wwPDB EM Map/Model Validation Report ⓘ

Jun 14, 2016 – 12:09 PM EDT

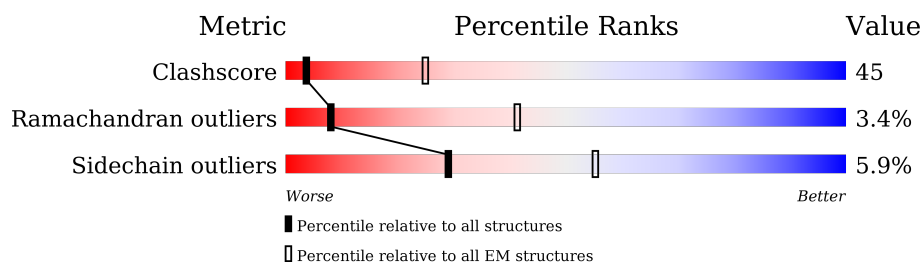|		
PDB ID	:	5IVW
EMDB ID:	:	EMD-8131
Title	:	Human core TFIIH bound to DNA within the PIC
Authors	:	He, Y.; Yan, C.; Fang, J.; Inouye, C.; Tjian, R.; Ivanov, I.; Nogales, E.
Deposited on	:	2016-03-21
Resolution	:	10.00 Å(reported)
Based on PDB ID	:	?

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 10.00 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	Percentile Ranks	Value
Clashscore		45
Ramachandran outliers		3.4%
Sidechain outliers		5.9%

Worse   *Better*

■ Percentile relative to all structures
▯ Percentile relative to all EM structures

Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the 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 <=5%

Mol	Chain	Length	Quality of chain
1	V	782	38% 17% 5% • 39%
2	W	760	59% 24% • • 13%
3	0	395	31% 15% • 52%
4	1	71	18% 59% 10% 13%
5	2	462	22% 34% • 41%
6	3	308	16% 41% 6% 37%
7	X	19	53% 47%
8	Y	20	40% 60%

2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 15688 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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 TFIIH basal transcription factor complex helicase XPB subunit.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	V	475	Total 3855	C 2454	N 663	O 712	S 26	0	0

- Molecule 2 is a protein called TFIIH basal transcription factor complex helicase XPD subunit.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	W	665	Total 5348	C 3415	N 932	O 975	S 26	0	0

- Molecule 3 is a protein called General transcription factor IIH subunit 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	0	188	Total 1479	C 935	N 258	O 276	S 10	0	0

- Molecule 4 is a protein called General transcription factor IIH subunit 5.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	1	62	Total 491	C 317	N 77	O 93	S 4	0	0

- Molecule 5 is a protein called General transcription factor IIH subunit 4.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	2	274	Total 2196	C 1417	N 377	O 392	S 10	0	0

- Molecule 6 is a protein called General transcription factor IIH subunit 3.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	3	193	Total 1526	C 978	N 252	O 284	S 12	0	0

- Molecule 7 is a DNA chain called scp-X.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	X	19	Total 392	C 187	N 83	O 104	P 18	0	0

- Molecule 8 is a DNA chain called scp-Y.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	Y	20	Total 401	C 196	N 59	O 127	P 19	0	0

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 errors 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: TFIIH basal transcription factor complex helicase XPB subunit

Chain V:

• Molecule 2: TFIIH basal transcription factor complex helicase XPD subunit

Chain W:

• Molecule 3: General transcription factor IIH subunit 2

Chain 0:

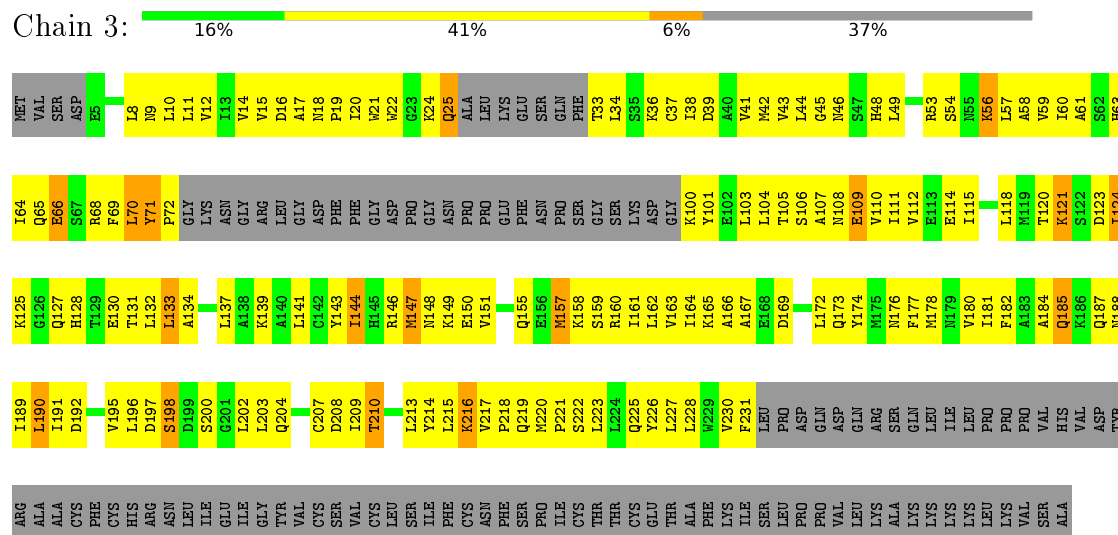• Molecule 4: General transcription factor IIH subunit 5

Chain 1:

• Molecule 5: General transcription factor IIH subunit 4

Chain 2:
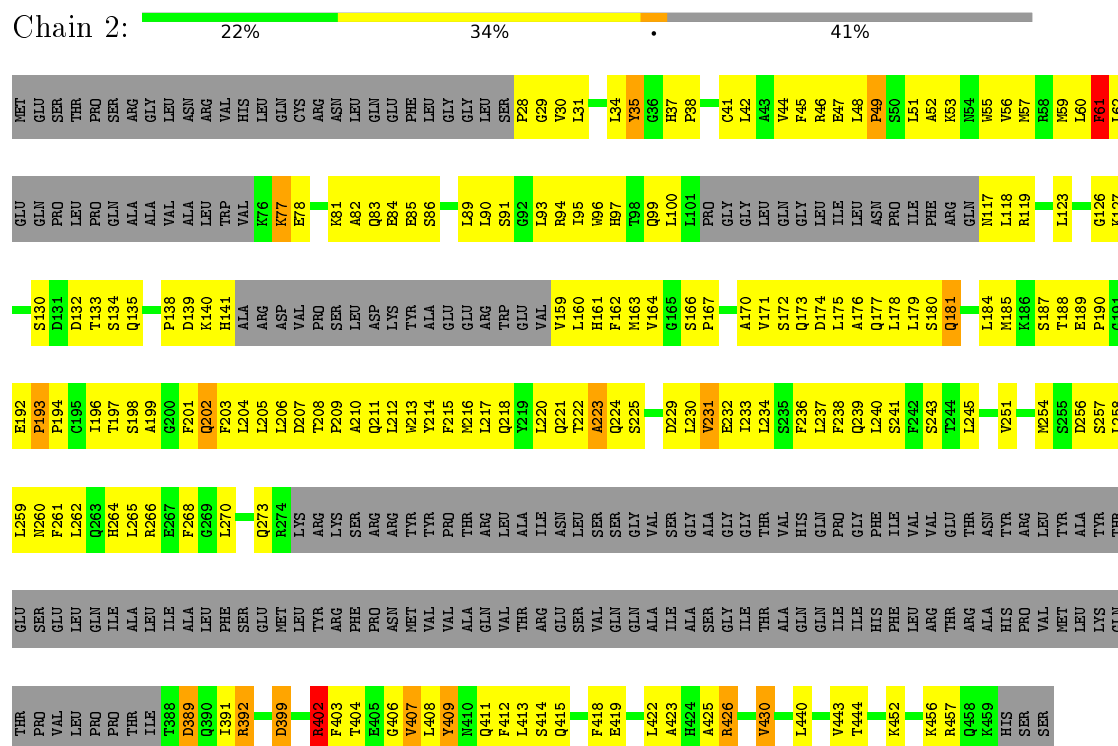
• Molecule 6: General transcription factor IIH subunit 3

Chain 3:

• Molecule 7: scp-X

Chain X:

• Molecule 8: scp-Y

Chain Y:
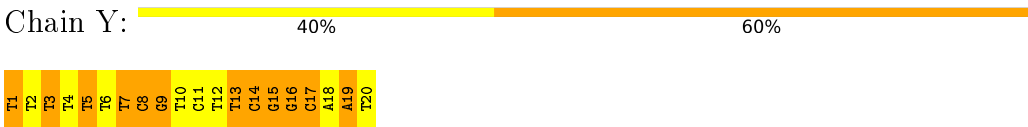
4 Experimental information (i)

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	219771	Depositor
Resolution determination method	FSC 0.5 CUT-OFF	Depositor
CTF correction method	Not provided	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose $(e^-/\text{Å}^2)$	Not provided	Depositor
Minimum defocus (nm)	2000	Depositor
Maximum defocus (nm)	4000	Depositor
Magnification	27500	Depositor
Image detector	Not provided	Depositor

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	$\#|Z|>2$	RMSZ	$\#|Z|>2$
1	V	1.44	15/3931 (0.4%)	1.95	110/5298 (2.1%)
2	W	1.50	21/5460 (0.4%)	2.00	150/7390 (2.0%)
3	0	1.45	5/1506 (0.3%)	1.92	42/2038 (2.1%)
4	1	0.83	0/496	1.12	0/669
5	2	0.88	0/2243	1.18	9/3024 (0.3%)
6	3	0.85	0/1548	1.13	2/2090 (0.1%)
7	X	3.20	45/443 (10.2%)	3.95	129/682 (18.9%)
8	Y	3.42	58/445 (13.0%)	4.25	126/685 (18.4%)
All	All	1.51	144/16072 (0.9%)	2.00	568/21876 (2.6%)

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	V	0	12
2	W	0	8
5	2	0	8
7	X	0	7
8	Y	0	12
All	All	0	47

The worst 5 of 144 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	X	10	DG	N7-C5	17.61	1.49	1.39
8	Y	6	DT	C5-C7	15.71	1.59	1.50
8	Y	2	DT	C5-C7	11.35	1.56	1.50
7	X	1	DA	N3-C4	10.46	1.41	1.34
8	Y	4	DT	C5-C6	9.83	1.41	1.34

The worst 5 of 568 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	0	77	LYS	C-N-CD	-24.19	67.39	120.60
6	3	71	TYR	C-N-CD	-20.68	75.09	120.60
2	W	335	ARG	NE-CZ-NH1	-19.48	110.56	120.30
2	W	26	ARG	NE-CZ-NH2	18.81	129.70	120.30
1	V	358	ARG	NE-CZ-NH2	17.52	129.06	120.30

There are no chirality outliers.

5 of 47 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	V	247	ASP	Mainchain
1	V	251	PHE	Sidechain
1	V	319	TYR	Sidechain
1	V	378	PHE	Sidechain
1	V	489	TYR	Sidechain

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	V	3855	0	3872	163	0
2	W	5348	0	5371	175	0
3	0	1479	0	1524	69	0
4	1	491	0	507	227	0
5	2	2196	0	2206	587	0
6	3	1526	0	1561	454	0
7	X	392	0	207	1	0
8	Y	401	0	231	0	0
All	All	15688	0	15479	1391	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 45.

The worst 5 of 1391 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)
3:0:77:LYS:HB2	3:0:83:CYS:CB	1.26	1.61

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:W:110:SER:HA	2:W:207:TYR:CE1	1.10	1.58
5:2:31:LEU:HD11	6:3:33:THR:CB	1.33	1.55
3:0:54:ARG:HG3	6:3:182:PHE:CE1	1.42	1.55
1:V:315:VAL:HG13	2:W:500:ASP:CB	1.21	1.54

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 PDB entries followed by that with respect to all EM entries.

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	V	473/782 (60%)	401 (85%)	47 (10%)	25 (5%)	2	29
2	W	661/760 (87%)	569 (86%)	68 (10%)	24 (4%)	4	38
3	0	186/395 (47%)	166 (89%)	15 (8%)	5 (3%)	6	45
4	1	60/71 (84%)	49 (82%)	9 (15%)	2 (3%)	5	40
5	2	264/462 (57%)	246 (93%)	14 (5%)	4 (2%)	13	57
6	3	187/308 (61%)	175 (94%)	10 (5%)	2 (1%)	17	63
All	All	1831/2778 (66%)	1606 (88%)	163 (9%)	62 (3%)	8	40

5 of 62 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	V	385	ASP
1	V	427	MET
1	V	461	HIS
1	V	491	ALA
1	V	499	ASN

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 PDB entries followed by that with respect to all EM entries.

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	V	422/688 (61%)	400 (95%)	22 (5%)	29	65
2	W	577/664 (87%)	540 (94%)	37 (6%)	22	58
3	0	171/352 (49%)	162 (95%)	9 (5%)	28	64
4	1	56/64 (88%)	51 (91%)	5 (9%)	12	44
5	2	238/399 (60%)	229 (96%)	9 (4%)	40	73
6	3	171/272 (63%)	157 (92%)	14 (8%)	14	49
All	All	1635/2439 (67%)	1539 (94%)	96 (6%)	29	61

5 of 96 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	W	461	LEU
2	W	620	MET
6	3	124	ILE
2	W	489	CYS
2	W	544	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 28 such sidechains are listed below:

Mol	Chain	Res	Type
5	2	117	ASN
5	2	202	GLN
6	3	185	GLN
5	2	161	HIS
5	2	181	GLN

5.3.3 RNA (i)

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.