Supplementary material for

Structure determination of the minimal complex between Tfb5 and Tfb2, two subunits of the yeast TFIIH transcription/DNA repair factor: a retrospective study

Denis E. Kainov, Vincent Cura, Marc Vitorino, Helène Nierengarten, Pierre Poussin, Bruno Kieffer, Jean Cavarelli, Arnaud Poterszman

Amino acid sequence of $Tfb2_{435-513}$, $Tfb2_{412-513}$ and $Tfb5_{2-72}$

$Tfb2_{435-513}$

80 residues, 670 non-hydrogen atoms

GPTVVDQIRL WQLELDRVIT YEGSLYSDFE TSQEYNLLSK YAQDIGVLLW
KDDKKKKFFI SKEGNSQVLD FAKRKLKKKQ

$Tfb2_{412-513}$

108 residues, 891 non-hydrogen atoms

GPHMASAEEK LEKKLELDPN CKEPLQVLPP TVVDQIRLWQ LELDRVITYE
GSLYSDFETS QEYNLLSKYA QDIGVLLWKD DKKKKFFISK EGNSQVLDFA
KRKLKKKQ

$Tfb5_{2-72}$

71 residues, 567 non-hydrogen atoms

ARARKGALVQ CDPSIKALIL QIDAKMSDIV LEELDDTHLL VNPSKVEFVK
HELNRLLSKN IYNPMDEEEN Q

Supplementary Table 1: Predicted and measured molecular weight

Constructs	Theoretical	Measured
$Tfb2_{435-513}$	9459.83	9459.86
$Tfb2_{435-513}$ (Se-Met)	9459.83	9459.86
$Tfb2_{412-513}$	12629.52	12629.08
$Tfb2_{412-513}$ (Se-Met)	12676.41	12676.99
$Tfb5_{2-72}$ (N-term M1 processed)	8101.28	8100.79
$Tfb5_{2-72}$ (N-term M1 processed, SeMet)	8195.06	8195.32

Supplementary Table 2: Summary of crystallographic data

	Hexagonal form	Orthorombic Form
Construct	$Tfb2_{435-513}$: $Tfb5_{2-72}$	$Tfb2_{412-513}$: $Tfb5_{2-72}$
SPG	$P6_1$	$P2_12_12_1$
Max res (Å)	1.7	2.6
Solvent content (%)	59	54
Vm (A^3/Da)	3.06	2.68
Tfb2 vis res	447-508	437-510 and 432-507
Tfb5 vis res	2-64	2-59 and 2-66

Supplementary Table 3: Crystal contacts in the orthorhombic and hexagonal crystal forms $^{i}N_{at}$ indicates the number of interfacing atoms in the corresponding structure and $^{i}N_{res}$ the number of interfacing residues. Symmetry op-n indicates the symmetry operation that should be applied to 2nd interfacing structure in order to obtain the respective interface. Interface area (Å^{2,}) calculated as difference in total accessible surface areas of isolated and interfacing structures divided by two. *, biological relevant interface; # crystallographic interface observed in both orthorhombic and hexagonal crystal forms; & crystallographic interface that bridges the two copies of the Tfb2C:Tfb5 complex leading to the formation of a probable heterotetramer. Data were generated with PISA (Krissinel and Henrick, 2007) from PDB 3DOM and 3DGP.

Interfaces in $P2_12_12_1$

		chain	$^{i}N_{at}$	$^{i}N_{res}$	chain	Symmetry op-n	$^{i}N_{at}$	$^{i}N_{res}$	Interface (A^2)
1	Tfb2/Tfb5 (*)	D	111	32	C	x,y,z	127	33	1152.4
2	Tfb2/Tfb5 (*)	B	110	29	A	x,y,z	122	29	1073.9
3	Tfb2/Tfb2 (#)	C	68	23	C	x-1/2,-y-1/2,-z	68	18	664.6
4	Tfb2/Tfb2 (#)	A	69	20	A	x-1/2,-y+1/2,-z	68	19	655.6
5	Tfb2/Tfb2 (&)	C	29	10	A	-x-1/2,-y,z-1/2	33	9	298.5
6	Tfb5/Tfb5	D	34	11	B	x,y,z	29	10	287.5
7	Tfb5/Tfb5 (&)	D	13	4	B	-x-1/2,-y,z-1/2	21	7	202.5
8	Tfb2/Tfb5 (&)	D	22	5	A	-x-1/2,-y,z-1/2	24	5	187.9
9	Tfb2/Tfb5 (&)	C	13	4	B	-x-1/2,-y,z-1/2	15	5	119.3
10	Tfb2/Tfb5	B	10	3	A	x-1/2,-y+1/2,-z	6	2	58.1
11	Tfb2/Tfb5	D	4	1	C	x-1/2,-y-1/2,-z	4	1	39.5
12	Tfb2/Tfb2	C	4	2	A	-x,y-1/2,-z-1/2	4	2	27.6

Interfaces in $P6_1$

		chain	$^{i}N_{at}$	$^{i}N_{res}$	chain	Symmetry op-n	$^{i}N_{at}$	$^{i}N_{res}$	Interface (A^2)
1	Tfb2/Tfb5 (*)	B	94	28	A	x,y,z	107	28	951.4
2	Tfb2/Tfb2 (#)	A	69	19	A	-x+1,-y+1,z-1/2	65	16	637.9
3	Tfb5/Tfb5	B	33	7	B	-y+1,x-y+1,z-2/3	27	5	270.4
4	Tfb2/Tfb5	B	18	6	A	-y+1,x-y+1,z-2/3	27	10	206.5
5	Tfb5/Tfb5	B	15	6	B	-y+1,x-y+1,z+1/3	18	6	149.5
6	Tfb2/Tfb5	B	4	2	A	-x+1,-y+1,z-1/2	7	3	44.3

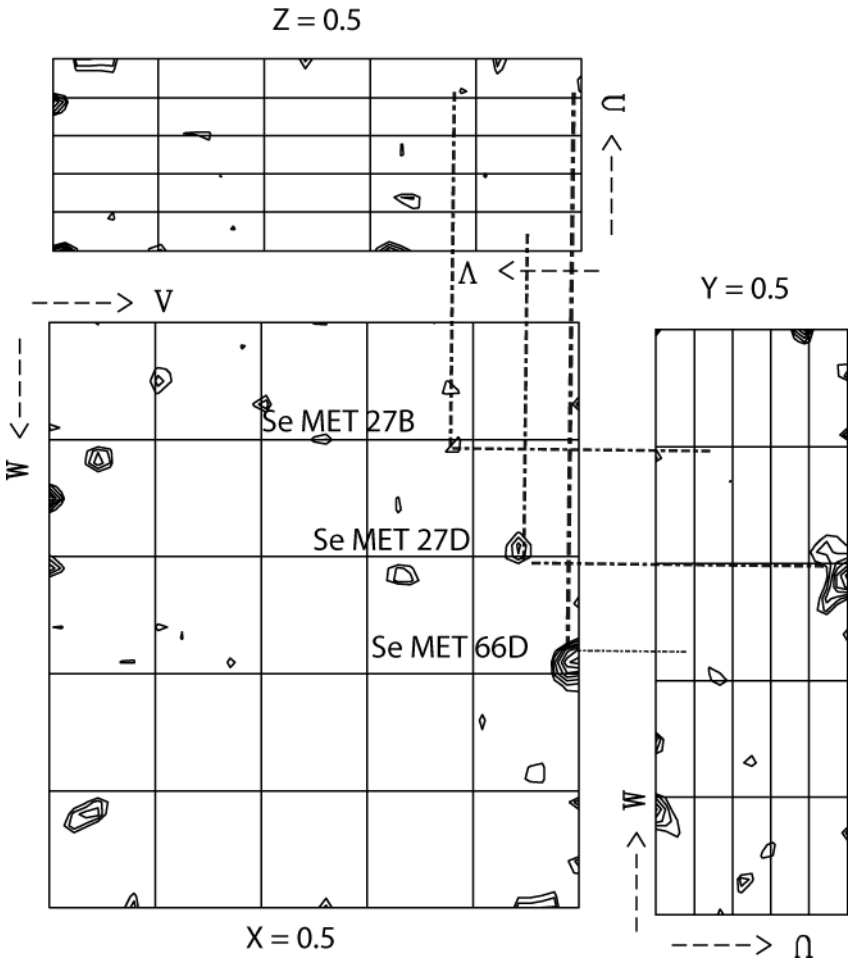
Supplementary Table 4: Selected interfaces from $Tfb2_{412-513}$: $Tfb5_{2-72}$ (Orthorhombic form). Analysis of interfaces 5, 8 and 9 from PDB 3DOM obtained with SPIDER (Kabsch and Sander, 1983; Porollo and Meller, 2007). Criteria used for the interaction site (interacting residue) assignment were: RSA change (loss) at least 4% and not less than 5Å^2 upon complex formation *, Total change in surface area of the interface for a given chain as calculated by DSSP (Kabsch and Sander, 1983); ** Hydrophobicity index (mean) of the interface according to the ARGP820101 scale from AAIndex database. With the exception of R442, L446 for interface 5 and I16 for interface 9 interacting residues are identical that those identified with PISA

Interface	Chains	Residues	ISA, Å^2*	HIP**
5	Tfb2/Tfb2 chainA/chainC	V437 D439 Q440 L443 W444 E447 R450	339	1.00±0.83
5	Tfb2/Tfb2 chainC/chainA	Q431 V432 L433 P434 P435 T436 V437 Q440 I441 W444	336	1.30±0.92
8	Tfb2/Tfb5 chainA/chainD	V437 V438 Q440 I441 W444	202	1.30±0.92
8	Tfb5/Tfb2 chainD/chainA	L19 N61 I62 Y63 P65	202	1.50±0.91
9	Tfb5/Tfb2 chainB/chainC	D13 P14 S15 A18	139	0.77±0.71
9	Tfb2/Tfb5 chainC/chainB	V432 P434 P435	122	1.74±0.30

Supplementary Table 5: Anomalous differences evaluated with SOLVE at the peak for Selenium ((λ= 0.98055) for $Tfb2_{435-513}$: $Tfb5_{2-72}$(hexagonal form).

shell	dmin	nobs	Fbar	R	scale	SIGNAL	NOISE	S/N
1	4.800	865	255.039	0.035	1.000	10.43	4.72	2.21
2	3.600	1244	290.113	0.025	0.999	7.23	5.31	1.36
3	3.360	473	244.546	0.022	1.000	4.43	5.11	0.87
4	3.180	478	205.551	0.024	1.000	3.59	5.19	0.69
5	3.000	581	174.831	0.029	0.999	3.63	5.08	0.71
6	2.880	469	144.575	0.031	0.999	1.14	5.48	0.21
7	2.760	565	131.252	0.037	1.000	1.69	5.68	0.30
8	2.640	652	115.836	0.043	1.001	0.77	6.08	0.13
9	2.520	780	107.862	0.045	1.000	0.00	6.54	0.00
10	2.400	585	98.790	0.058	1.000	0.00	7.38	0.00
Total:		6692	187.750	0.032	1.000	5.18	5.69	0.76

Supplementary Figure 1: Anomalous difference Patterson contoured at 3.0σ with 0.5 σ intervals, calculated at the peak wavelength for $Tfb2_{412-513}$: $Tfb5_{2-72}$(orthorhombic form) between 35 and 3.5 Å. Dashed lines show the expected positions of the harker peaks.



References

Kabsch, W., and Sander, C. (1983). Dictionary of Protein Secondary Structure - Pattern-Recognition of Hydrogen-Bonded and Geometrical Features. *Biopolymers* **22**, 2577-2637.

Krissinel, E., and Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774-797.

Porollo, A., and Meller, J. (2007). Versatile annotation and publication quality visualization of protein complexes using POLYVIEW-3D. *BMC Bioinformatics* **8**, 316.