



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 10:13 PM GMT

PDB ID : 4WYM  
Title : Structural basis of HIV-1 capsid recognition by CPSF6  
Authors : Battacharya, A.; Taylor, A.B.; Hart, P.J.; Ivanov, D.N.  
Deposited on : 2014-11-17  
Resolution : 2.60 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

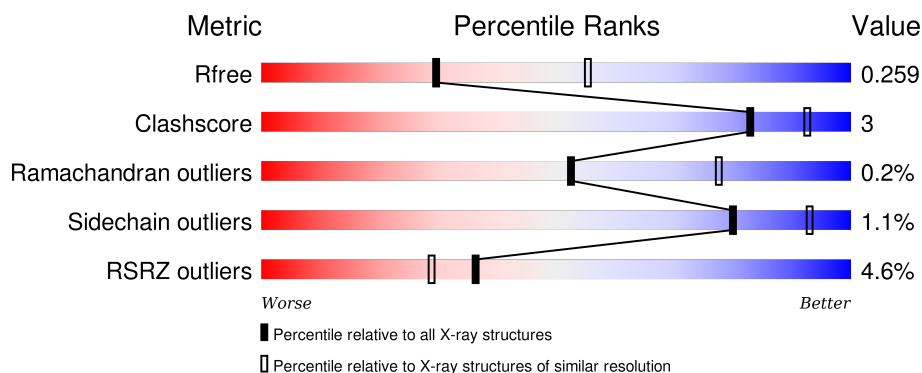
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.60 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	2328 (2.60-2.60)
Clashscore	102246	2679 (2.60-2.60)
Ramachandran outliers	100387	2635 (2.60-2.60)
Sidechain outliers	100360	2635 (2.60-2.60)
RSRZ outliers	91569	2334 (2.60-2.60)

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	231	<div> <div>3%</div> <div> <div></div> <div>84%</div> <div>8%</div> <div>7%</div> </div> </div>
1	B	231	<div> <div>5%</div> <div> <div></div> <div>84%</div> <div>8%</div> <div>8%</div> </div> </div>
1	C	231	<div> <div>%</div> <div> <div></div> <div>88%</div> <div>•</div> <div>8%</div> </div> </div>
1	D	231	<div> <div>4%</div> <div> <div></div> <div>85%</div> <div>•</div> <div>12%</div> </div> </div>
1	E	231	<div> <div>3%</div> <div> <div></div> <div>88%</div> <div>5%</div> <div>7%</div> </div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	231	
1	G	231	
1	H	231	
1	I	231	
1	J	231	
1	K	231	
1	L	231	
2	M	17	
2	N	17	
2	O	17	
2	P	17	
2	Q	17	
2	R	17	
2	S	17	
2	T	17	
2	U	17	
2	V	17	
2	W	17	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 20913 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 Capsid protein p24.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	214	Total	C	N	O	S	0	0	0
			1659	1044	291	310	14			
1	B	213	Total	C	N	O	S	0	0	0
			1658	1041	290	313	14			
1	C	212	Total	C	N	O	S	0	0	0
			1654	1039	289	312	14			
1	D	204	Total	C	N	O	S	0	0	0
			1597	1006	278	299	14			
1	E	215	Total	C	N	O	S	0	0	0
			1671	1050	292	315	14			
1	F	220	Total	C	N	O	S	0	0	0
			1704	1071	299	320	14			
1	G	211	Total	C	N	O	S	0	0	0
			1642	1035	287	306	14			
1	H	211	Total	C	N	O	S	0	0	0
			1642	1035	287	306	14			
1	I	212	Total	C	N	O	S	0	0	0
			1647	1038	288	307	14			
1	J	213	Total	C	N	O	S	0	0	0
			1661	1044	290	313	14			
1	K	210	Total	C	N	O	S	0	0	0
			1638	1032	286	306	14			
1	L	208	Total	C	N	O	S	0	0	0
			1630	1026	284	306	14			

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	14	CYS	ALA	engineered mutation	UNP P12497
A	45	CYS	GLU	engineered mutation	UNP P12497
A	184	ALA	TRP	engineered mutation	UNP P12497
A	185	ALA	MET	engineered mutation	UNP P12497
B	14	CYS	ALA	engineered mutation	UNP P12497

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	45	CYS	GLU	engineered mutation	UNP P12497
B	184	ALA	TRP	engineered mutation	UNP P12497
B	185	ALA	MET	engineered mutation	UNP P12497
C	14	CYS	ALA	engineered mutation	UNP P12497
C	45	CYS	GLU	engineered mutation	UNP P12497
C	184	ALA	TRP	engineered mutation	UNP P12497
C	185	ALA	MET	engineered mutation	UNP P12497
D	14	CYS	ALA	engineered mutation	UNP P12497
D	45	CYS	GLU	engineered mutation	UNP P12497
D	184	ALA	TRP	engineered mutation	UNP P12497
D	185	ALA	MET	engineered mutation	UNP P12497
E	14	CYS	ALA	engineered mutation	UNP P12497
E	45	CYS	GLU	engineered mutation	UNP P12497
E	184	ALA	TRP	engineered mutation	UNP P12497
E	185	ALA	MET	engineered mutation	UNP P12497
F	14	CYS	ALA	engineered mutation	UNP P12497
F	45	CYS	GLU	engineered mutation	UNP P12497
F	184	ALA	TRP	engineered mutation	UNP P12497
F	185	ALA	MET	engineered mutation	UNP P12497
G	14	CYS	ALA	engineered mutation	UNP P12497
G	45	CYS	GLU	engineered mutation	UNP P12497
G	184	ALA	TRP	engineered mutation	UNP P12497
G	185	ALA	MET	engineered mutation	UNP P12497
H	14	CYS	ALA	engineered mutation	UNP P12497
H	45	CYS	GLU	engineered mutation	UNP P12497
H	184	ALA	TRP	engineered mutation	UNP P12497
H	185	ALA	MET	engineered mutation	UNP P12497
I	14	CYS	ALA	engineered mutation	UNP P12497
I	45	CYS	GLU	engineered mutation	UNP P12497
I	184	ALA	TRP	engineered mutation	UNP P12497
I	185	ALA	MET	engineered mutation	UNP P12497
J	14	CYS	ALA	engineered mutation	UNP P12497
J	45	CYS	GLU	engineered mutation	UNP P12497
J	184	ALA	TRP	engineered mutation	UNP P12497
J	185	ALA	MET	engineered mutation	UNP P12497
K	14	CYS	ALA	engineered mutation	UNP P12497
K	45	CYS	GLU	engineered mutation	UNP P12497
K	184	ALA	TRP	engineered mutation	UNP P12497
K	185	ALA	MET	engineered mutation	UNP P12497
L	14	CYS	ALA	engineered mutation	UNP P12497
L	45	CYS	GLU	engineered mutation	UNP P12497
L	184	ALA	TRP	engineered mutation	UNP P12497

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
L	185	ALA	MET	engineered mutation	UNP P12497

- Molecule 2 is a protein called ISOFORM 2 OF CLEAVAGE AND POLYADENYLATION SPECIFICITY FACTOR SUBUNIT 6.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	M	13	Total	C	N	O	0	0	0
			98	68	15	15			
2	N	14	Total	C	N	O	0	0	0
			105	72	16	17			
2	O	13	Total	C	N	O	0	0	0
			98	68	15	15			
2	P	14	Total	C	N	O	0	0	0
			105	72	16	17			
2	Q	13	Total	C	N	O	0	0	0
			98	68	15	15			
2	R	14	Total	C	N	O	0	0	0
			105	72	16	17			
2	S	15	Total	C	N	O	0	0	0
			109	74	17	18			
2	T	13	Total	C	N	O	0	0	0
			98	68	15	15			
2	U	13	Total	C	N	O	0	0	0
			98	68	15	15			
2	V	13	Total	C	N	O	0	0	0
			98	68	15	15			
2	W	13	Total	C	N	O	0	0	0
			98	68	15	15			

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
M	311	GLY	-	expression tag	UNP Q16630
M	312	THR	-	expression tag	UNP Q16630
N	311	GLY	-	expression tag	UNP Q16630
N	312	THR	-	expression tag	UNP Q16630
O	311	GLY	-	expression tag	UNP Q16630
O	312	THR	-	expression tag	UNP Q16630
P	311	GLY	-	expression tag	UNP Q16630
P	312	THR	-	expression tag	UNP Q16630
Q	311	GLY	-	expression tag	UNP Q16630
Q	312	THR	-	expression tag	UNP Q16630
R	311	GLY	-	expression tag	UNP Q16630

*Continued on next page...*

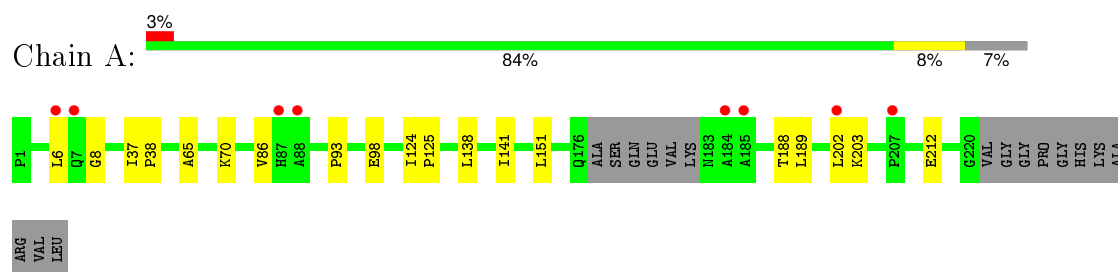
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
R	312	THR	-	expression tag	UNP Q16630
S	311	GLY	-	expression tag	UNP Q16630
S	312	THR	-	expression tag	UNP Q16630
T	311	GLY	-	expression tag	UNP Q16630
T	312	THR	-	expression tag	UNP Q16630
U	311	GLY	-	expression tag	UNP Q16630
U	312	THR	-	expression tag	UNP Q16630
V	311	GLY	-	expression tag	UNP Q16630
V	312	THR	-	expression tag	UNP Q16630
W	311	GLY	-	expression tag	UNP Q16630
W	312	THR	-	expression tag	UNP Q16630

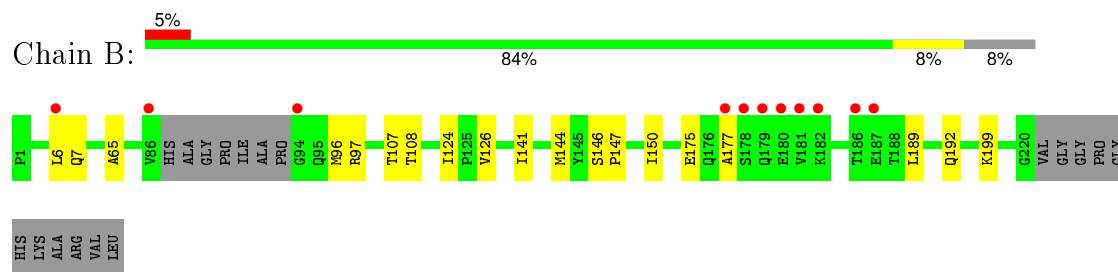
### 3 Residue-property plots [i](#)

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.

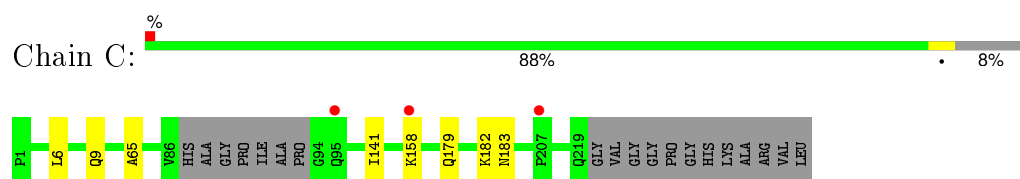
- Molecule 1: Capsid protein p24



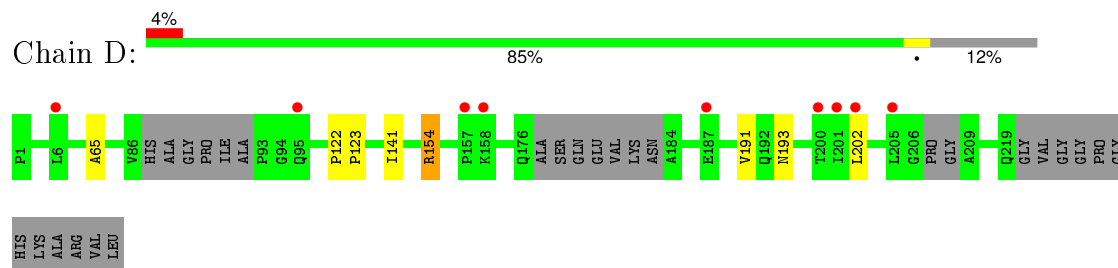
- Molecule 1: Capsid protein p24



- Molecule 1: Capsid protein p24

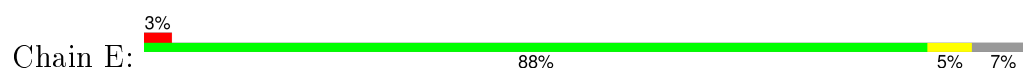


- Molecule 1: Capsid protein p24

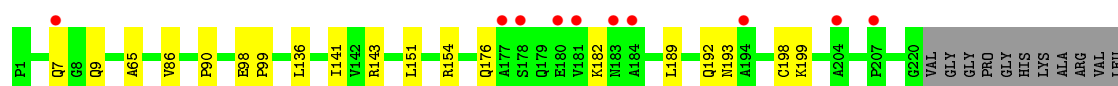
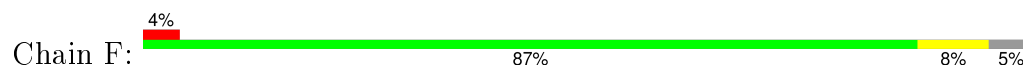


- Molecule 1: Capsid protein p24

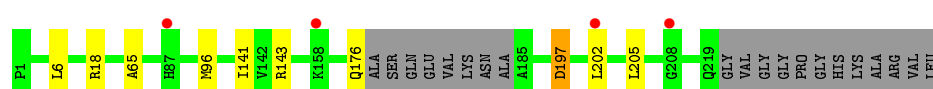
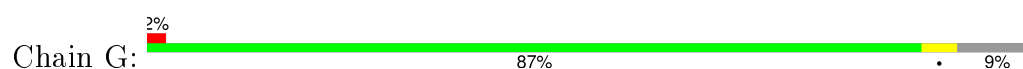




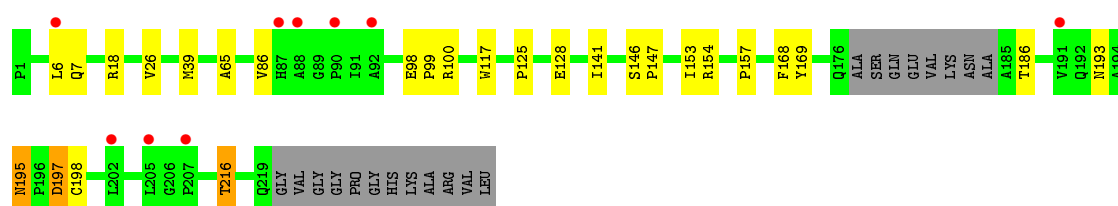
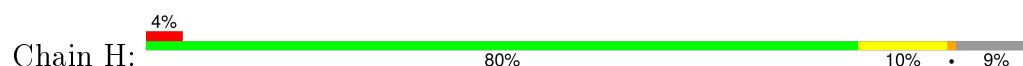
- Molecule 1: Capsid protein p24



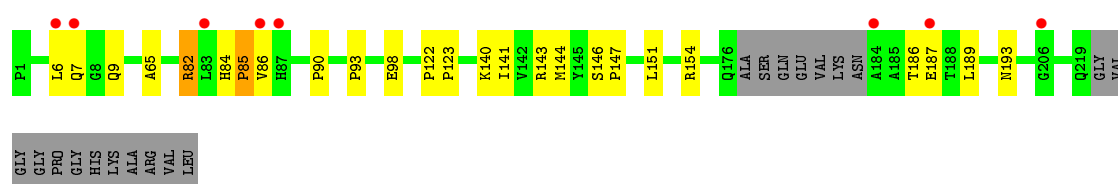
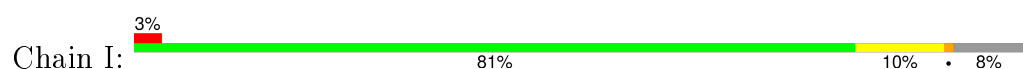
- Molecule 1: Capsid protein p24



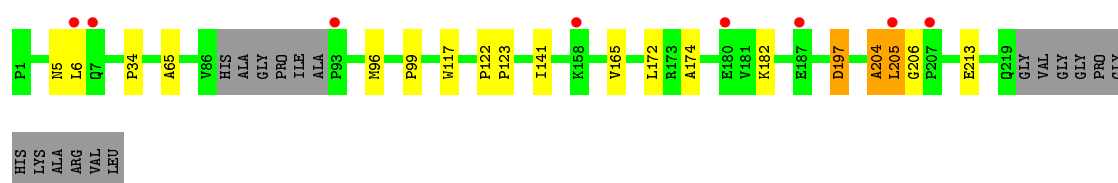
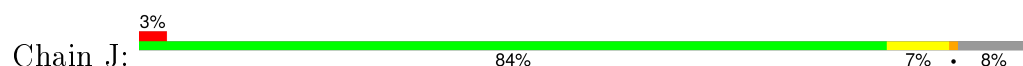
- Molecule 1: Capsid protein p24



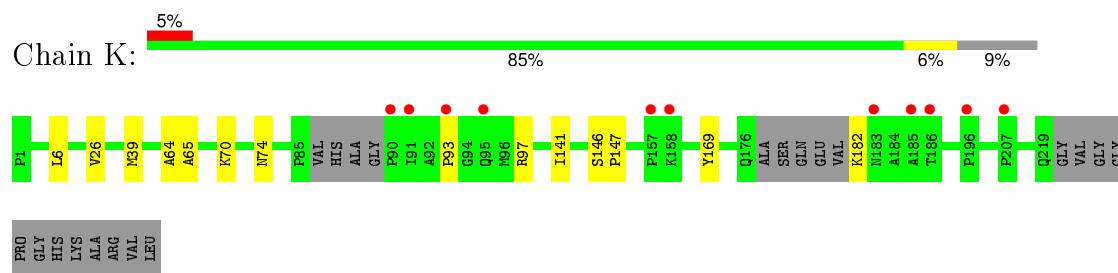
- Molecule 1: Capsid protein p24



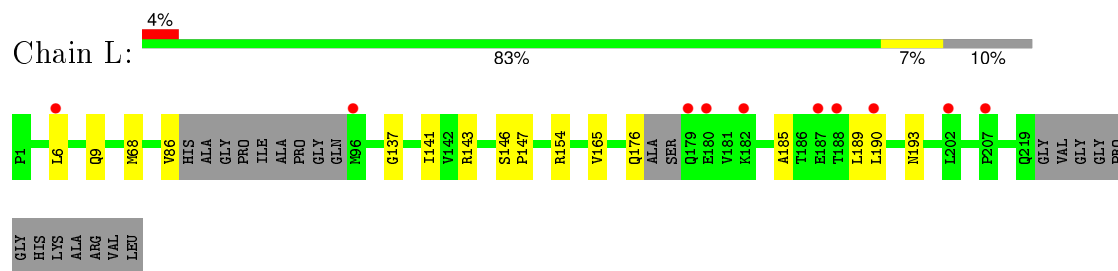
- Molecule 1: Capsid protein p24



- Molecule 1: Capsid protein p24



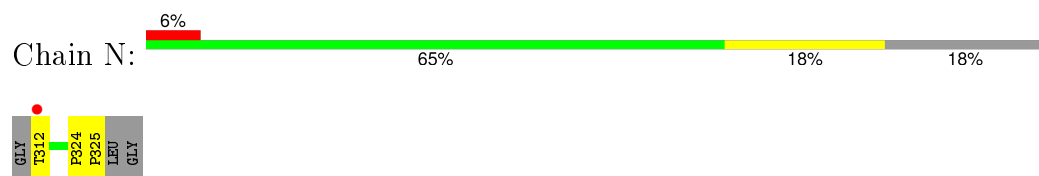
- Molecule 1: Capsid protein p24



- Molecule 2: ISOFORM 2 OF CLEAVAGE AND POLYADENYLATION SPECIFICITY FACTOR SUBUNIT 6



- Molecule 2: ISOFORM 2 OF CLEAVAGE AND POLYADENYLATION SPECIFICITY FACTOR SUBUNIT 6

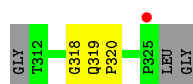


- Molecule 2: ISOFORM 2 OF CLEAVAGE AND POLYADENYLATION SPECIFICITY FACTOR SUBUNIT 6

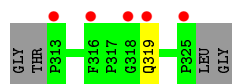


- Molecule 2: ISOFORM 2 OF CLEAVAGE AND POLYADENYLATION SPECIFICITY FACTOR SUBUNIT 6





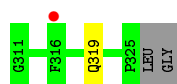
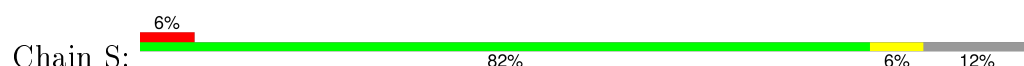
- Molecule 2: ISOFORM 2 OF CLEAVAGE AND POLYADENYLATION SPECIFICITY FACTOR SUBUNIT 6



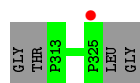
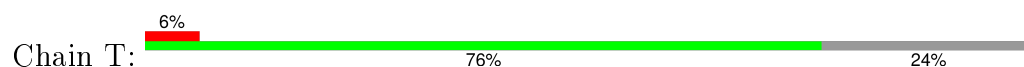
- Molecule 2: ISOFORM 2 OF CLEAVAGE AND POLYADENYLATION SPECIFICITY FACTOR SUBUNIT 6



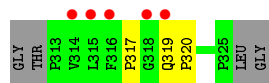
- Molecule 2: ISOFORM 2 OF CLEAVAGE AND POLYADENYLATION SPECIFICITY FACTOR SUBUNIT 6



- Molecule 2: ISOFORM 2 OF CLEAVAGE AND POLYADENYLATION SPECIFICITY FACTOR SUBUNIT 6

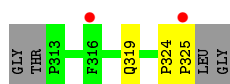


- Molecule 2: ISOFORM 2 OF CLEAVAGE AND POLYADENYLATION SPECIFICITY FACTOR SUBUNIT 6

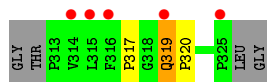


- Molecule 2: ISOFORM 2 OF CLEAVAGE AND POLYADENYLATION SPECIFICITY FACTOR SUBUNIT 6





- Molecule 2: ISOFORM 2 OF CLEAVAGE AND POLYADENYLATION SPECIFICITY FACTOR SUBUNIT 6



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	134.79Å 136.03Å 207.26Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.47 – 2.60 48.47 – 2.60	Depositor EDS
% Data completeness (in resolution range)	100.0 (48.47-2.60) 99.9 (48.47-2.60)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.07	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.15 (at 2.61Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, $R_{free}$	0.219 , 0.259 0.219 , 0.259	Depositor DCC
$R_{free}$ test set	2008 reflections (1.71%)	DCC
Wilson B-factor (Å <sup>2</sup> )	62.7	Xtriage
Anisotropy	0.229	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 46.4	EDS
Estimated twinning fraction	0.027 for k,h,l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	1 of 117427 reflections (0.001%)	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	20913	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	73.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.86% 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](#)

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	0.21	0/1695	0.38	0/2303
1	B	0.21	0/1691	0.37	0/2294
1	C	0.22	0/1687	0.36	0/2289
1	D	0.20	0/1628	0.37	0/2206
1	E	0.21	0/1705	0.37	0/2314
1	F	0.21	0/1741	0.36	0/2366
1	G	0.20	0/1678	0.37	0/2280
1	H	0.21	0/1678	0.37	0/2280
1	I	0.27	1/1683 (0.1%)	0.40	1/2287 (0.0%)
1	J	0.21	0/1695	0.39	0/2300
1	K	0.21	0/1672	0.36	0/2268
1	L	0.22	0/1662	0.36	0/2254
2	M	0.21	0/104	0.39	0/143
2	N	0.25	0/111	0.51	0/154
2	O	0.22	0/104	0.39	0/143
2	P	0.26	0/111	0.52	0/154
2	Q	0.22	0/104	0.37	0/143
2	R	0.23	0/111	0.48	0/154
2	S	0.23	0/115	0.43	0/159
2	T	0.23	0/104	0.41	0/143
2	U	0.22	0/104	0.39	0/143
2	V	0.21	0/104	0.39	0/143
2	W	0.22	0/104	0.35	0/143
All	All	0.22	1/21391 (0.0%)	0.37	1/29063 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	I	85	PRO	N-CD	5.61	1.55	1.47

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	I	84	HIS	C-N-CD	5.41	139.75	128.40

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	1659	0	1659	13	0
1	B	1658	0	1660	13	0
1	C	1654	0	1657	6	0
1	D	1597	0	1601	4	0
1	E	1671	0	1674	9	0
1	F	1704	0	1704	11	0
1	G	1642	0	1643	6	0
1	H	1642	0	1643	19	0
1	I	1647	0	1650	20	0
1	J	1661	0	1665	20	0
1	K	1638	0	1645	9	0
1	L	1630	0	1633	10	0
2	M	98	0	95	3	0
2	N	105	0	101	1	0
2	O	98	0	95	2	0
2	P	105	0	101	3	0
2	Q	98	0	95	0	0
2	R	105	0	101	2	0
2	S	109	0	104	0	0
2	T	98	0	95	0	0
2	U	98	0	95	2	0
2	V	98	0	95	1	0
2	W	98	0	95	2	0
All	All	20913	0	20906	129	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 3.

The worst 5 of 129 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:L:154:ARG:HG2	1:L:193:ASN:HB3	1.67	0.77
1:C:179:GLN:O	1:C:179:GLN:NE2	2.25	0.69
1:H:6:LEU:HD13	1:I:6:LEU:HD12	1.74	0.69
1:E:91:ILE:HG22	1:E:92:ALA:N	2.09	0.68
1:H:154:ARG:HG2	1:H:193:ASN:HB3	1.77	0.66

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	210/231 (91%)	204 (97%)	6 (3%)	0	100	100
1	B	209/231 (90%)	203 (97%)	5 (2%)	1 (0%)	34	60
1	C	208/231 (90%)	204 (98%)	4 (2%)	0	100	100
1	D	196/231 (85%)	192 (98%)	4 (2%)	0	100	100
1	E	211/231 (91%)	206 (98%)	5 (2%)	0	100	100
1	F	218/231 (94%)	213 (98%)	5 (2%)	0	100	100
1	G	207/231 (90%)	204 (99%)	3 (1%)	0	100	100
1	H	207/231 (90%)	204 (99%)	3 (1%)	0	100	100
1	I	208/231 (90%)	206 (99%)	2 (1%)	0	100	100
1	J	209/231 (90%)	194 (93%)	13 (6%)	2 (1%)	19	39
1	K	204/231 (88%)	197 (97%)	6 (3%)	1 (0%)	34	60
1	L	202/231 (87%)	198 (98%)	4 (2%)	0	100	100
2	M	11/17 (65%)	11 (100%)	0	0	100	100
2	N	12/17 (71%)	12 (100%)	0	0	100	100
2	O	11/17 (65%)	11 (100%)	0	0	100	100
2	P	12/17 (71%)	12 (100%)	0	0	100	100

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	Q	11/17 (65%)	11 (100%)	0	0	100	100
2	R	12/17 (71%)	12 (100%)	0	0	100	100
2	S	13/17 (76%)	13 (100%)	0	0	100	100
2	T	11/17 (65%)	11 (100%)	0	0	100	100
2	U	11/17 (65%)	11 (100%)	0	0	100	100
2	V	11/17 (65%)	11 (100%)	0	0	100	100
2	W	11/17 (65%)	11 (100%)	0	0	100	100
All	All	2615/2959 (88%)	2551 (98%)	60 (2%)	4 (0%)	52	77

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	J	204	ALA
1	B	177	ALA
1	J	205	LEU
1	K	93	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	181/193 (94%)	180 (99%)	1 (1%)	90	97
1	B	182/193 (94%)	181 (100%)	1 (0%)	92	98
1	C	182/193 (94%)	181 (100%)	1 (0%)	92	98
1	D	176/193 (91%)	175 (99%)	1 (1%)	90	97
1	E	183/193 (95%)	183 (100%)	0	100	100
1	F	186/193 (96%)	183 (98%)	3 (2%)	70	89
1	G	180/193 (93%)	178 (99%)	2 (1%)	80	93
1	H	180/193 (93%)	175 (97%)	5 (3%)	51	78
1	I	180/193 (93%)	179 (99%)	1 (1%)	90	97

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	J	183/193 (95%)	182 (100%)	1 (0%)	92	98
1	K	180/193 (93%)	180 (100%)	0	100	100
1	L	180/193 (93%)	178 (99%)	2 (1%)	80	93
2	M	11/13 (85%)	11 (100%)	0	100	100
2	N	12/13 (92%)	11 (92%)	1 (8%)	14	27
2	O	11/13 (85%)	10 (91%)	1 (9%)	12	22
2	P	12/13 (92%)	12 (100%)	0	100	100
2	Q	11/13 (85%)	10 (91%)	1 (9%)	12	22
2	R	12/13 (92%)	11 (92%)	1 (8%)	14	27
2	S	12/13 (92%)	11 (92%)	1 (8%)	14	27
2	T	11/13 (85%)	11 (100%)	0	100	100
2	U	11/13 (85%)	11 (100%)	0	100	100
2	V	11/13 (85%)	10 (91%)	1 (9%)	12	22
2	W	11/13 (85%)	10 (91%)	1 (9%)	12	22
All	All	2298/2459 (94%)	2273 (99%)	25 (1%)	80	93

5 of 25 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	H	197	ASP
1	H	216	THR
2	V	319	GLN
1	H	198	CYS
1	I	82	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	G	50	GLN
2	U	319	GLN
1	I	9	GLN
1	C	112	GLN
2	S	319	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 carbohydrates 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 ⓘ

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, 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	214/231 (92%)	0.19	8 (3%) 45 37	39, 63, 113, 131	0
1	B	213/231 (92%)	0.10	11 (5%) 31 24	42, 66, 109, 141	0
1	C	212/231 (91%)	0.06	3 (1%) 78 74	44, 63, 96, 112	0
1	D	204/231 (88%)	0.27	9 (4%) 38 30	40, 66, 118, 133	0
1	E	215/231 (93%)	0.22	8 (3%) 45 37	41, 66, 102, 139	0
1	F	220/231 (95%)	0.32	10 (4%) 37 29	42, 66, 115, 159	0
1	G	211/231 (91%)	0.18	4 (1%) 70 64	47, 71, 109, 119	0
1	H	211/231 (91%)	0.34	9 (4%) 39 31	48, 79, 117, 135	0
1	I	212/231 (91%)	0.15	8 (3%) 44 36	47, 72, 106, 125	0
1	J	213/231 (92%)	0.24	8 (3%) 44 36	45, 64, 113, 130	0
1	K	210/231 (90%)	0.27	11 (5%) 31 24	43, 73, 115, 128	0
1	L	208/231 (90%)	0.31	10 (4%) 34 27	45, 70, 111, 143	0
2	M	13/17 (76%)	0.75	1 (7%) 16 11	60, 84, 99, 100	0
2	N	14/17 (82%)	0.73	1 (7%) 19 13	56, 88, 102, 117	0
2	O	13/17 (76%)	1.13	2 (15%) 3 1	55, 85, 101, 102	0
2	P	14/17 (82%)	0.67	1 (7%) 19 13	54, 92, 99, 99	0
2	Q	13/17 (76%)	1.66	5 (38%) 0 0	70, 100, 109, 115	0
2	R	14/17 (82%)	0.26	1 (7%) 19 13	62, 82, 98, 108	0
2	S	15/17 (88%)	0.50	1 (6%) 21 15	63, 91, 112, 115	0
2	T	13/17 (76%)	0.93	1 (7%) 16 11	72, 93, 110, 116	0
2	U	13/17 (76%)	1.76	5 (38%) 0 0	72, 100, 111, 113	0
2	V	13/17 (76%)	1.13	2 (15%) 3 1	64, 92, 103, 104	0
2	W	13/17 (76%)	2.54	5 (38%) 0 0	69, 101, 117, 125	0
All	All	2691/2959 (90%)	0.27	124 (4%) 36 29	39, 70, 113, 159	0

The worst 5 of 124 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	180	GLU	9.6
1	B	178	SER	8.7
2	W	316	PHE	8.4
1	D	6	LEU	7.1
1	J	6	LEU	7.1

## 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.