)	350 (92%)	25 (7%)	5 (1%)	12 48
1	L	373/395 (94%)	348 (93%)	17 (5%)	8 (2%)	7 39
2	M	81/83 (98%)	65 (80%)	8 (10%)	8 (10%)	0 10
3	N	112/126 (89%)	94 (84%)	6 (5%)	12 (11%)	0 8
4	P	76/117 (65%)	57 (75%)	11 (14%)	8 (10%)	0 8
All	All	4842/5066 (96%)	4528 (94%)	209 (4%)	105 (2%)	6 38

5 of 105 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	186	PRO
1	A	244	GLN
1	A	353	LEU
1	B	8	THR
1	B	186	PRO

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	312/326 (96%)	310 (99%)	2 (1%)	86 92

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	B	324/326 (99%)	322 (99%)	2 (1%)	86 92
1	C	318/326 (98%)	315 (99%)	3 (1%)	78 87
1	D	324/326 (99%)	323 (100%)	1 (0%)	92 95
1	E	312/326 (96%)	308 (99%)	4 (1%)	69 82
1	F	319/326 (98%)	317 (99%)	2 (1%)	86 92
1	G	311/326 (95%)	310 (100%)	1 (0%)	92 95
1	H	324/326 (99%)	323 (100%)	1 (0%)	92 95
1	I	324/326 (99%)	322 (99%)	2 (1%)	86 92
1	J	321/326 (98%)	318 (99%)	3 (1%)	78 87
1	K	317/326 (97%)	312 (98%)	5 (2%)	62 79
1	L	313/326 (96%)	309 (99%)	4 (1%)	69 82
2	M	66/66 (100%)	66 (100%)	0	100 100
3	N	92/102 (90%)	92 (100%)	0	100 100
4	P	65/94 (69%)	65 (100%)	0	100 100
All	All	4042/4174 (97%)	4012 (99%)	30 (1%)	84 90

5 of 30 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	315	ASP
1	L	162	LEU
1	I	315	ASP
1	L	355	TYR
1	K	188	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 123 such sidechains are listed below:

Mol	Chain	Res	Type
1	G	61	ASN
1	L	41	GLN
1	H	79	ASN
1	K	354	GLN
2	M	58	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

There are no monosaccharides in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	I	15
1	D	14
1	B	14
1	J	11
1	F	9
1	H	9
3	N	9
1	L	8
1	G	7
1	K	7
1	C	6
1	A	6
2	M	5
4	P	4
1	E	3

The worst 5 of 127 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	J	383:ARG	C	384:THR	N	3.74
1	C	190:GLN	C	191:ARG	N	3.53
1	D	17:ASN	C	18:GLN	N	3.14
1	I	17:ASN	C	18:GLN	N	3.14
1	L	35:SER	C	36:TYR	N	2.61

6 Fit of model and data

6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

6.4 Ligands

EDS was not executed - this section is therefore empty.

6.5 Other polymers

EDS was not executed - this section is therefore empty.