



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 11, 2024 – 10:35 PM EDT

PDB ID : 1PQV  
Title : RNA polymerase II-TFIIS complex  
Authors : Kettenberger, H.; Armache, K.-J.; Cramer, P.  
Deposited on : 2003-06-19  
Resolution : 3.80 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

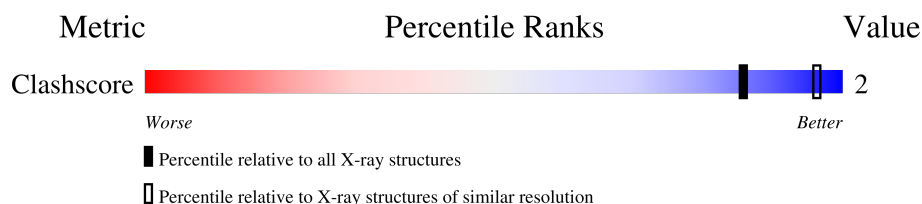
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*





The reported resolution of this entry is 3.80 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	1288 (4.00-3.60)

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 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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	1733	 81% 19%
2	B	1224	 90% 10%
3	C	318	 84% 16%
4	D	221	 60% . 39%
5	E	215	 99% .
6	F	155	 54% 46%
7	G	171	 99% .
8	H	146	 91% 9%
9	I	122	 98% .
10	J	70	 93% 7%
11	K	120	 95% 5%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
12	L	70	 66% 34%
13	S	309	 54% • 44%

## 2 Entry composition

There are 15 unique types of molecules in this entry. The entry contains 4041 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-directed RNA polymerase II largest subunit.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf	Trace
1	A	1409	Total	C	0	0	1409
			1409	1409			

- Molecule 2 is a protein called DNA-directed RNA polymerase II 140 kDa polypeptide.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf	Trace
2	B	1103	Total	C	0	0	1103
			1103	1103			

- Molecule 3 is a protein called DNA-directed RNA polymerase II 45 kDa polypeptide.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf	Trace
3	C	266	Total	C	0	0	266
			266	266			

- Molecule 4 is a protein called DNA-directed RNA polymerase II 32 kDa polypeptide.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf	Trace
4	D	135	Total	C	0	0	135
			135	135			

There are 29 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	86	UNK	-	SEE REMARK 999	UNP P20433
D	87	UNK	-	SEE REMARK 999	UNP P20433
D	88	UNK	-	SEE REMARK 999	UNP P20433
D	89	UNK	-	SEE REMARK 999	UNP P20433
D	90	UNK	-	SEE REMARK 999	UNP P20433
D	91	UNK	-	SEE REMARK 999	UNP P20433
D	92	UNK	-	SEE REMARK 999	UNP P20433
D	93	UNK	-	SEE REMARK 999	UNP P20433

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	94	UNK	-	SEE REMARK 999	UNP P20433
D	95	UNK	-	SEE REMARK 999	UNP P20433
D	96	UNK	-	SEE REMARK 999	UNP P20433
D	97	UNK	-	SEE REMARK 999	UNP P20433
D	98	UNK	-	SEE REMARK 999	UNP P20433
D	99	UNK	-	SEE REMARK 999	UNP P20433
D	100	UNK	-	SEE REMARK 999	UNP P20433
D	101	UNK	-	SEE REMARK 999	UNP P20433
D	102	UNK	-	SEE REMARK 999	UNP P20433
D	103	UNK	-	SEE REMARK 999	UNP P20433
D	104	UNK	-	SEE REMARK 999	UNP P20433
D	105	UNK	-	SEE REMARK 999	UNP P20433
D	106	UNK	-	SEE REMARK 999	UNP P20433
D	107	UNK	-	SEE REMARK 999	UNP P20433
D	108	UNK	-	SEE REMARK 999	UNP P20433
D	109	UNK	-	SEE REMARK 999	UNP P20433
D	110	UNK	-	SEE REMARK 999	UNP P20433
D	111	UNK	-	SEE REMARK 999	UNP P20433
D	112	UNK	-	SEE REMARK 999	UNP P20433
D	113	UNK	-	SEE REMARK 999	UNP P20433
D	114	UNK	-	SEE REMARK 999	UNP P20433

- Molecule 5 is a protein called DNA-directed RNA polymerases I, II, and III 27 kDa polypeptide.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
5	E	214	Total C 214 214	0	0	214

- Molecule 6 is a protein called DNA-directed RNA polymerases I, II, and III 23 kDa polypeptide.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
6	F	84	Total C 84 84	0	0	84

- Molecule 7 is a protein called DNA-directed RNA polymerase II 19 kDa polypeptide.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
7	G	169	Total C 169 169	0	0	169

- Molecule 8 is a protein called DNA-directed RNA polymerases I, II, and III 14.5 kDa polypeptide.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf	Trace
8	H	133	Total	C	0	0	133
			133	133			

- Molecule 9 is a protein called DNA-directed RNA polymerase II 14.2 kDa polypeptide.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf	Trace
9	I	119	Total	C	0	0	119
			119	119			

- Molecule 10 is a protein called DNA-directed RNA polymerases I, II, and III 8.3 kDa polypeptide.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf	Trace
10	J	65	Total	C	0	0	65
			65	65			

- Molecule 11 is a protein called DNA-directed RNA polymerase II 13.6 kDa polypeptide.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf	Trace
11	K	114	Total	C	0	0	114
			114	114			

- Molecule 12 is a protein called DNA-directed RNA polymerases I, II, and III 7.7 kDa polypeptide.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf	Trace
12	L	46	Total	C	0	0	46
			46	46			

- Molecule 13 is a protein called Transcription elongation factor S-II.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf	Trace
13	S	174	Total	C	0	0	174
			174	174			

- Molecule 14 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
14	A	1	Total	Mg	0	0
			1	1		

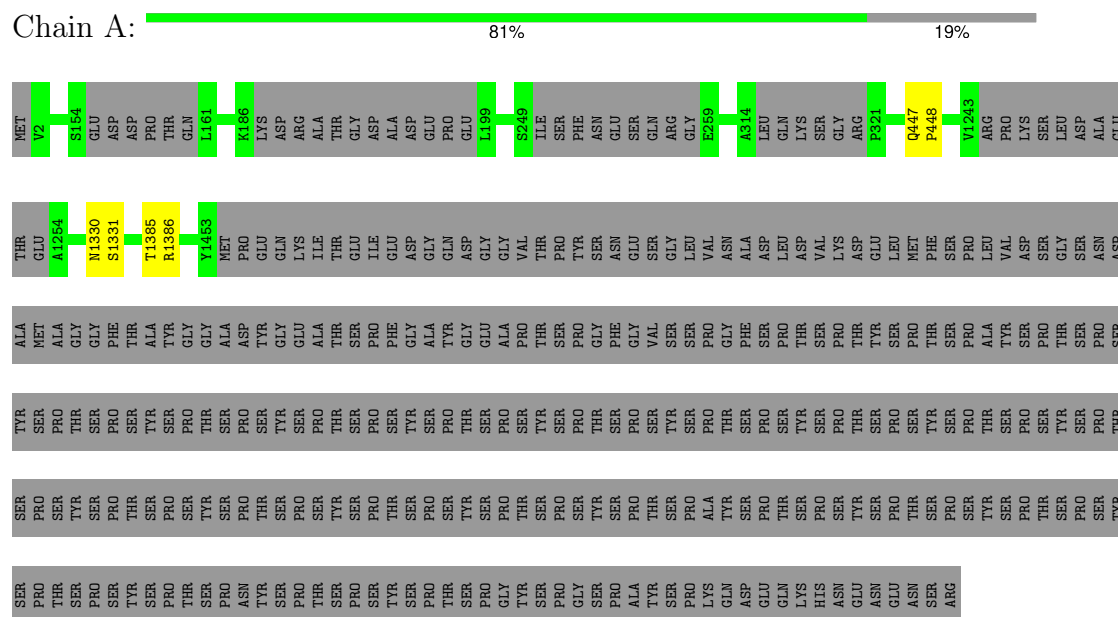
- Molecule 15 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
15	A	2	Total 2	Zn 2	0	0
15	B	1	Total 1	Zn 1	0	0
15	C	1	Total 1	Zn 1	0	0
15	I	2	Total 2	Zn 2	0	0
15	J	1	Total 1	Zn 1	0	0
15	L	1	Total 1	Zn 1	0	0
15	S	1	Total 1	Zn 1	0	0

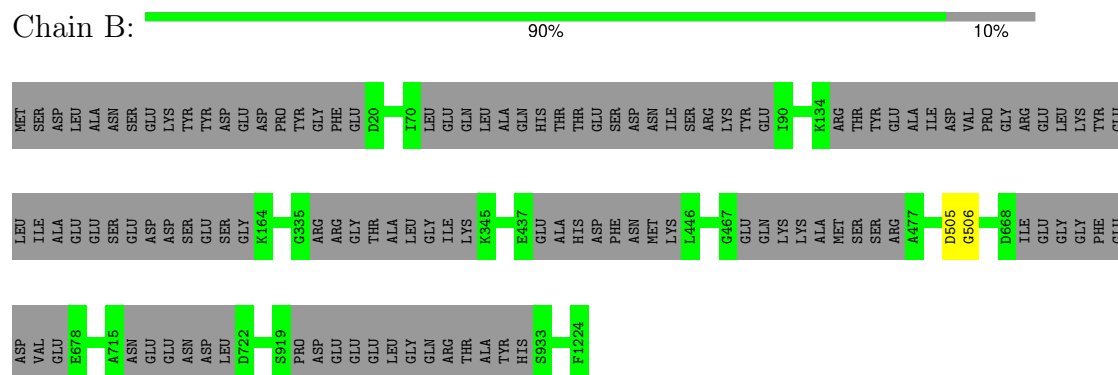
### 3 Residue-property plots [i](#)

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. 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.

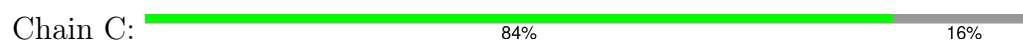
- Molecule 1: DNA-directed RNA polymerase II largest subunit



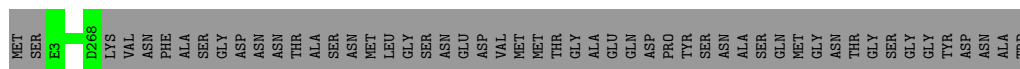
- Molecule 2: DNA-directed RNA polymerase II 140 kDa polypeptide



- Molecule 3: DNA-directed RNA polymerase II 45 kDa polypeptide

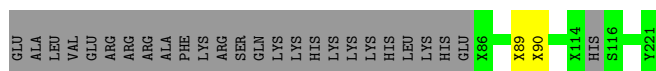
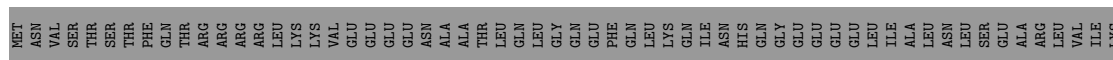






- Molecule 4: DNA-directed RNA polymerase II 32 kDa polypeptide

Chain D: 60% 39%



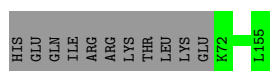
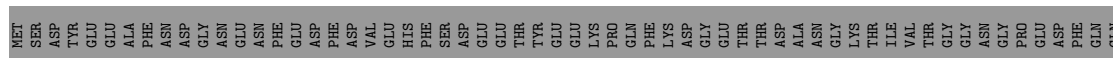
- Molecule 5: DNA-directed RNA polymerases I, II, and III 27 kDa polypeptide

Chain E: 99%



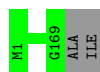
- Molecule 6: DNA-directed RNA polymerases I, II, and III 23 kDa polypeptide

Chain F: 54% 46%



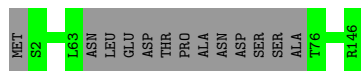
- Molecule 7: DNA-directed RNA polymerase II 19 kDa polypeptide

Chain G: 99%



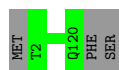
- Molecule 8: DNA-directed RNA polymerases I, II, and III 14.5 kDa polypeptide

Chain H: 91% 9%



- Molecule 9: DNA-directed RNA polymerase II 14.2 kDa polypeptide

Chain I: 98%



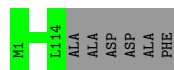
- Molecule 10: DNA-directed RNA polymerases I, II, and III 8.3 kDa polypeptide

Chain J:  93% 7%



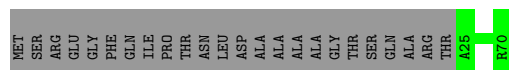
- Molecule 11: DNA-directed RNA polymerase II 13.6 kDa polypeptide

Chain K:  95% 5%



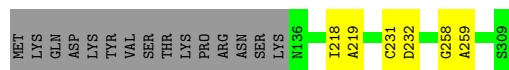
- Molecule 12: DNA-directed RNA polymerases I, II, and III 7.7 kDa polypeptide

Chain L:  66% 34%



- Molecule 13: Transcription elongation factor S-II

Chain S:  54% 44%



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	218.90Å 395.30Å 281.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 3.80 49.41 – 3.79	Depositor EDS
% Data completeness (in resolution range)	96.2 (50.00-3.80) 89.4 (49.41-3.79)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.09	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.07 (at 3.77Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	(Not available) , (Not available) 0.449 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	91.0	Xtriage
Anisotropy	0.459	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.27 , 91.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.015 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.019 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
$F_o, F_c$ correlation	0.62	EDS
Total number of atoms	4041	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.16% 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: MG, ZN

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.

There are no bond length outliers.

There are no bond angle outliers.

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	1409	0	0	3	0
2	B	1103	0	0	1	0
3	C	266	0	0	0	0
4	D	135	0	0	1	0
5	E	214	0	0	1	0
6	F	84	0	0	0	0
7	G	169	0	0	0	0
8	H	133	0	0	0	0
9	I	119	0	0	0	0
10	J	65	0	0	0	0
11	K	114	0	0	0	0
12	L	46	0	0	0	0
13	S	174	0	0	3	0
14	A	1	0	0	0	0
15	A	2	0	0	0	0
15	B	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
15	C	1	0	0	0	0
15	I	2	0	0	0	0
15	J	1	0	0	0	0
15	L	1	0	0	0	0
15	S	1	0	0	0	0
All	All	4041	0	0	9	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 2.

The worst 5 of 9 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:1330:ASN:CA	1:A:1331:SER:CA	2.54	0.85
4:D:89:UNK:CA	4:D:90:UNK:CA	2.57	0.83
1:A:1385:THR:CA	1:A:1386:ARG:CA	2.57	0.81
2:B:505:ASP:CA	2:B:506:GLY:CA	2.63	0.76
13:S:218:ILE:CA	13:S:219:ALA:CA	2.73	0.66

There are no symmetry-related clashes.

## 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](#)

Of 10 ligands modelled in this entry, 10 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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 ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.3 Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.4 Ligands ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.5 Other polymers ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.