



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 13, 2024 – 01:19 AM EDT

PDB ID : 3KGV  
Title : Crystal Structure of Human DNA-dependent Protein Kinase Catalytic Subunit (DNA-PKcs)  
Authors : Sibanda, B.L.; Chirgadze, D.Y.; Blundell, T.L.  
Deposited on : 2009-10-29  
Resolution : 6.60 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

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Xtriage (Phenix) : 1.20.1  
EDS : 2.36.2  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.36.2

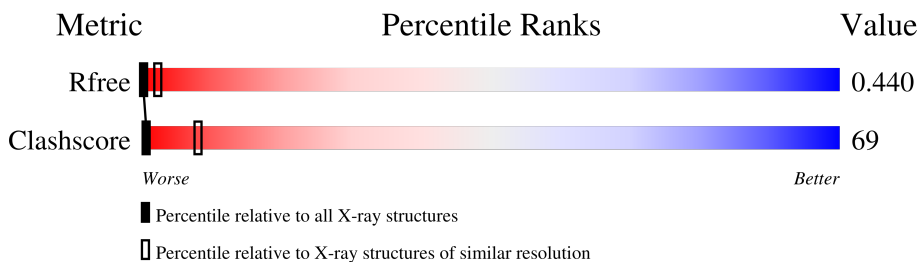
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

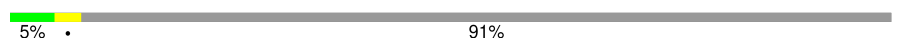



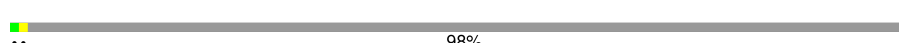




The reported resolution of this entry is 6.60 Å.

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.



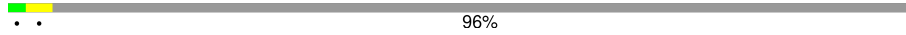
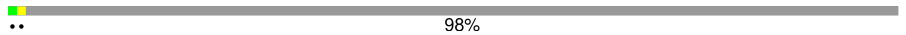

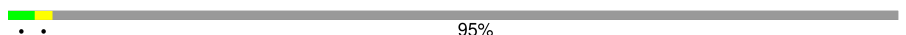
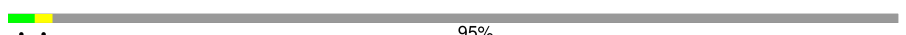
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	1000 (9.00-3.90)
Clashscore	141614	1064 (9.00-3.90)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	4128	 5% . 91%
1	B	4128	 . . 93%
1	C	4128	 5% 5% 90%
1	D	4128	 . . 96%
1	E	4128	 .. 98%
1	F	4128	 7% 6% 87%
1	O	4128	 5% . 91%
1	P	4128	 . . 93%
1	Q	4128	 5% 5% 90%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	R	4128	 96%
1	S	4128	 98%
1	T	4128	 7% 6% 87%
1	X	4128	 95%
1	Y	4128	 95%

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 20320 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 DNA-dependent Protein Kinase Catalytic Subunit.

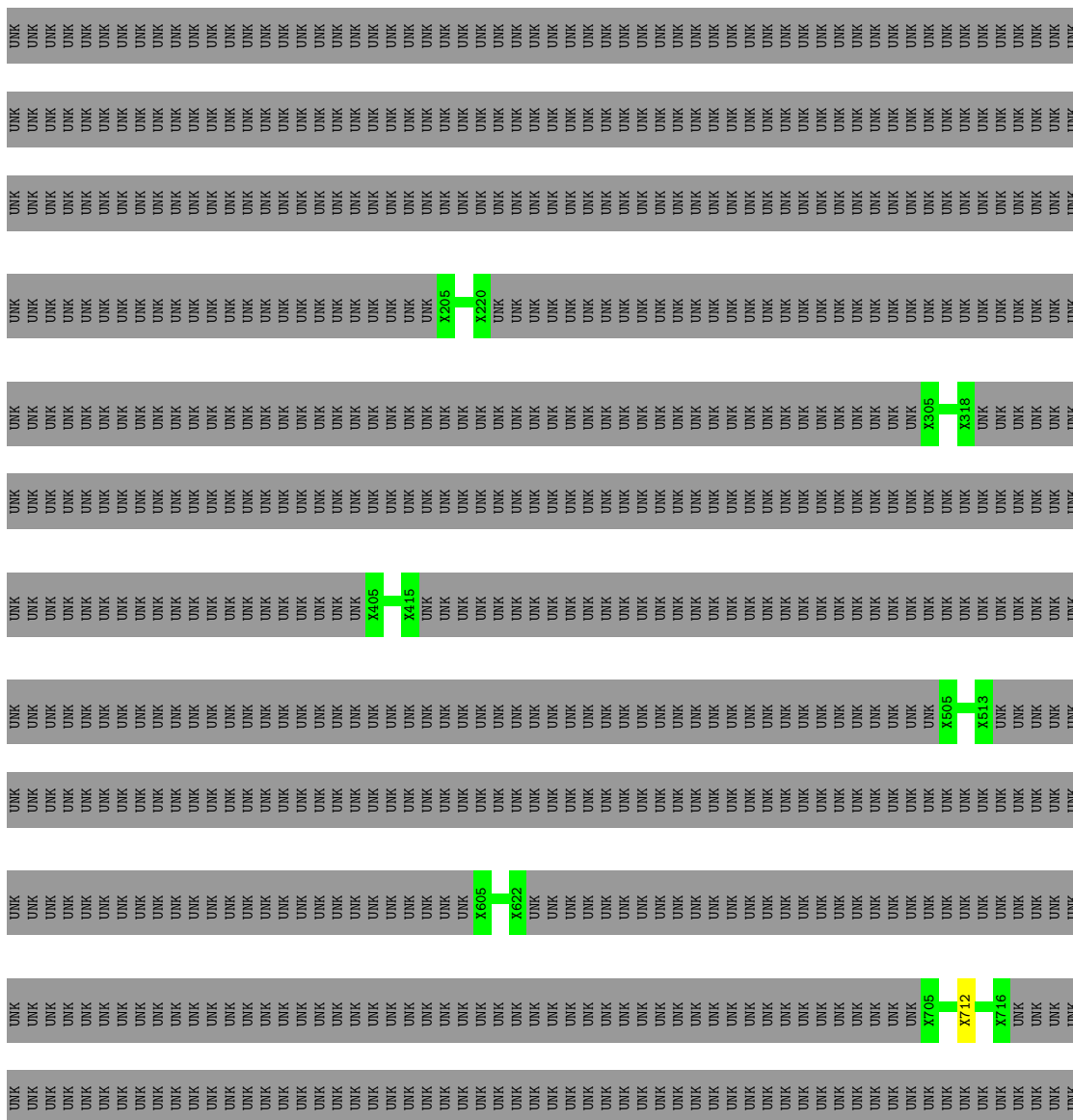
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
1	A	354	Total	C	N	O	0	0	0
			1770	1062	354	354			
1	B	283	Total	C	N	O	0	0	0
			1415	849	283	283			
1	C	406	Total	C	N	O	0	0	0
			2030	1218	406	406			
1	D	182	Total	C	N	O	0	0	0
			910	546	182	182			
1	E	65	Total	C	N	O	0	0	0
			325	195	65	65			
1	F	531	Total	C	N	O	0	0	0
			2655	1593	531	531			
1	X	211	Total	C	N	O	0	0	0
			1055	633	211	211			
1	O	354	Total	C	N	O	0	0	0
			1770	1062	354	354			
1	P	283	Total	C	N	O	0	0	0
			1415	849	283	283			
1	Q	406	Total	C	N	O	0	0	0
			2030	1218	406	406			
1	R	182	Total	C	N	O	0	0	0
			910	546	182	182			
1	S	65	Total	C	N	O	0	0	0
			325	195	65	65			
1	T	531	Total	C	N	O	0	0	0
			2655	1593	531	531			
1	Y	211	Total	C	N	O	0	0	0
			1055	633	211	211			

### 3 Residue-property plots i

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 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: DNA-dependent Protein Kinase Catalytic Subunit

Chain A:  5% . 91%















This page contains 30 horizontal bars, each representing a set of validation metrics for a specific structural component. Each bar contains 45 columns of data, with every cell containing the text 'UNK', indicating missing or unknown values for all metrics in this dataset.

UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK									
UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK				
UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK	UNK









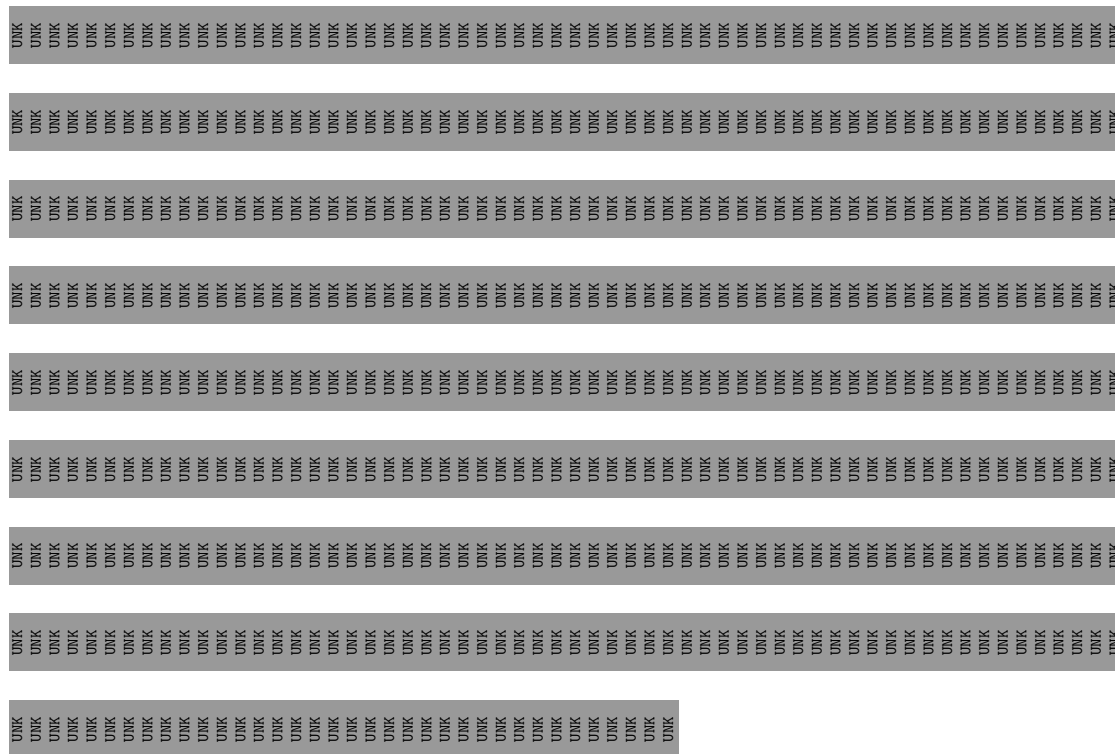






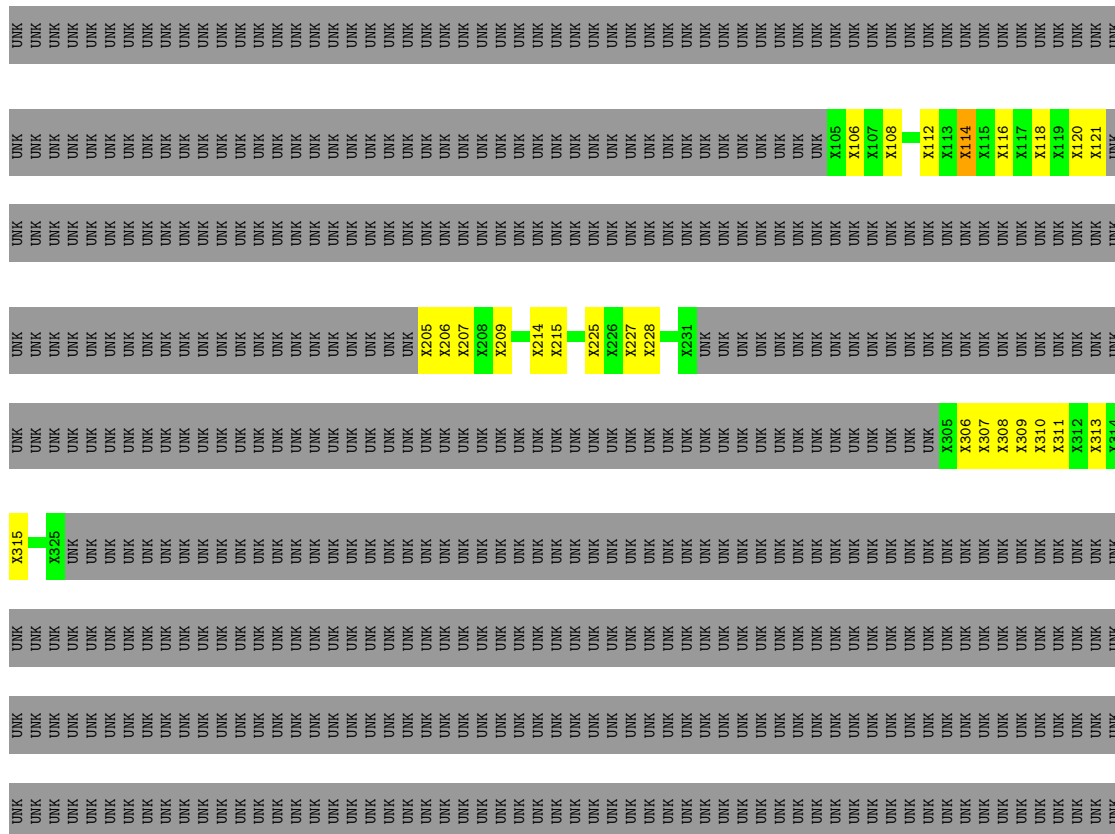






● Molecule 1: DNA-dependent Protein Kinase Catalytic Subunit

Chain E: ..  98%





This section contains 20 rows of data, each row representing a different validation metric. Each row consists of 15 columns of data, all of which are 'UNK'. The rows are as follows:

- Row 1: UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK
- Row 2: UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK
- Row 3: UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK
- Row 4: UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK
- Row 5: UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK
- Row 6: UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK
- Row 7: UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK
- Row 8: UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK
- Row 9: UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK
- Row 10: UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK
- Row 11: UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK
- Row 12: UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK
- Row 13: UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK
- Row 14: UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK
- Row 15: UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK
- Row 16: UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK
- Row 17: UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK
- Row 18: UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK
- Row 19: UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK
- Row 20: UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK



Chain F: 7% 6% 87%

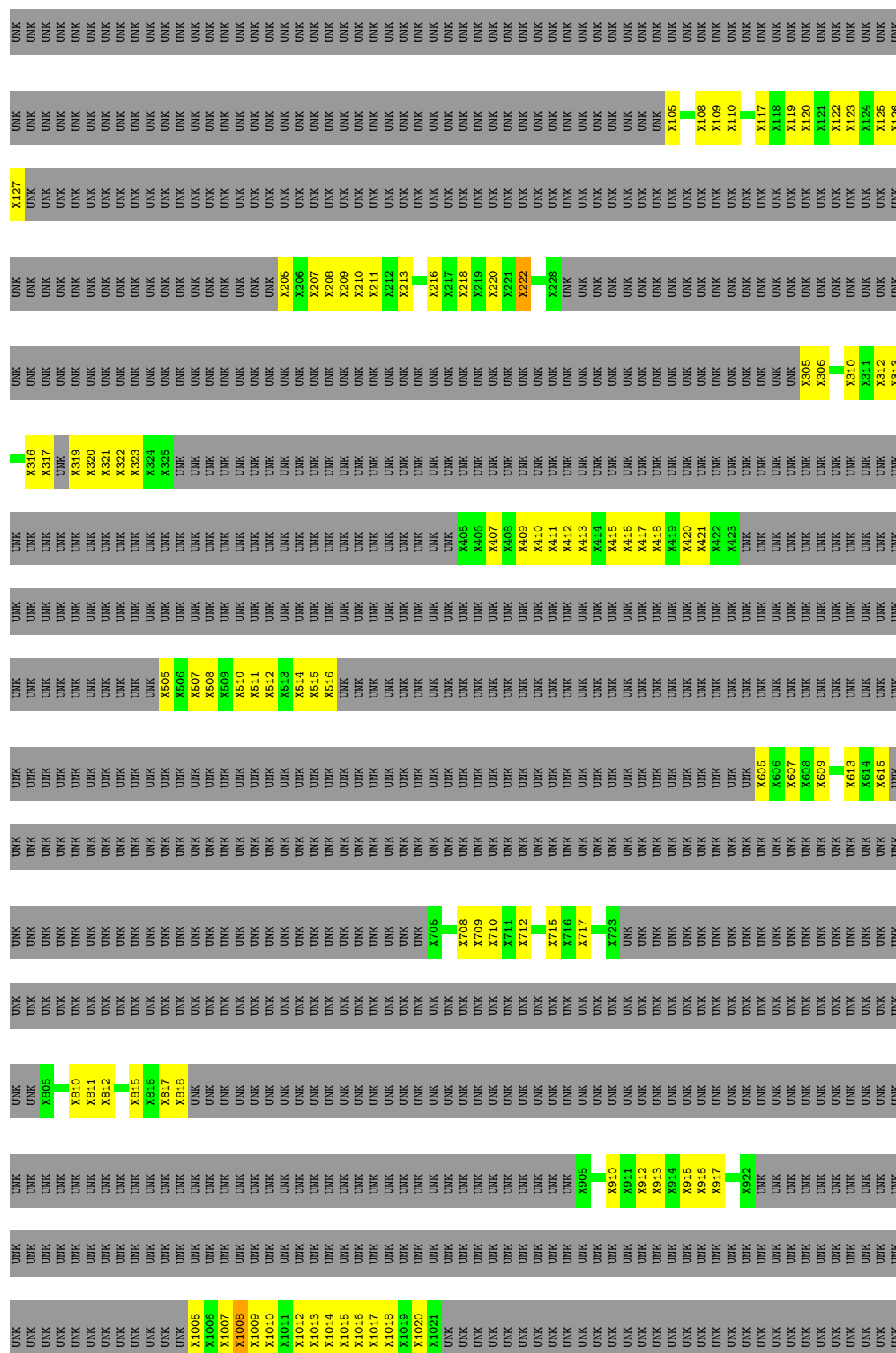






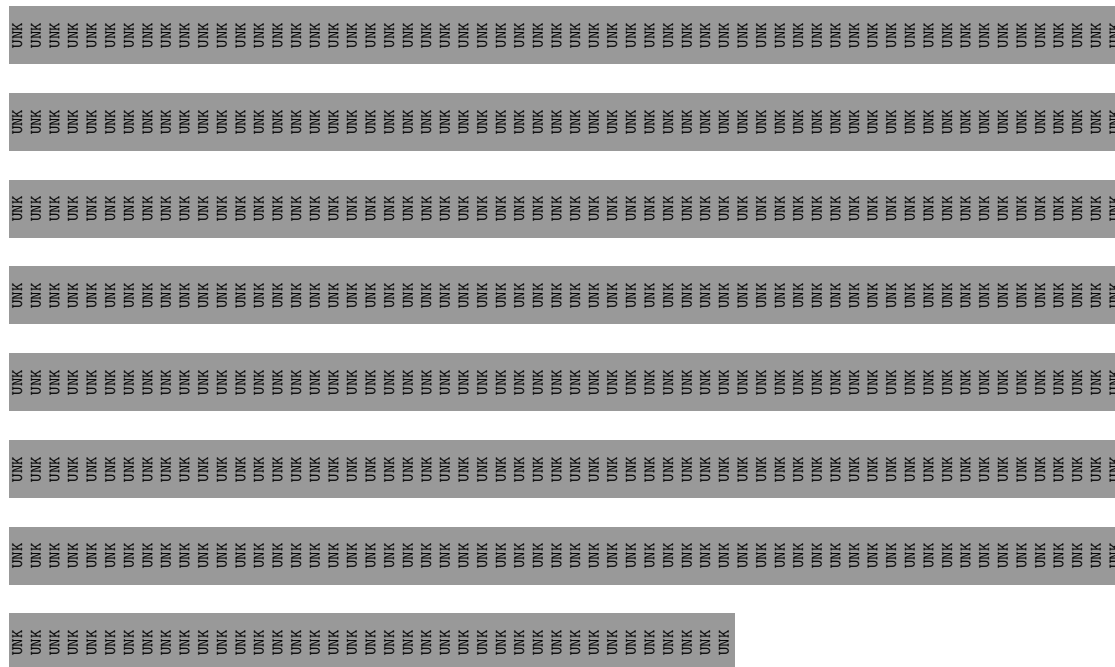




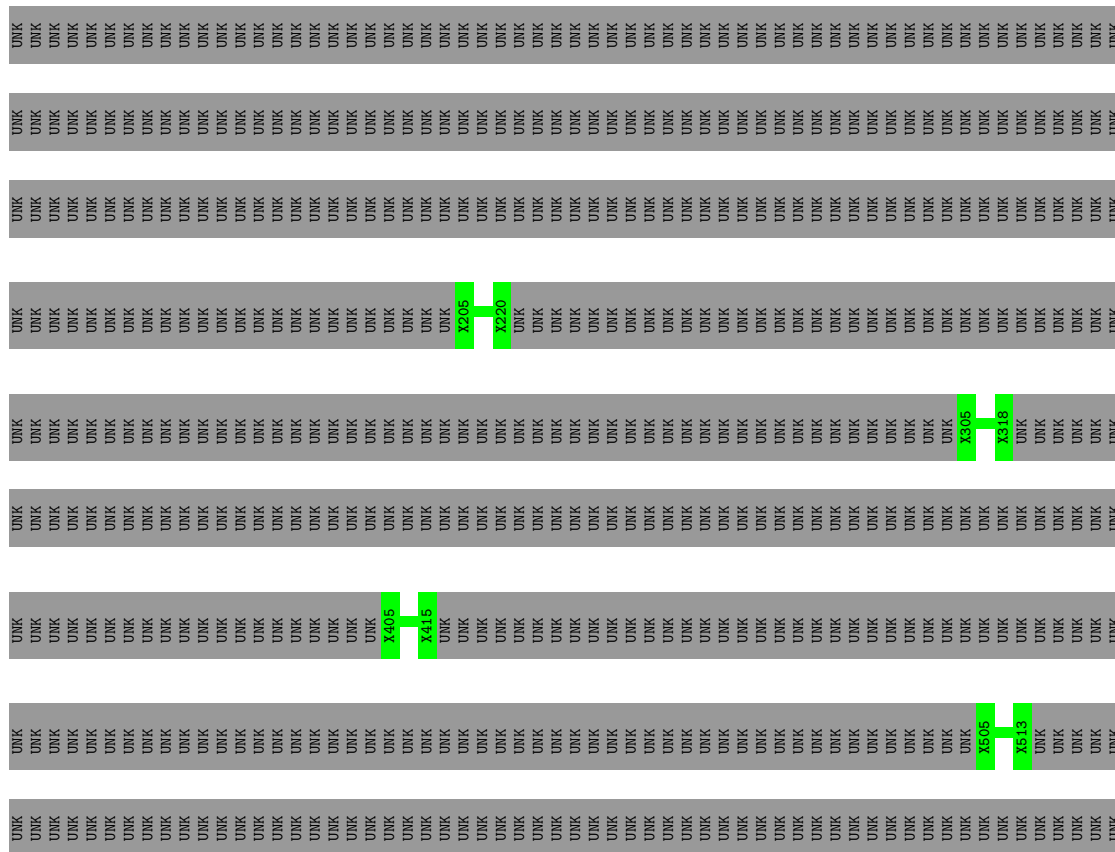




Table with 16 rows and 24 columns. Each cell contains text such as 'UNK' or 'UNK UNK'.



● Molecule 1: DNA-dependent Protein Kinase Catalytic Subunit













• Molecule 1: DNA-dependent Protein Kinase Catalytic Subunit

Chain P: . . . 93%

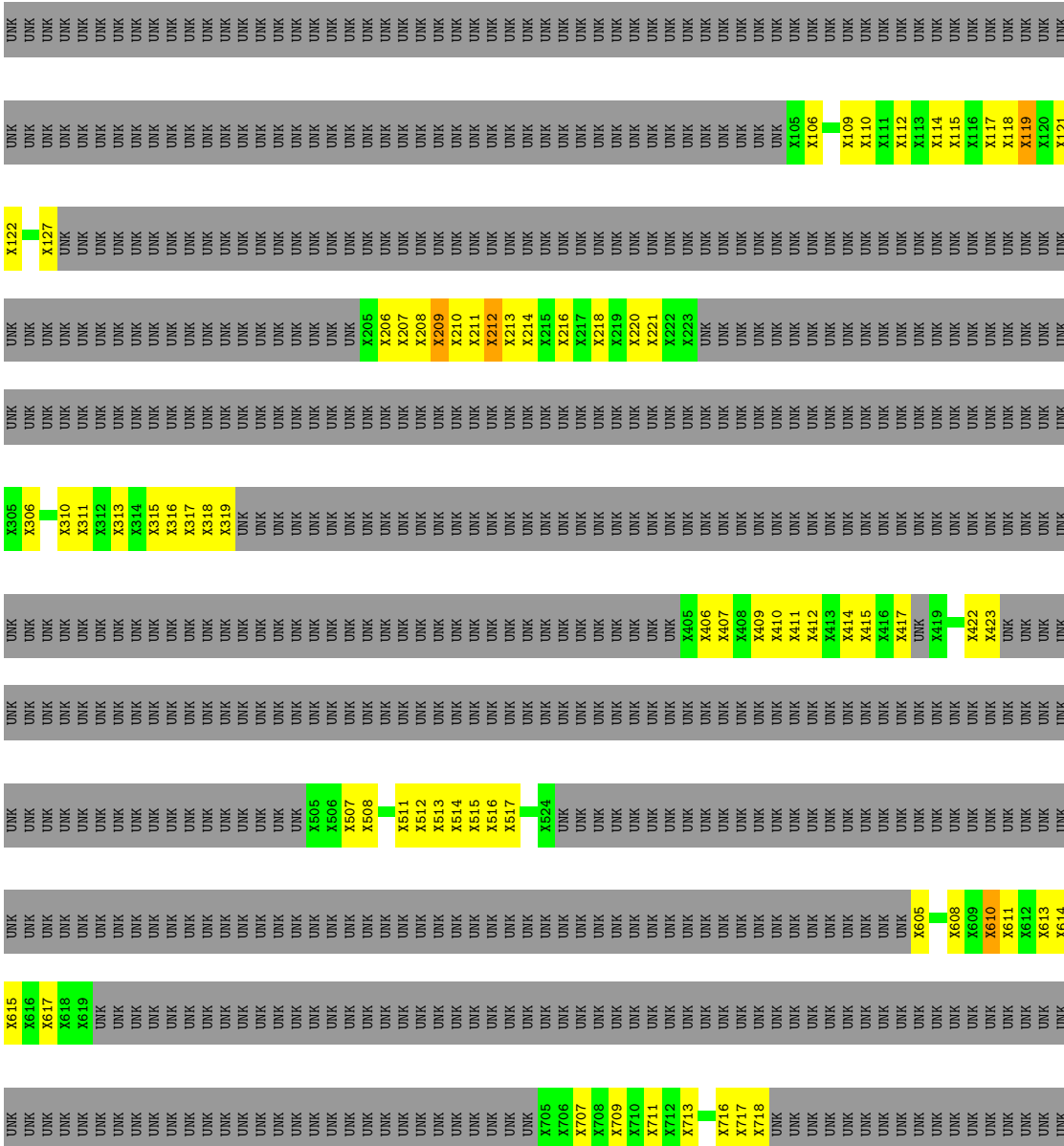
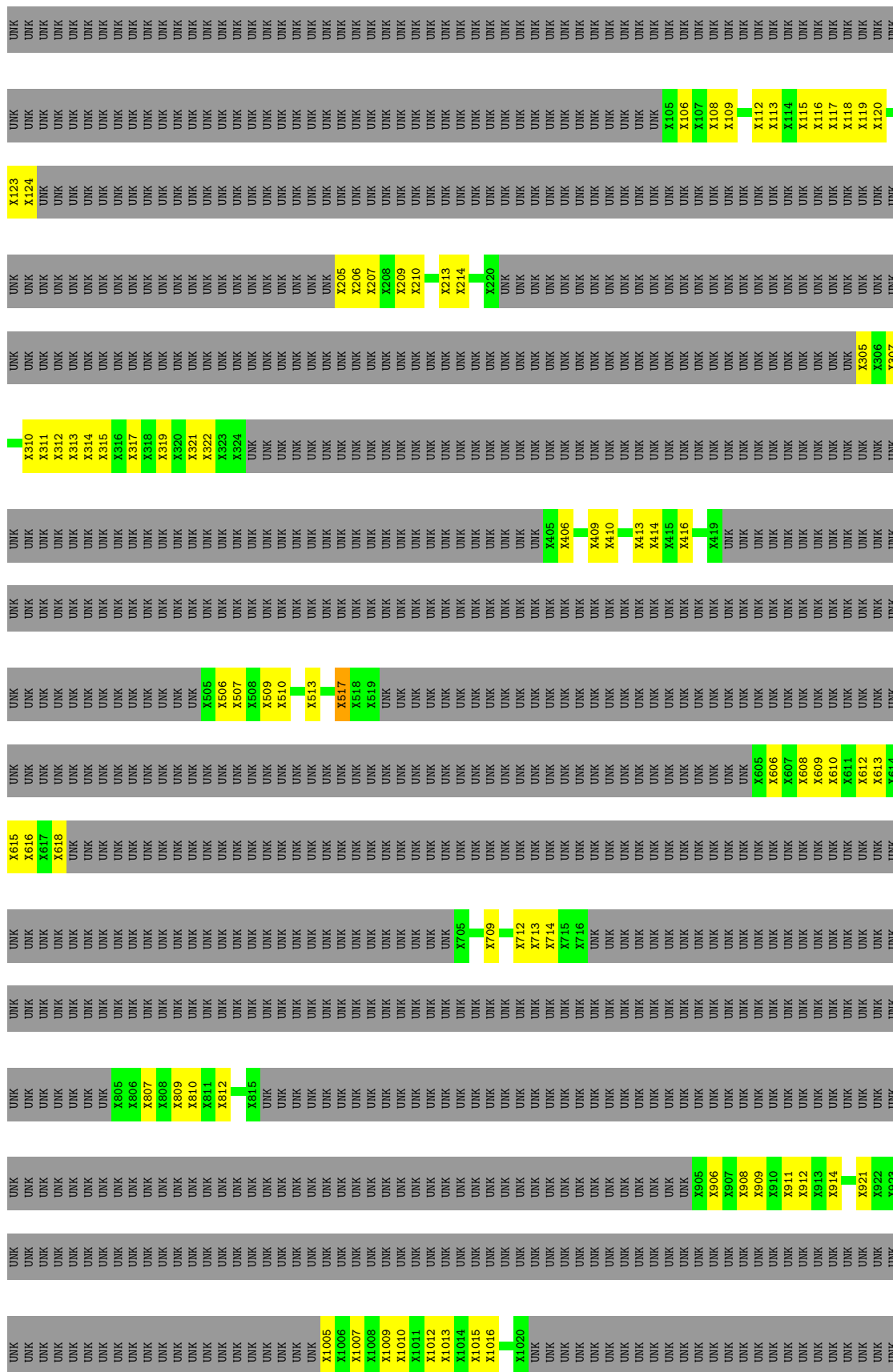




Table with 37 rows. Each row contains 32 columns of data, all of which are the text 'UNK'.



Chain Q: 5% 5% 90%



















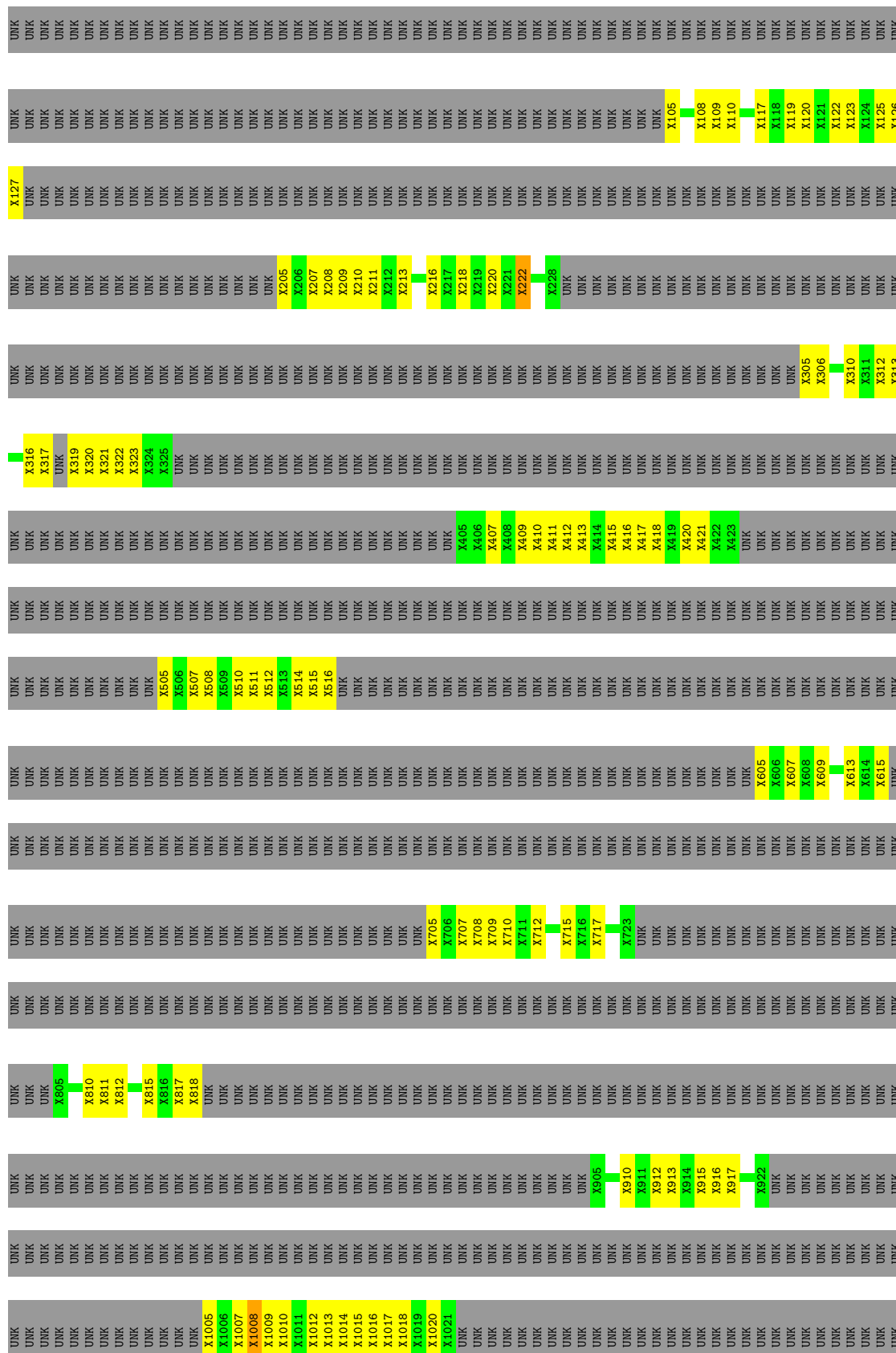








Chain T: 7% 6% 87%



















## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	168.60Å 132.99Å 297.00Å 90.00° 105.58° 90.00°	Depositor
Resolution (Å)	100.00 – 6.60 91.97 – 6.58	Depositor EDS
% Data completeness (in resolution range)	97.6 (100.00-6.60) 97.6 (91.97-6.58)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	0.08	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	12.86 (at 6.72Å)	Xtrriage
Refinement program	REFMAC 5.5.0102	Depositor
R, $R_{free}$	0.442 , 0.441 0.441 , 0.440	Depositor DCC
$R_{free}$ test set	1223 reflections (5.12%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	269.3	Xtrriage
Anisotropy	0.049	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.23 , 620.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	0.010 for h,-k,-h-l	Xtrriage
$F_o, F_c$ correlation	0.63	EDS
Total number of atoms	20320	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	80.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.94% 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, 0.333 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).

There are no protein, RNA or DNA chains available to summarize Z scores of covalent bonds and angles.

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	2
1	B	0	18
1	C	0	8
1	D	0	8
1	E	0	3
1	F	0	6
1	O	0	2
1	P	0	18
1	Q	0	8
1	R	0	8
1	S	0	3
1	T	0	6
1	X	0	7
1	Y	0	7
All	All	0	104

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 104 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1214	UNK	Peptide
1	A	2306	UNK	Peptide
1	B	119	UNK	Peptide
1	B	122	UNK	Peptide

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Group
1	B	209	UNK	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	1770	0	429	123	0
1	B	1415	0	356	149	0
1	C	2030	0	489	178	0
1	D	910	0	239	117	0
1	E	325	0	78	20	0
1	F	2655	0	648	203	0
1	O	1770	0	429	123	48
1	P	1415	0	358	153	0
1	Q	2030	0	489	180	0
1	R	910	0	239	119	0
1	S	325	0	78	19	0
1	T	2655	0	648	208	48
1	X	1055	0	263	74	0
1	Y	1055	0	264	75	0
All	All	20320	0	5007	1740	48

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 69.

The worst 5 of 1740 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:B:1412:UNK:CA	1:B:1412:UNK:C	1.75	1.64
1:P:1412:UNK:C	1:P:1412:UNK:CA	1.74	1.58
1:P:1412:UNK:HA	1:P:1415:UNK:CB	1.69	1.22
1:B:1412:UNK:HA	1:B:1415:UNK:CB	1.69	1.20
1:E:116:UNK:O	1:E:118:UNK:N	1.76	1.17

The worst 5 of 48 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)
1:O:619:UNK:C	1:T:2418:UNK:O[1_545]	0.56	1.64
1:O:620:UNK:C	1:T:2419:UNK:CA[1_545]	0.57	1.63
1:O:622:UNK:N	1:T:2418:UNK:N[1_545]	0.67	1.53
1:O:622:UNK:C	1:T:2417:UNK:O[1_545]	0.68	1.52
1:O:621:UNK:O	1:T:2417:UNK:CA[1_545]	0.91	1.29

### 5.3 Torsion angles [i](#)

#### 5.3.1 Protein backbone [i](#)

There are no protein backbone outliers to report in this entry.

#### 5.3.2 Protein sidechains [i](#)

There are no protein residues with a non-rotameric sidechain to report in this entry.

#### 5.3.3 RNA [i](#)

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 monosaccharides 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

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	0/4128	-	-	-	-
1	B	0/4128	-	-	-	-
1	C	0/4128	-	-	-	-
1	D	0/4128	-	-	-	-
1	E	0/4128	-	-	-	-
1	F	0/4128	-	-	-	-
1	O	0/4128	-	-	-	-
1	P	0/4128	-	-	-	-
1	Q	0/4128	-	-	-	-
1	R	0/4128	-	-	-	-
1	S	0/4128	-	-	-	-
1	T	0/4128	-	-	-	-
1	X	0/4128	-	-	-	-
1	Y	0/4128	-	-	-	-
All	All	0/57792	-	-	-	-

There are no RSRZ outliers to report.

### 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 monosaccharides 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.