Supplementary Information

Polycistronic co-expression and non-denaturing purification of Histone octamers

Yoonjung Shim[a], Ming-Rui Duan[b], Xuejing Chen[a], Michael J. Smerdon[b] and Jung-Hyun Min[a*]

[a]*Department of Chemistry, University of Illinois at Chicago, Chicago, IL 60607, USA*

[b]*Department of Biochemistry and Biophysics, School of Molecular Biosciences, Washington State University, Pullman, WA 99164, , USA*

Table of Contents

1. Methods

Cloning of histones and polycistronic plasmid construction. Xenopus laevis histone genes were amplified by PCR from the plasmids originally made by the Luger laboratory [6]. To construct a polycistronic plasmid as depicted in **Fig. 1B**, we incorporated RBS along with a spacer sequence into the primers for H2B, H3 and H4 (**Table S1**) [14]. A hexahistidine (His_6) tag was also added to H2A N-terminus between pre-existing S-tag and thrombin site through PCR. The stop codon in H4 C-terminus was deleted, and an in-frame thrombin site was added between H4 and a pre-existing His_6-tag in a similar fashion (**Table S1**). The primers were purified using standard urea-PAGE protocols [16]. We then inserted these genes into the pET29a vector using standard cloning techniques using restriction enzymes as indicated in **Table S1 and S2**.

Histone expression. The resulting plasmid encoding all four core histones was transformed into BL21(DE3)pLysS cells and plated on to an LB agar plate containing kanamycin (50 μg/ml) and chloramphenicol (25 μg/ml). The plate was incubated overnight (~16 h) at 37 °C. For a 1 L scale preparation, one colony was inoculated into 10 mL 2xTY media containing kanamycin (50 μg/ml) and chloramphenicol (25 μg/ml). This starter culture was shaken at 170 rpm for ~4-5 h at 37 °C until slightly cloudy. Subsequently, the culture was amplified into 1 L of the same media and was grown for another 6-7 h at 37 °C. When the OD600 reaches ~0.4, histone co-expression was induced by adding 0.4 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). The culture was further shaken at 170 rpm at 37 °C for overnight (~20 h). Cells were harvested by centrifugation at 4,500 x g for 10 min at 4 °C. Cell pellets were processed immediately or stored at -80 °C for future purification.

Cell lysis and Nickel-affinity chromatography. Cell pellets were resuspended in 60 ml of lysis buffer (20 mM Tris-HCl pH 8.0, 2.0 M sodium chloride (NaCl), 1 mM phenylmethanesulfonylfluoride or Pefabloc SC (Centerchem) and 0.5 mM tris(2-carboxyethyl)phosphine (TCEP)). Resuspended cells were lysed by EmulsiFlex-C3 high pressure

homogenizer (Avestin) and clarified by centrifugation at 38,500 x g at 4 °C for 1 h. The supernatant was collected and imidazole stock solution was added to adjust the concentration to 30 mM. The clarified lysate was then loaded onto a 5 ml HisTrap FF column (GE Healthcare) pre-equilibrated in the Ni-buffer A (20 mM Tris-HCl pH 8.0, 2.0 M sodium chloride, 0.5 mM TCEP) followed by a 10 column volumes (CV) was with Ni-buffer A containing 30 mM imidazole and a 10 CV of Ni-buffer A. The bound proteins were eluted by increasing the imidazole concentration from 30 mM to 500 mM linearly over 23.5 CV. Each faction was analysed by 18% SDS-PAGE: 10 μl of sample was mixed with the same volume of 2x SDS loading buffer, incubated at 98 °C for 10 min and centrifuged at 17,000 x g for 10 min, after which all supernatant was loaded on the gel. Electrophoresis was run at 150 V for 80 min.

Thrombin digestion. Thrombin digestion was carried out by adding purified thrombin (Sigma) in 25:1 mass ratio and incubating the samples at 4 °C for ~ 3 h. The digestion was confirmed by SDS-PAGE.

Size exclusion chromatography. Thrombin-digested histones were then concentrated up to 3 mg/ml with ultrafiltration using Amicon YM50 membrane (MWCO 50 kDa) at 4 °C (Millipore). The concentrated sample was then injected onto a Superdex 200 10/300GL column. The histone octamer peak was eluted at an elution volume of 12.8 ml. The peak fractions were pooled and concentrated up to 8 mg/ml, aliquotted and flash-frozen in the presence of 50% glycerol for long-term storage.

Nucleosome reconstitution. Nucleosome reconstitution was done essentially as described [7; 8]. The histones and DNA were mixed in the molar ratio of octamer:DNA at 1.1:1 and dialyzed sequentially against TE buffers (10 mM Tris-HCl pH 8.0 and 1 mM EDTA) containing 1 M, 600 mM and 0 mM NaCl, each for at least one hour. The last dialysis was usually done overnight.

The results were verified on 5% native polyacrylamide gels run in 0.25x TBE at 130 V for 1 h at room temperature.

FRET measurements. The DNA substrate was internally labeled at positions 6 and 81 in a 601 sequence with Cy3 and Cy5 respectively [10]. The FRET by the dye pairs was measured as described to validate the formation of NCP by the histones [10].

2. Entire sequence of the histone-encoding polycistronic vector used in this study, depicted in Fig. 1B.

TGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTT
CTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACT
TGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAAC
AACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAA
CAAAATATTAACGTTTACAATTTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAATTA
ATTCTTAGAAAAACTCATCGAGCATCAAATGAAACTGCAATTTATTCATATCAGGATTATCAATACCATATTTTTGAAAAAGCCGTTTCTGTAATGAAGGAGAAAACTCAC
CGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTCTGCGATTCCGACTCGTCCAACATCAATACAACCTATTAATTTCCCCTCGTCAAAAATAAGGTTATCAAGTG
AGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGGCAAAAGTTTATGCATTTCTTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAAAATCACTCG
CATCAACCAAACCGTTATTCATTCGTGATTGCGCCTGAGCGAGACGAAATACGCGATCGCTGTTAAAAGGACAATTACAAACAGGAATCGAATGCAACCGGCGCAGGAACA
CTGCCAGCGCATCAACAATATTTTCACCTGAATCAGGATATTCTTCTAATACCTGGAATGCTGTTTTCCCGGGGATCGCAGTGGTGAGTAACCATGCATCATCAGGAGTAC
GGATAAAATGCTTGATGGTCGGAAGAGGCATAAATTCCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAACATCATTGGCAACGCTACCTTTGCCATGTTTCAGAAACA
ACTCTGGCGCATCGGGCTTCCCATACAATCGATAGATTGTCGCACCTGATTGCCCGACATTATCGCGAGCCCATTTATACCCATATAAATCAGCATCCATGTTGGAATTTA
ATCGCGGCCTAGAGCAAGACGTTTCCCGTTGAATATGGCTCATAACACCCCTTGTATTACTGTTTATGTAAGCAGACAGTTTTATTGTTCATGACCAAAATCCCTTAACGT
GAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTA
CCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTA
GGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCA
AGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAG
CTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGG
TATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTA
CGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGC
AGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATATATGGTGCA
CTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATACACTCCGCTATCGCTACGTGACTGGGTCATGGCTGCGCCCCGACACCCGCCAACACCCGCTGACGC
GCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGTCATCACCGAAACGCGCGAGGCA
GCTGCGGTAAAGCTCATCAGCGTGGTCGTGAAGCGATTCACAGATGTCTGCCTGTTCATCCGCGTCCAGCTCGTTGAGTTTCTCCAGAAGCGTTAATGTCTGGCTTCTGAT
AAAGCGGGCCATGTTAAGGGCGGTTTTTTCCTGTTTGGTCACTGATGCCTCCGTGTAAGGGGGATTTCTGTTCATGGGGGTAATGATACCGATGAAACGAGAGAGGATGCT
CACGATACGGGTTACTGATGATGAACATGCCCGGTTACTGGAACGTTGTGAGGGTAAACAACTGGCGGTATGGATGCGGCGGGACCAGAGAAAAATCACTCAGGGTCAATG
CCAGCGCTTCGTTAATACAGATGTAGGTGTTCCACAGGGTAGCCAGCAGCATCCTGCGATGCAGATCCGGAACATAATGGTGCAGGGCGCTGACTTCCGCGTTTCCAGACT
TTACGAAACACGGAAACCGAAGACCATTCATGTTGTTGCTCAGGTCGCAGACGTTTTGCAGCAGCAGTCGCTTCACGTTCGCTCGCGTATCGGTGATTCATTCTGCTAACC
AGTAAGGCAACCCCGCCAGCCTAGCCGGGTCCTCAACGACAGGAGCACGATCATGCGCACCCGTGGGGCCGCCATGCCGGCGATAATGGCCTGCTTCTCGCCGAAACGTTT
GGTGGCGGGACCAGTGACGAAGGCTTGAGCGAGGGCGTGCAAGATTCCGAATACCGCAAGCGACAGGCCGATCATCGTCGCGCTCCAGCGAAAGCGGTCCTCGCCGAAAAT
GACCCAGAGCGCTGCCGGCACCTGTCCTACGAGTTGCATGATAAAGAAGACAGTCATAAGTGCGGCGACGATAGTCATGCCCCGCGCCCACCGGAAGGAGCTGACTGGGTT
GAAGGCTCTCAAGGGCATCGGTCGAGATCCCGGTGCCTAATGAGTGAGCTAACTTACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCA
GCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCCAGGGTGGTTTTTCTTTTCACCAGTGAGACGGGCAACAGCTGATTGCCCTTCACCG
CCTGGCCCTGAGAGAGTTGCAGCAAGCGGTCCACGCTGGTTTGCCCCAGCAGGCGAAAATCCTGTTTGATGGTGGTTAACGGCGGGATATAACATGAGCTGTCTTCGGTAT
CGTCGTATCCCACTACCGAGATGTCCGCACCAACGCGCAGCCCGGACTCGGTAATGGCGCGCATTGCGCCCAGCGCCATCTGATCGTTGGCAACCAGCATCGCAGTGGGAA
CGATGCCCTCATTCAGCATTTGCATGGTTTGTTGAAAACCGGACATGGCACTCCAGTCGCCTTCCCGTTCCGCTATCGGCTGAATTTGATTGCGAGTGAGATATTTATGCC
AGCCAGCCAGACGCAGACGCGCCGAGACAGAACTTAATGGGCCCGCTAACAGCGCGATTTGCTGGTGACCCAATGCGACCAGATGCTCCACGCCCAGTCGCGTACCGTCTT
CATGGGAGAAAATAATACTGTTGATGGGTGTCTGGTCAGAGACATCAAGAAATAACGCCGGAACATTAGTGCAGGCAGCTTCCACAGCAATGGCATCCTGGTCATCCAGCG
GATAGTTAATGATCAGCCCACTGACGCGTTGCGCGAGAAGATTGTGCACCGCCGCTTTACAGGCTTCGACGCCGCTTCGTTCTACCATCGACACCACCACGCTGGCACCCA
GTTGATCGGCGCGAGATTTAATCGCCGCGACAATTTGCGACGGCGCGTGCAGGGCCAGACTGGAGGTGGCAACGCCAATCAGCAACGACTGTTTGCCCGCCAGTTGTTGTG
CCACGCGGTTGGGAATGTAATTCAGCTCCGCCATCGCCGCTTCCACTTTTTCCCGCGTTTTCGCAGAAACGTGGCTGGCCTGGTTCACCACGCGGGAAACGGTCTGATAAG
AGACACCGGCATACTCTGCGACATCGTATAACGTTACTGGTTTCACATTCACCACCCTGAATTGACTCTCTTCCGGGCGCTATCATGCCATACCGCGAAAGGTTTTGCGCC
ATTCGATGGTGTCCGGGATCTCGACGCTCTCCCTTATGCGACTCCTGCATTAGGAAGCAGCCCAGTAGTAGGTTGAGGCCGTTGAGCACCGCCGCCGCAAGGAATGGTGCA
TGCAAGGAGATGGCGCCCAACAGTCCCCCGGCCACGGGGCCTGCCACCATACCCACGCCGAAACAAGCGCTCATGAGCCCGAAGTGGCGAGCCCGATCTTCCCCATCGGTG
ATGTCGGCGATATAGGCGCCAGCAACCGCACCTGTGGCGCCGGTGATGCCGGCCACGATGCGTCCGGCGTAGAGGATCGAGATCGATCTCGATCCCGCGAAATTAATACGA
CTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGAAAGAAACCGCTGCTGCTAAATTCGAAC
GCCAGCACATGGACAGCCCAGATCTGCACCACCACCACCACCAGGTACCCTGGTGCCACGCGGTTCCATGGgtATGTCAGGAAGAGGCAAACAAGGCGGTAAAACCCGCG
CTAAGGCCAAGACTCGCTCATCTCGGGCTGGGCTACAGTTCCCTGTTGGCCGTGTTCACCGGCTGTTAAGGAAAGGCAATTATGCAGAGCGGGTGGGAGCTGGAGCTCCAG
TCTATCTGGCTGCAGTGTTGGAGTATCTGACCGCTGAGATTTTGGAATTGGCCGGGAATGCGGCCCGTGATAACAAGAAGACTCGCATTATCCCCAGACACCTGCAGCTCG
CTGTGCGCAACGATGAGGAACTGAACAAACTGCTCGGAAGAGTCACTATCGCTCAGGGCGGGGTCCTGCCCAACATCCAGTCCGTGCTGCTGCCCAAGAAAACCGAGAGTT
CCAAGTCGGCCAAGAGCAAGTGAGAATTCAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGGCCAAGTCCGCTCCAGCCCCGAAGAAAGGCTCCAAGAAAGCGG
TGACCAAGACTCAGAAGAAAGACGGGAAAAAGCGCAGGAAGACAAGGAAGGAGAGTTATGCCATTTACGTGTACAAGGTGCTGAAGCAGGTGCACCCCGATACCGGCATCT
CGTCCAAGGCCATGAGCATCATGAACTCCTTTGTCAACGATGTGTTTGAGCGCATCGCAGGGGAAGCCTCCCGCCTGGCTCATTACAACAAGCGCTCCACCATCACCTCCC
GGGAGATCCAGACCGCGGTCCGACTGCTGCTGCCTGGGGAGTTGGCCAAACACGCCGTGTCCGAGGGCACCAAGGCTGTCACCAAGTACACCAGCGCCAAGTAAGTCGACA
ATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGGCCCGTACCAAGCAGACCGCCCGTAAATCCACCGGAGGGAAGGCTCCCCGCAAGCAGCTGGCCACCAAGGCAG
CCAGGAAGTCCGCTCCTGCTACCGGCGGAGTCAAGAAACCTCACCGTTACCGGCCCGGCACAGTCGCTCTCCGCGAGATCCGCCGCTACCAGAAATCCACCGAGCTGCTCA
TCCGCAAACTGCCTTTCCAGCGCCTGGTCCGGGAGATCGCTCAGGACTTCAAGACCGACCTGCGCTTCCAGAGCTCGGCCGTTATGGCTCTGCAGGAGGCCAGCGAGGCTT
ATCTGGTCGCTCTCTTTGAGGACACCAACCTGTGCGCCATCCACGCCAAGAGGGTCACCATCATGCCCAAGGACATCCAGCTGGCCCGCAGAATCCGAGGCGAGAGGGCTT
AGGCGGCCGCAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGTCTGGTCGTGGTAAAGGTGGTAAAGGTCTGGGTAAAGGTGGTGCTAAACGTCACCGTAAAGT
TCTGCGTGACAACATCCAGGGTATCACCAAGCCGGCTATCCGTCGTCTGGCTCGTCGTGGTGGTGTTAAACGTATCTCCGGTCTGATCTACGAAGAAACCCGCGGTGTTCT
GAAAGTTTTCCTGGAAAACGTTATCCGTGACGCTGTTACCTACACCGAACACGCTAAACGTAAAACCGTTACCGCTATGGACGTTGTTTACGCTCTGAAACGTCAGGGTCG
TACCCTGTACGGTTTCGGTGGTAGCAGCGGCCTGGTGCCGCGCGGCAGCCTCGAGCACCACCACCACCACCACTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGA
GTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGAT

3. Supplementary Table 1 &2

Table S1. PCR primers used in this study.

Name	Sequence
BglII_ His_6_KpnI_F	C AGCCCAGATCTG CACCACCACCACCACCAC GGTACC CTGGTG
BglII_ His_6_KpnI_R	CACCAG GGTACC GTGGTGGTGGTGGTGGTG CAGATCTGGGCT G
NcoI_xeH2A_F	acttga CCATGGgt ATGTCAGGAAGAGGCAAA
EcoRI_xeH2A_R	cacaca GAATTC TCACTTGCTCTTGGCCGA
EcoRI_ RBS_xeH2B_F	acttga GAATTC AATAATTTTGTTTAACTT TAAGAAGGAGATATACAT ATGGCCAAGTCCGCTCCA
SalI_xeH2B_R	cacaca GTCGAC TTACTTGGCGCTGGTGTA
SalI_ RBS_xeH3_F	acttga GTCGAC AATAATTTTGTTTAACTT TAAGAAGGAGATATACAT ATGGCCCGTACCAAGCAG
NotI_xeH3_R	cacaca GCGGCCGC CTAAGCCCTCTCGCCTCG
NotI_ RBS_xeH4_F	acttga GCGGCCGC AATAATTTTGTTTAACTT TAAGAAGGAGATATACAT ATGTCTGGTCGTGGTAAA
XhoI_Thrombin-xeH4_R	cacaca CTCGAG GCTGCCGCGCGGCACCAGGCCGCTGCT ACCACCGAAACCGTACAG

*Restriction enzyme sites are underlined and RBS's are indicated in red.

Table S2. Restriction enzymes and affinity tags used for the histones.

Histone	(5') Restriction Enzyme	(3') Restriction Enzyme	Affinity tags
H2A	NcoI	EcoRI	An N-terminal S-tag followed by a His_6-tag and a thrombin cleavage site.
H2B	EcoRI	SalI	None
H3	SalI	NotI	None
H4	NotI	XhoI	A C-terminal His_6-tag preceded by a thrombin site.

4. Supplementary Figures 1-3

Fig. S1. Co-expression and purification of histone octamers. **(A)** All four histones were induced and overexpressed upon adding 0.4 mM IPTG. Induced and uninduced cells are indicated by (+) and (–) IPTG, respectively. M: molecular weight marker. (**B**) Most histones were recovered in the soluble fraction of the cell lysate. C: total cell lysate, S: cell supernatant, P: cell pellet fraction. All lanes represent 50 μl of cell culture. **(C)** The affinity tags on H2A and H4 could be digested completely within 3 hours of thrombin addition at 4 °C. Undigested and digested samples are indicated with (-) and (+) thrombin, respecitively. Numbers above the lanes 3-5 indicate the number of hours after thrombin addition.

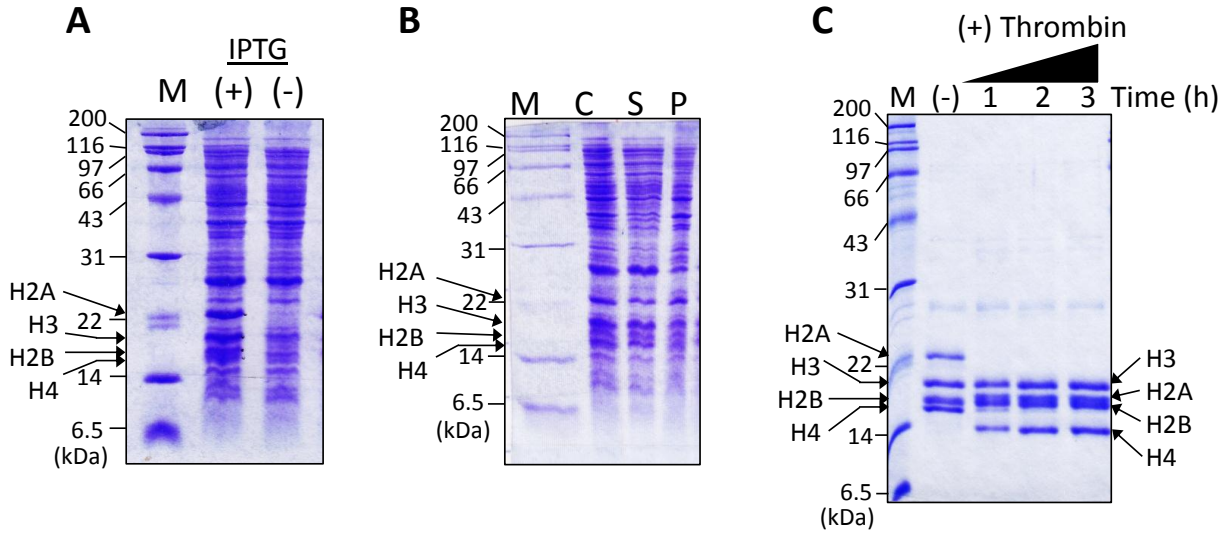

Fig. S2. Biochemical characterization of purified histones. **(A)** Size-exclusion chromatogram of purified histone products. Histones elute at 12.8 ml as in Fig. 2E with stoichiometric ratio of all histones. The void volume peak observed in Fig. 2E no longer exists. **(B)** Calibration curve for the Superdex 200 column used in this study. According to this curve, the molecular weight of the histone complexes eluting at 12.8 ml is 111 kDa, closely matching that of the octamer (109 kDa). K_{av} and Log MW were calculated as described in the product manual for the size-exclusion column molecular weight calibration kit (GE Healthcare). **(C)** SDS-PAGE gel of the histone fractions shown in (B). M: molecular weight marker, I: input, 7-14 indicate fraction numbers. **(D)** SDS-PAGE of histone octamers prepared by a conventional method [10] and the histones purified in Fig. 2.

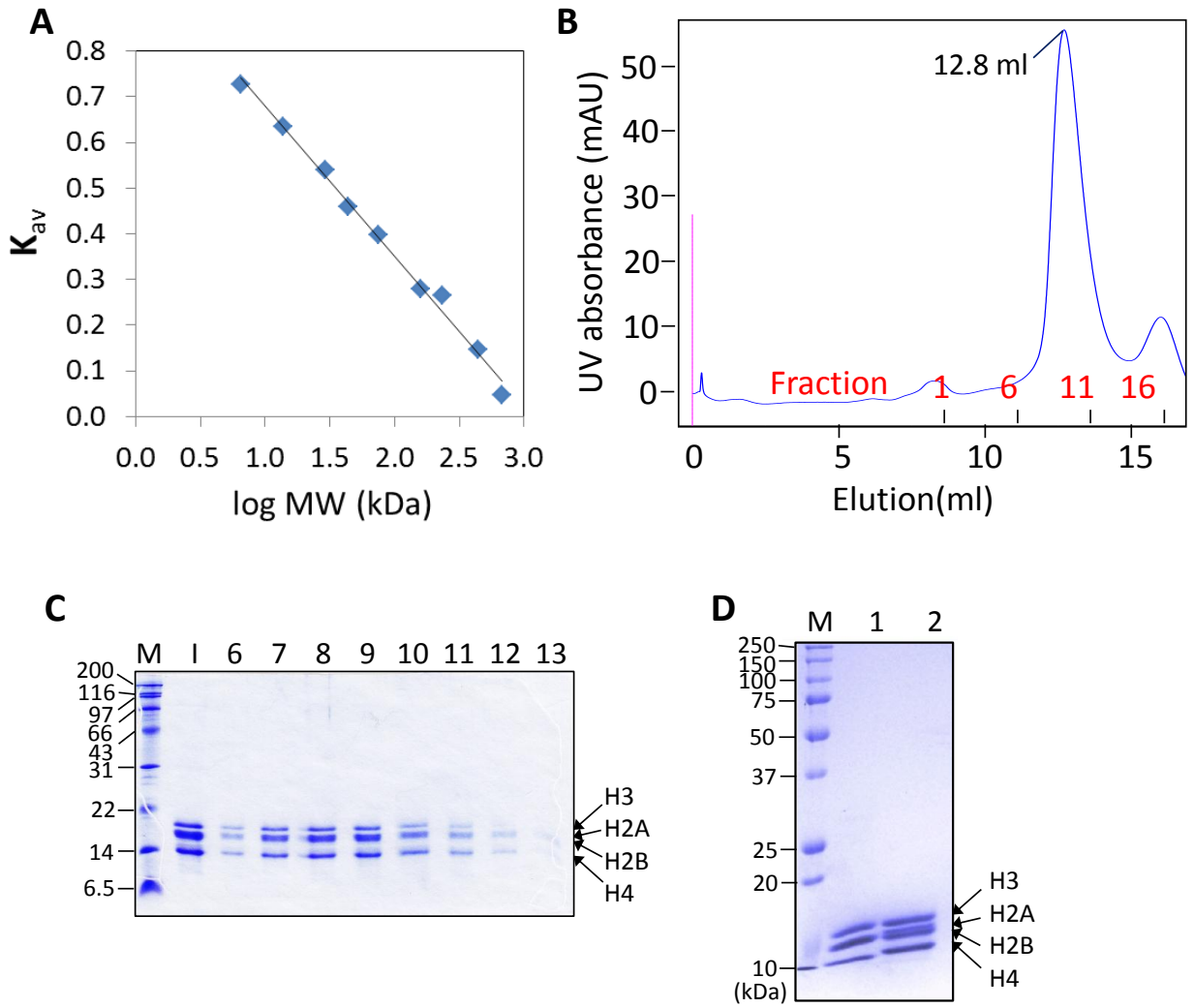

Fig. S3. Schematic diagrams of the locations of donor (Cy3) and acceptor (Cy5) on a naked DNA (**A**) and a nucleosome (**B**). The formation of NCPs with the purified histones can be confirmed by FRET peak of the acceptor as shown in Fig. 2B.