Свойства белков, способствующие жидко-фазному разделению хроматина

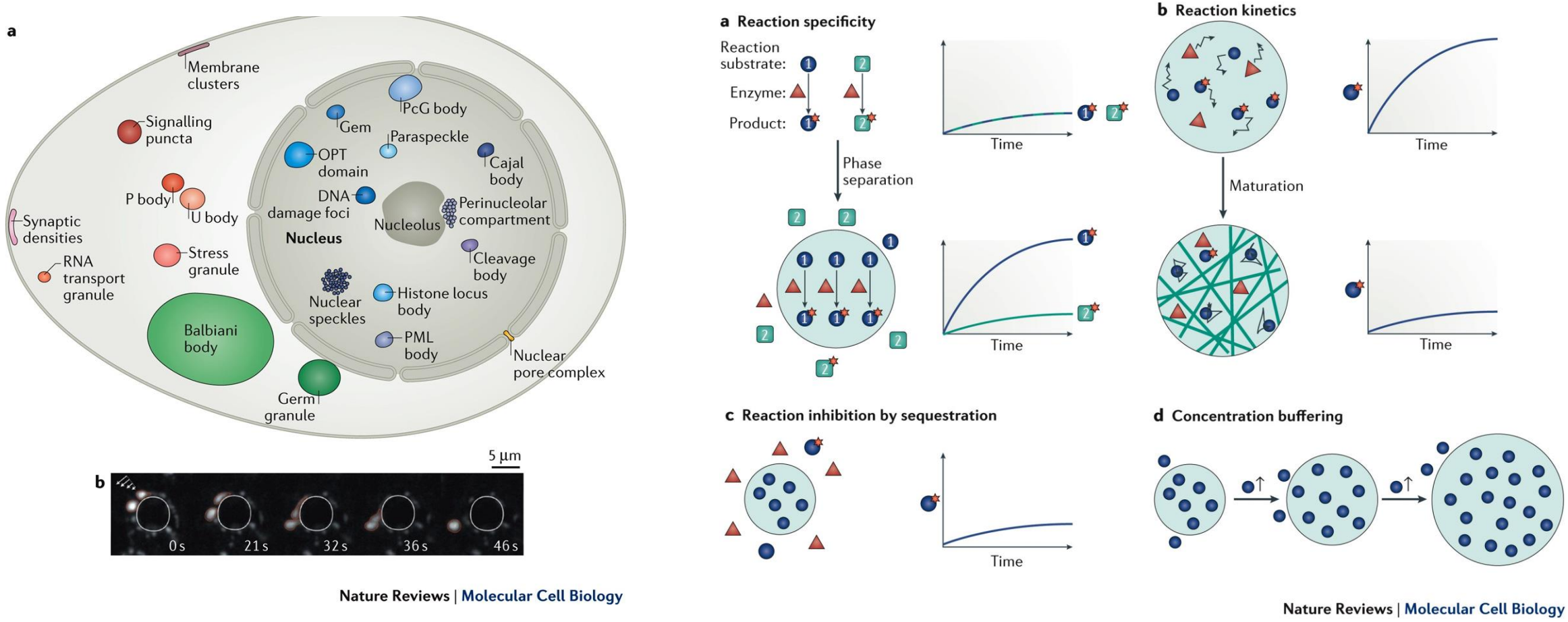
Анна Грибкова,

03.03.2021

План

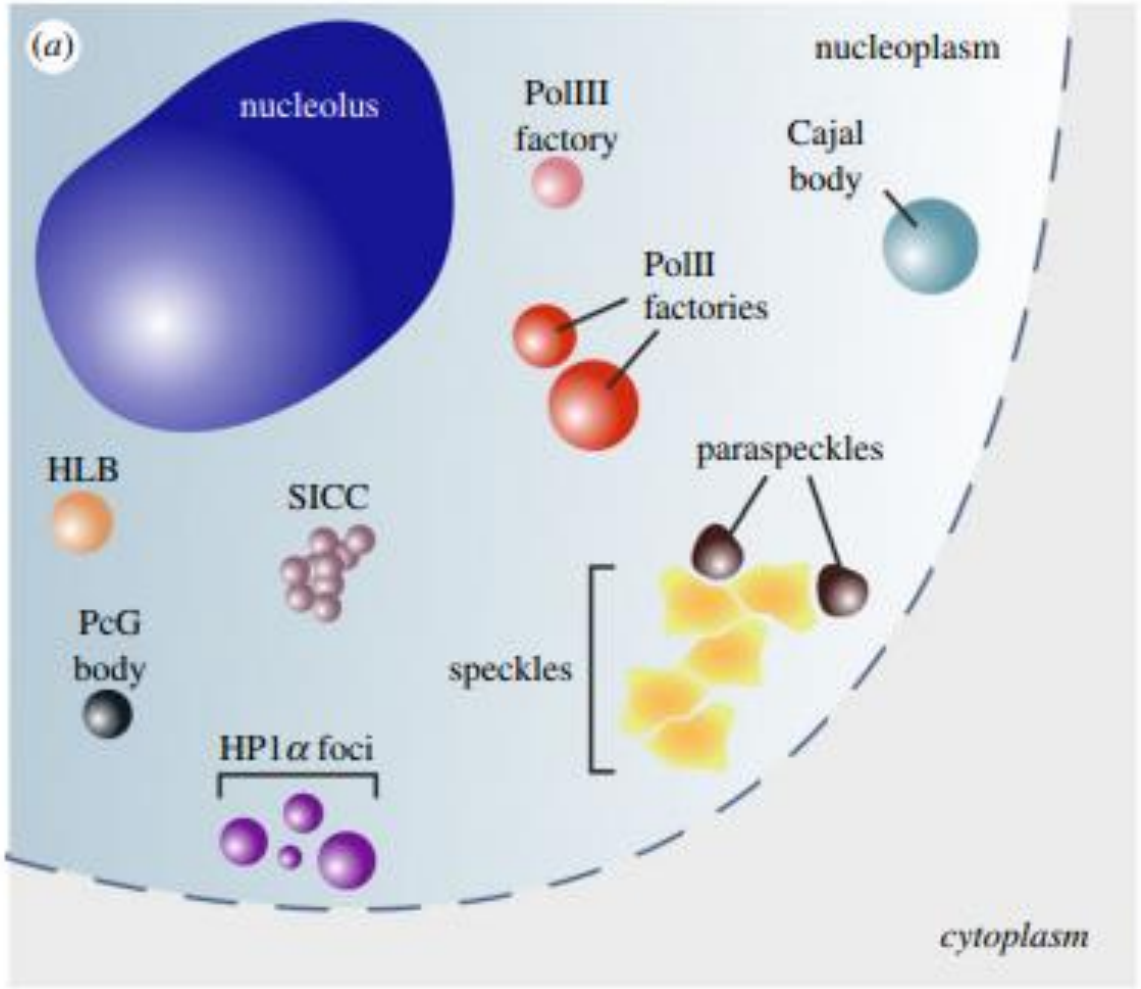
1. Немембранные органеллы
2. Жидко-фазное разделение в хроматине
3. Acetylated chromatin only phase separates upon binding multi-bromodomain proteins (Gibson)
4. Что влияет на разделение фаз (кроме длины линкера, гистона H1): Физико-химические свойства белков, доменная организация белков хроматина, IDR
5. HP1α, H3K9me – heterochromatin
6. Chromodomain
7. [алгоритмы предсказания белков, которые способствую разделению фаз]

Немембранные органеллы



Salman et al., 2017

Ядро



Spiros Palikyras and Argyris Papantonis, 2019

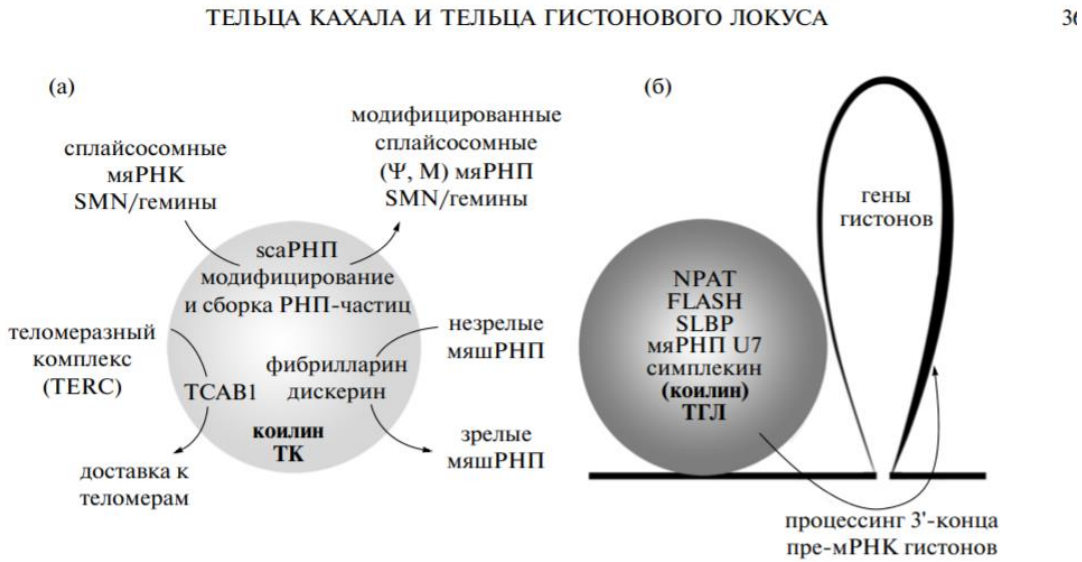
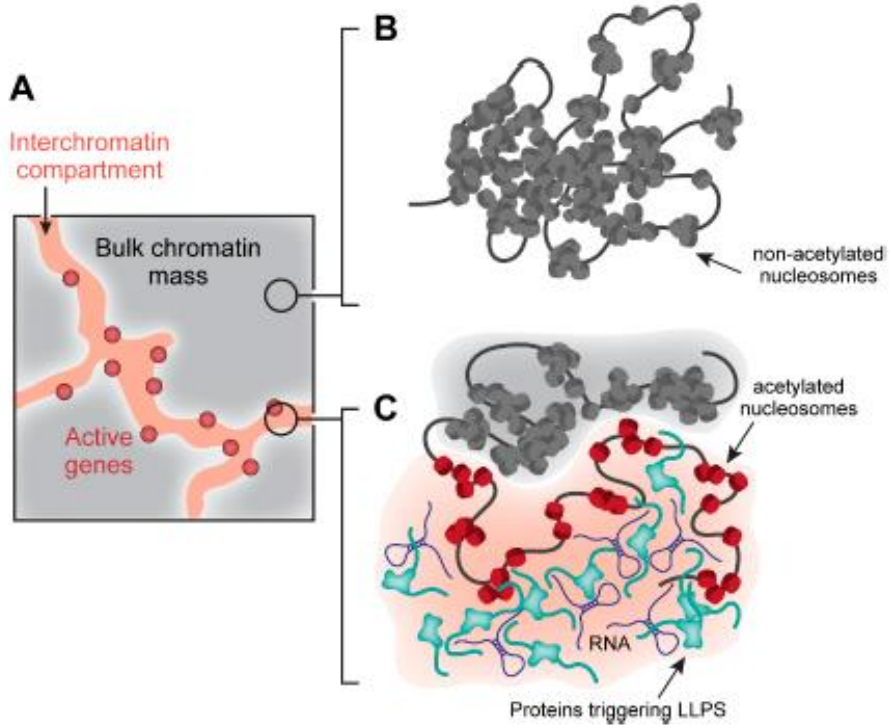
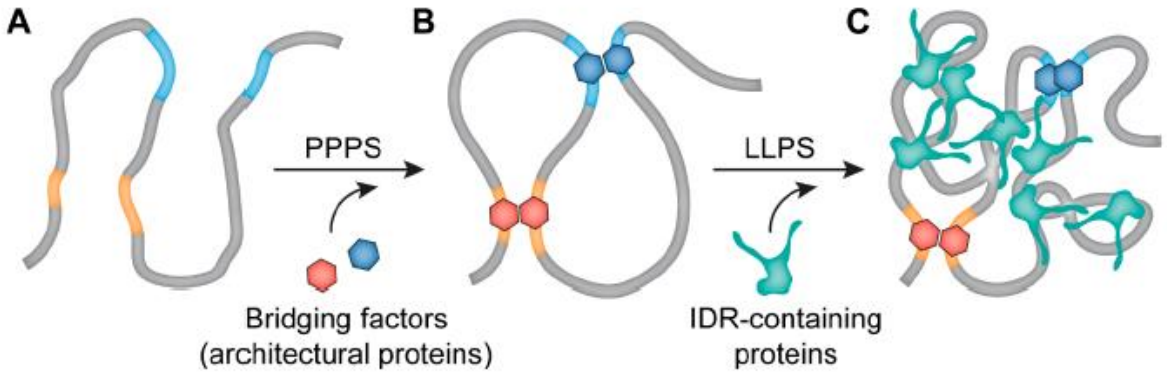


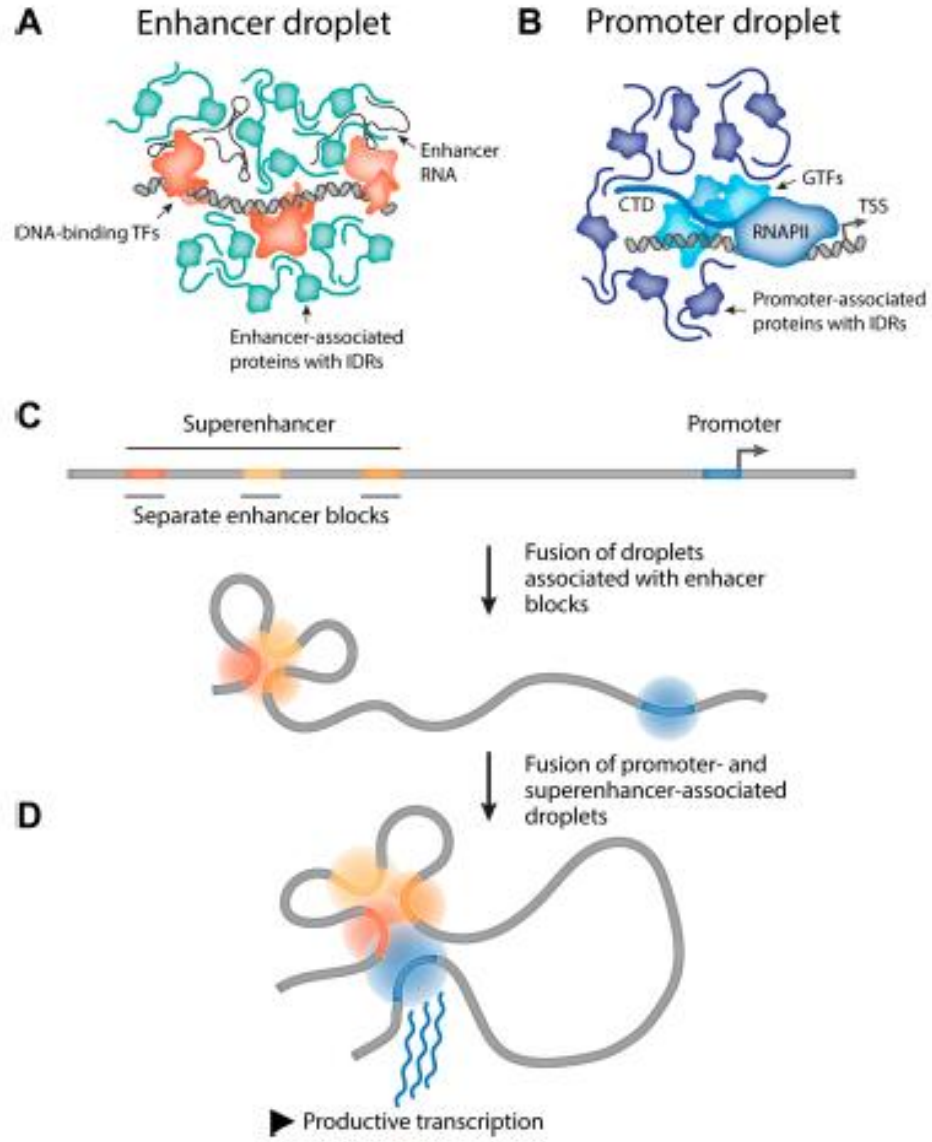
Рис. 1. Сравнительная схема телец Кахала и телец гистонового локуса соматических клеток животных. (а) Тельце Кахала (ТК) принимает участие в модификации сплайсосомных мяРНК (метилировании (М) и псевдоуридинилировании (Ψ)) и сборке мяРНП, созревании мяшРНП, определяет доставку теломеразного комплекса к теломерам. (б) Тельце гистонового локуса (ТГЛ) участвует в динамике факторов процессинга 3'-конца пре-мРНК гистонов. ТГЛ ассоциировано с кластером генов гистонов.

Iain A. Sawyer et al., 2018

Хроматин

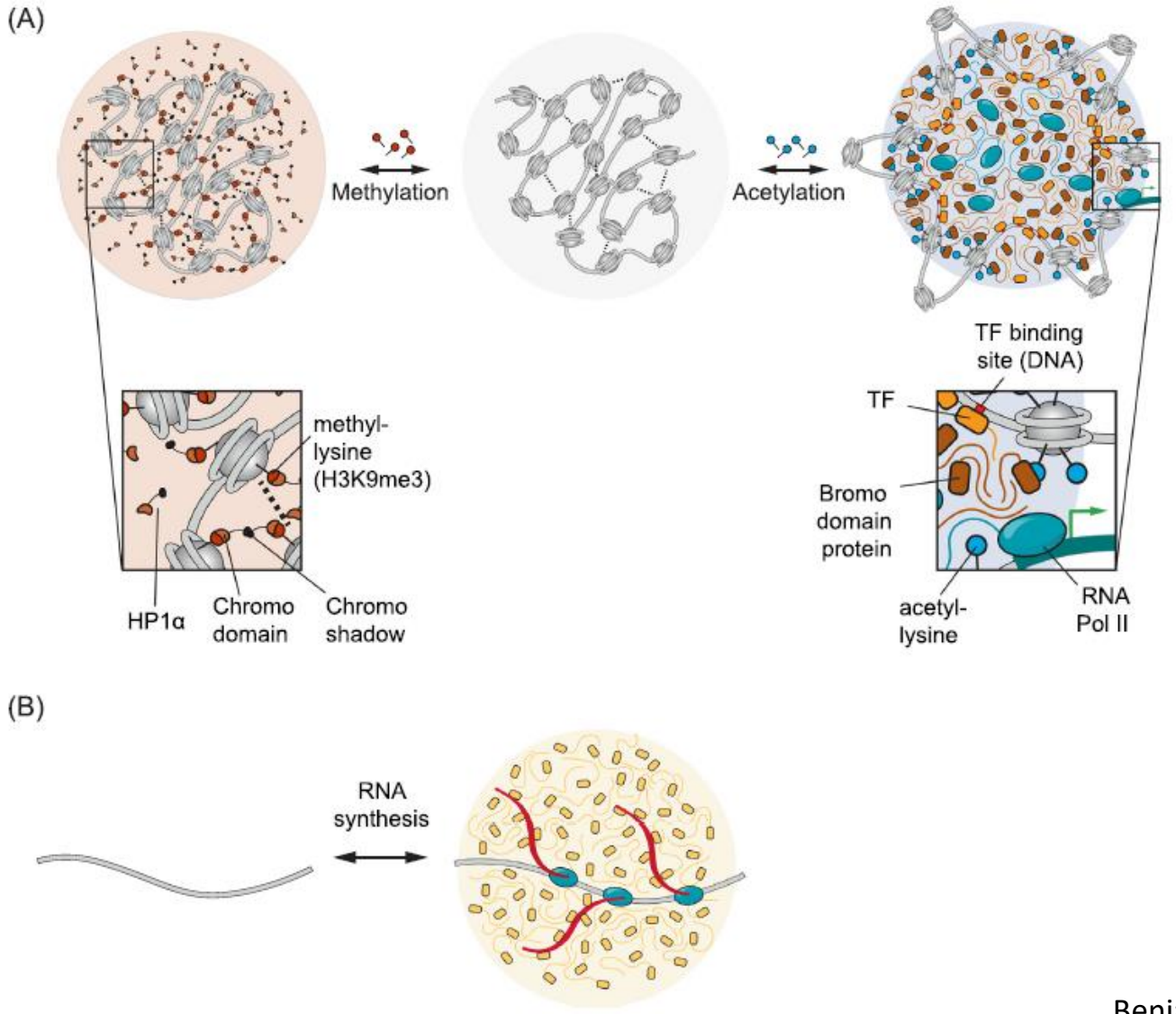






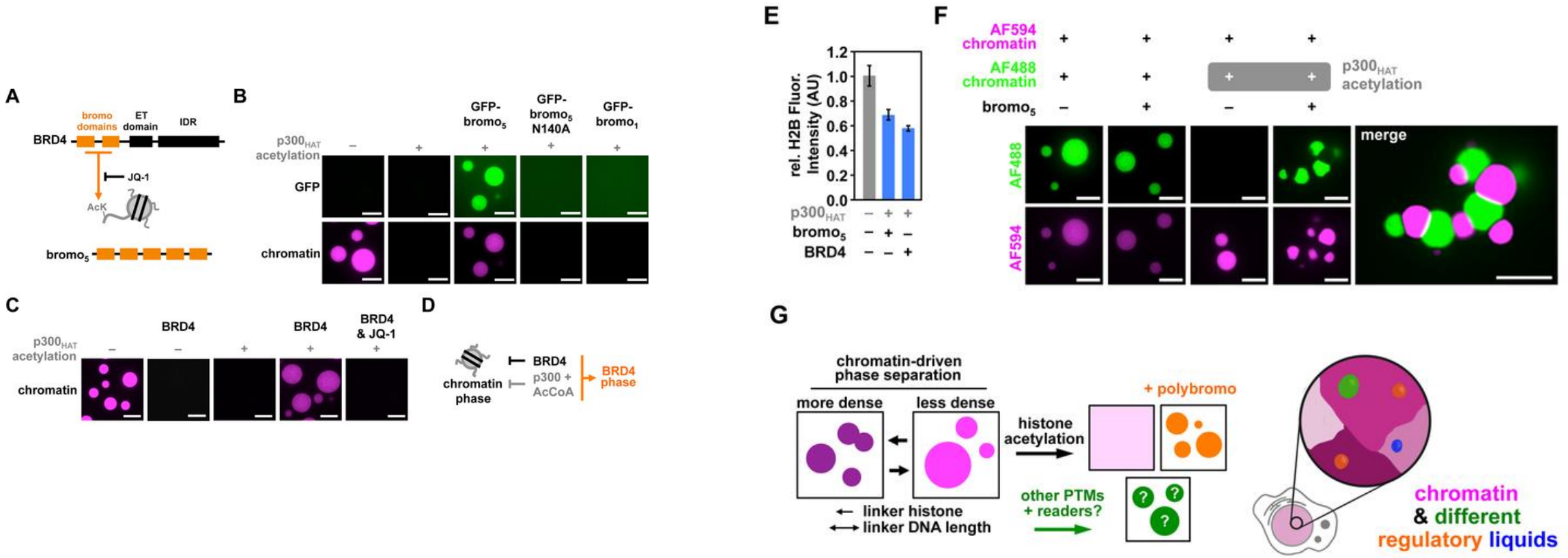
Sergey Razin and Sergey Ulianov, 2020

Some Examples of Nuclear Condensate Regulation



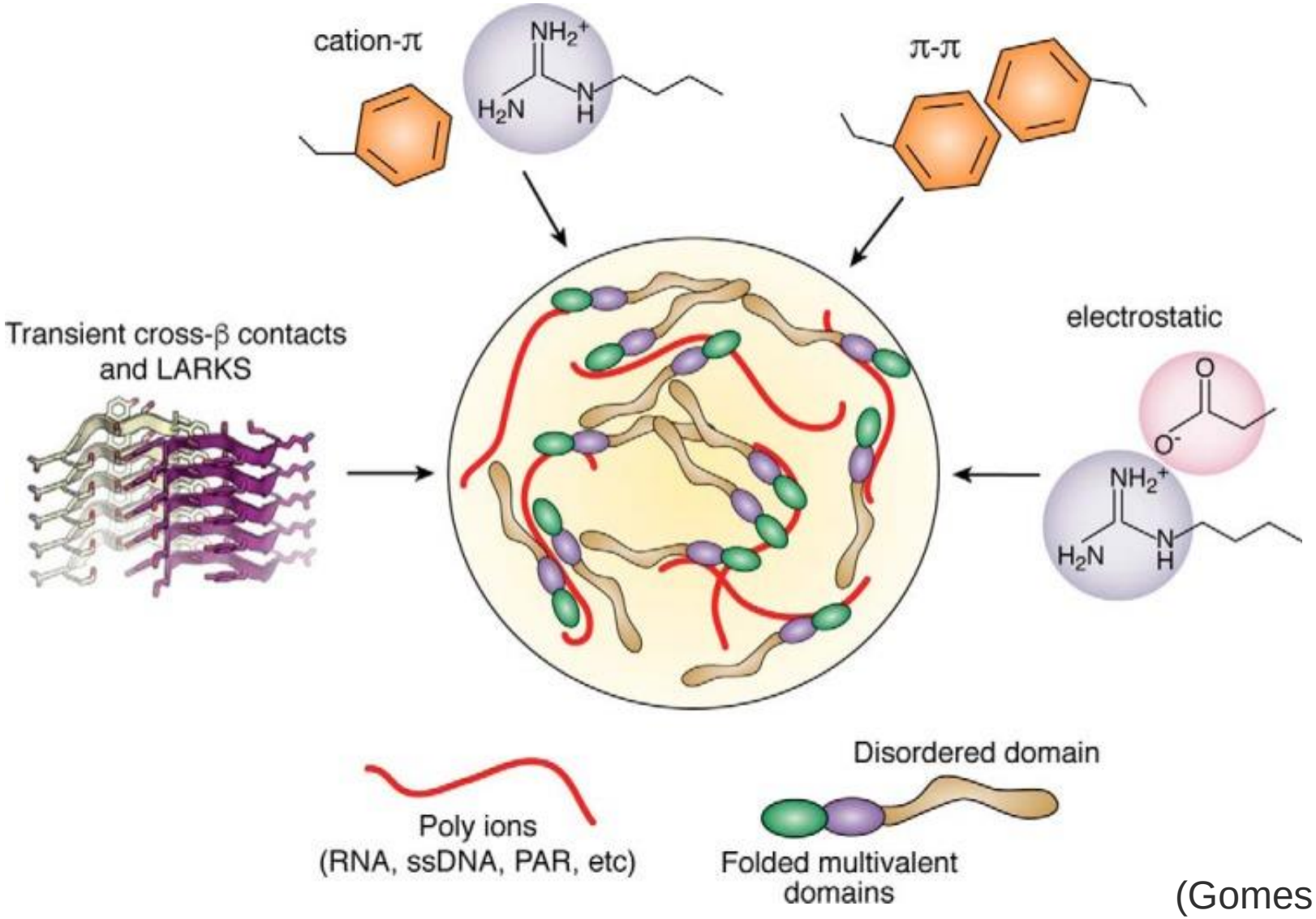
Benjamin R. Sabari et al., 2020

BRD4 promotes a new liquid phase of acetylated chromatin



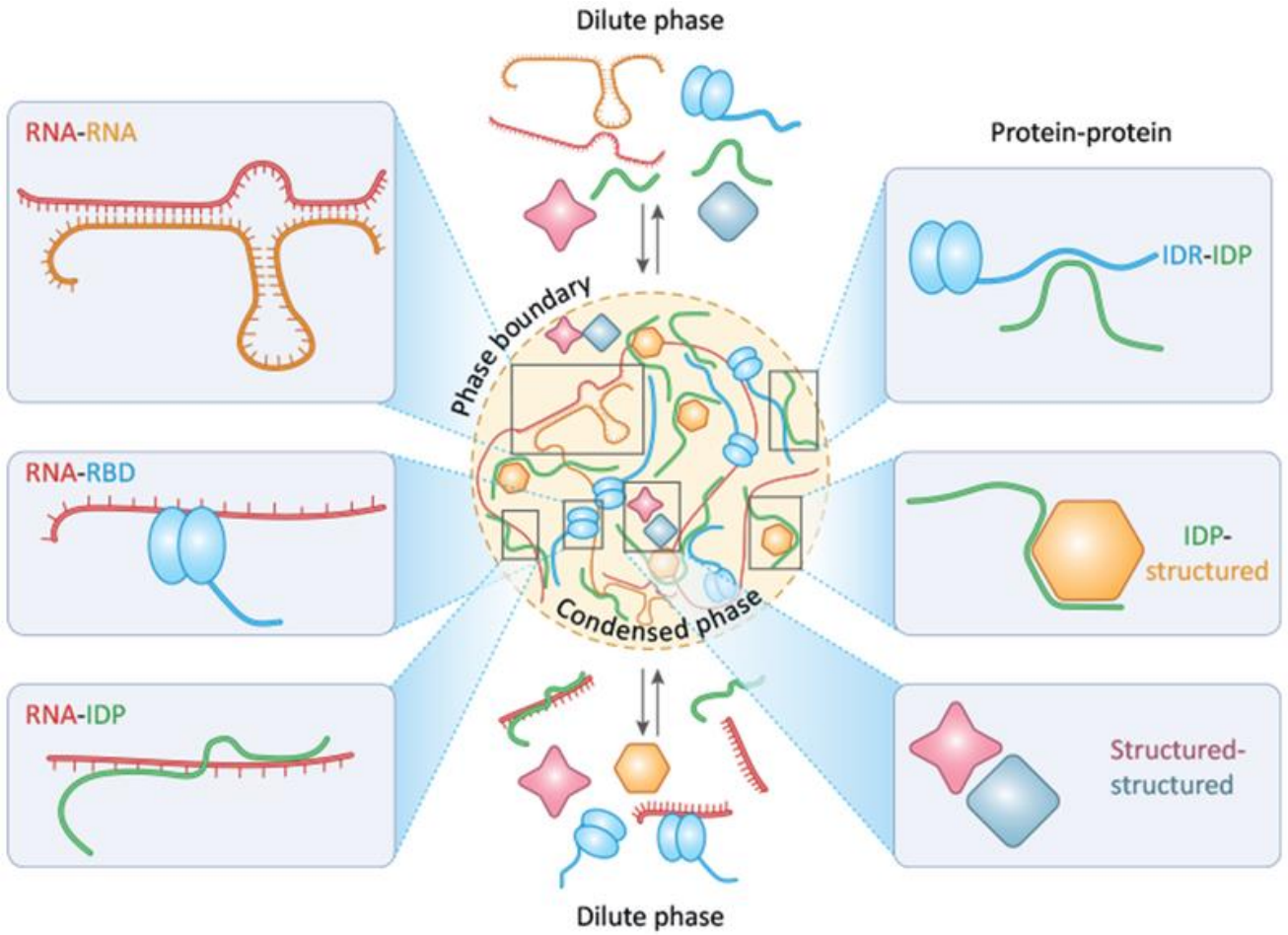
(Gibson et al. 2019)

Формирование немембранных органелл



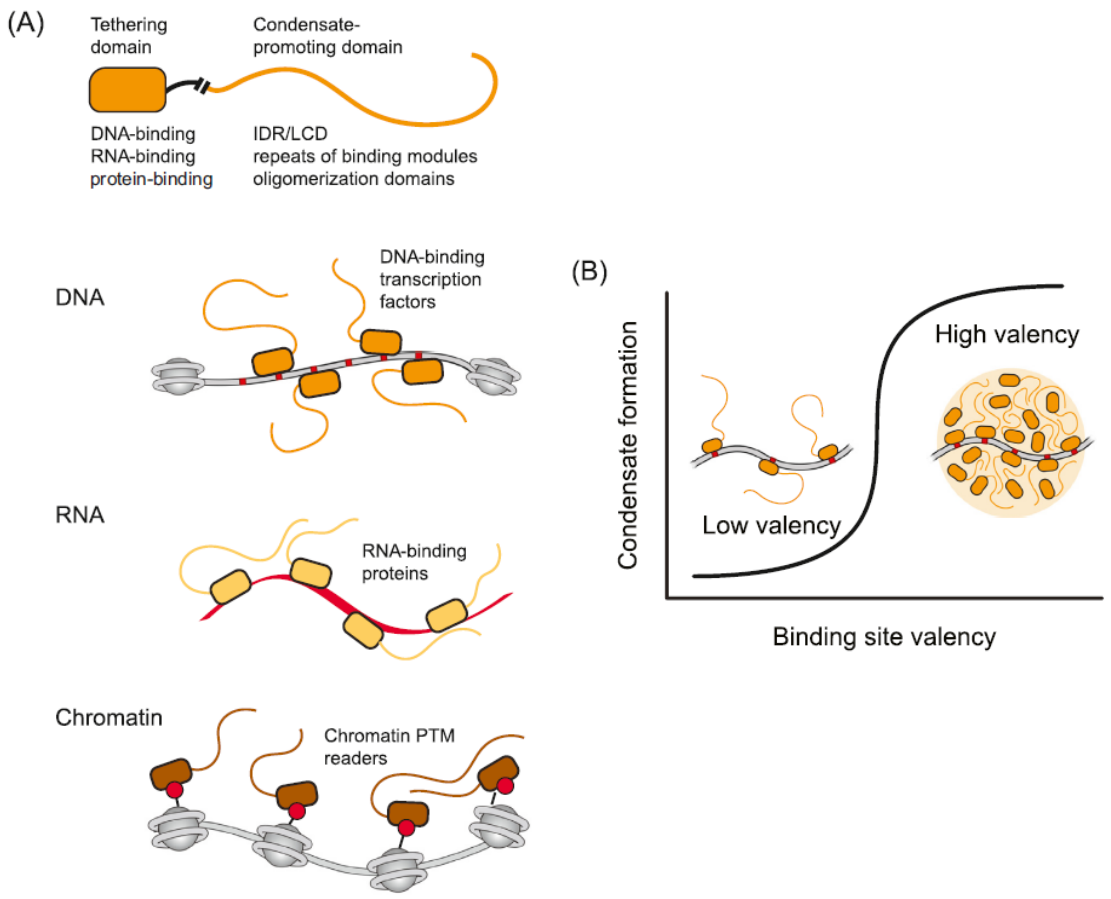
(Gomes E, Shorter J, 2019)

Формирование немембранных органелл

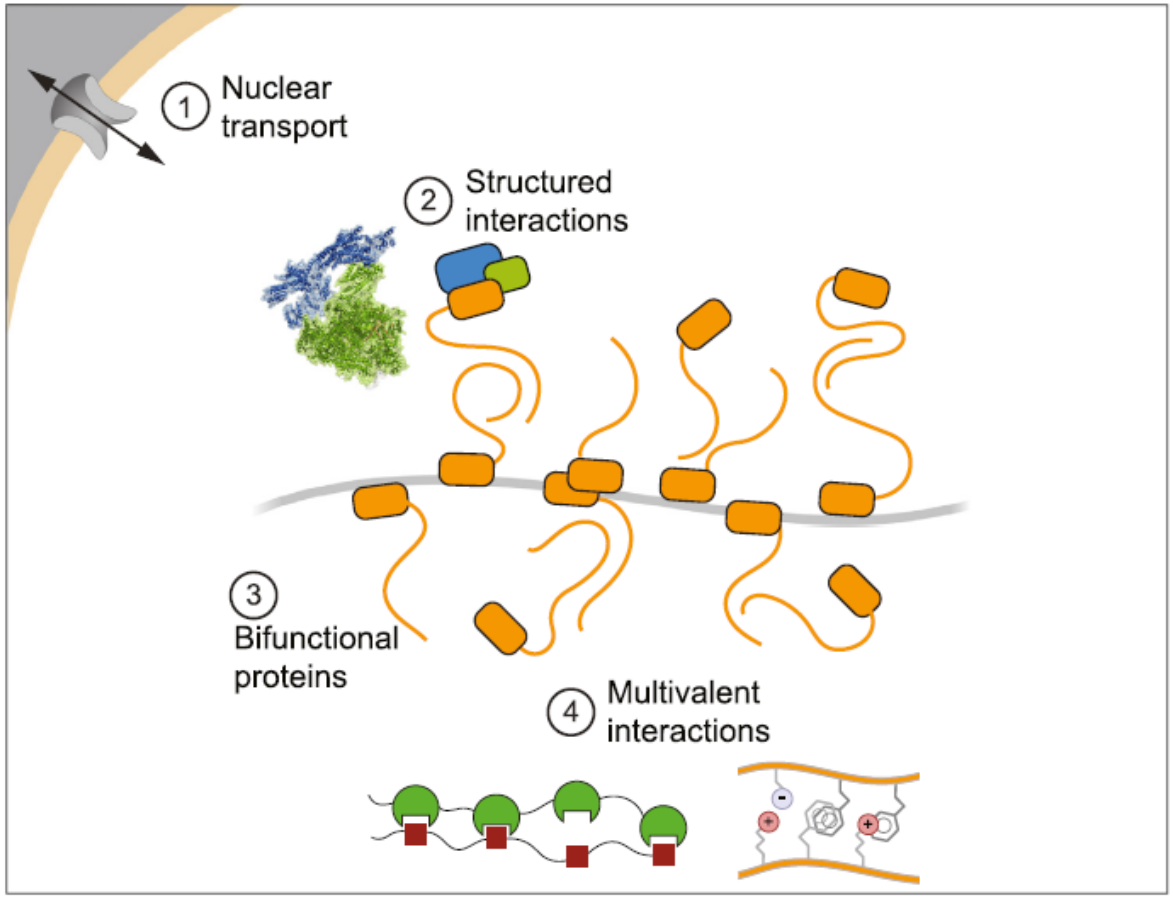


Veronica H. Ryan et al., 2019

Bifunctional Proteins
Promote Condensate Formation
at Specific Genomic Loci.



Sources of Compositional Specificity in Nuclear Condensates.



Benjamin R. Sabari et al., 2020

Метилирование гистонов и узнающие домены

Histone code hypothesis: multiple histone modifications, acting in a combinatorial or sequential fashion on one or multiple histone tails, specify unique downstream function (Strahl and Allis)

- ПТМ – влияние на взаимодействие гистон-гистон, гистон-ДНК или привлечение эффекторов
- Домен узнает разные метки
- Метка узнается разными доменами

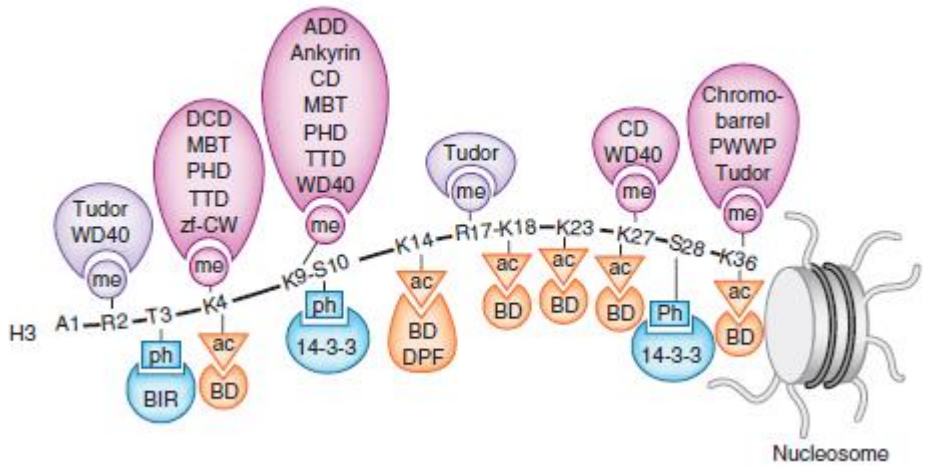


Table 1 Histone readers and their target PTMs

Recognition of	Reader	Histone PTM
Methyllysine	ADD	H3K9me3
	Ankyrin	H3K9me2, H3K9me1
	BAH	H4K20me2
	Chromo-barrel	H3K36me3, H3K36me2, H4K20me1, H3K4me1
	Chromodomain	H3K9me3, H3K9me2, H3K27me3, H3K27me2
	DCD	H3K4me3, H3K4me2, H3K4me1
	MBT	H3Kme1, H3Kme2, H4Kme1, H4Kme2
	PHD	H3K4me3, H3K4me2, H3K9me3
	PWWP	H3K36me3, H4K20me1, H4K20me3, H3K79me3
	TTD	H3K4me3, H3K9me3, H4K20me2
	Tudor	H3K36me3
	WD40	H3K27me3, H3K9me3
	zf-CW	H3K4me3
Methylarginine	ADD	H4R3me2s
	Tudor	H3Rme2, H4Rme2
	WD40	H3R2me2
Acetyllysine	Bromodomain	H3Kac, H4Kac, H2AKac, H2BKac
	DBD	H3KacKac, H4KacKac
	DPF	H3Kac
	Double PH	H3K56ac
Phosphoserine or phosphothreonine	14-3-3	H3S10ph, H3S28ph
	BIR	H3T3ph
	Tandem BRCT	H2AXS139ph
Unmodified histone	ADD	H3un
	PHD	H3un
	WD40	H3un

(Musselman et al. 2012)

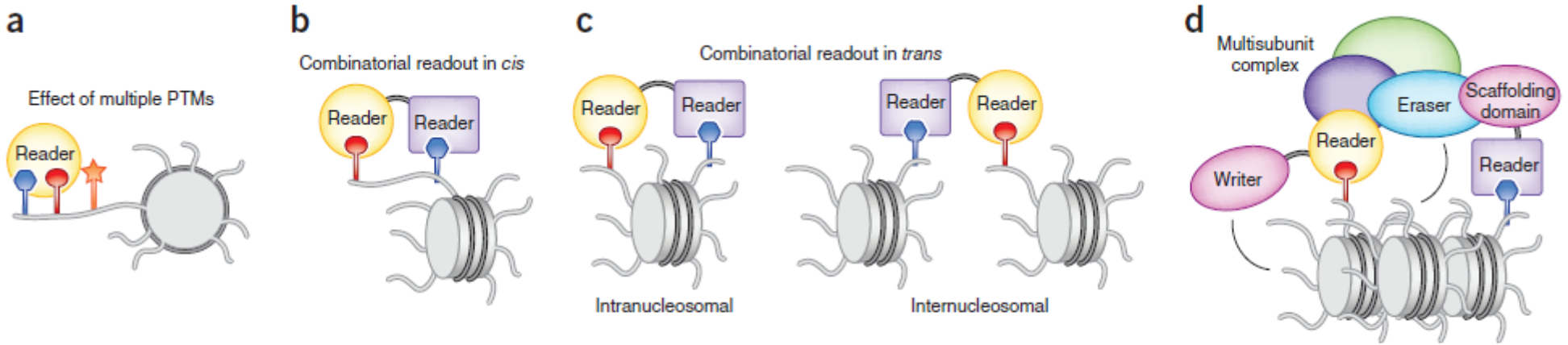
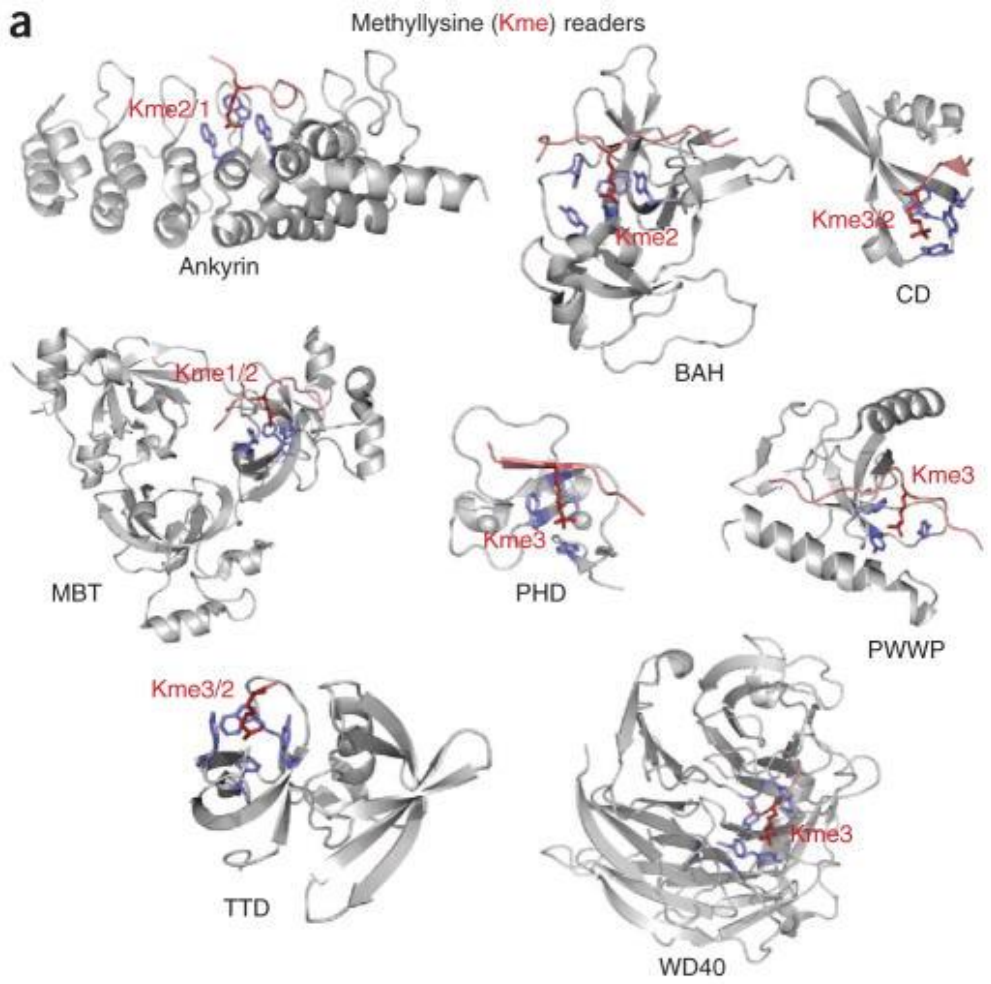
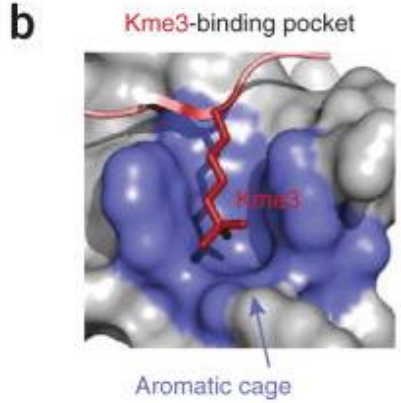
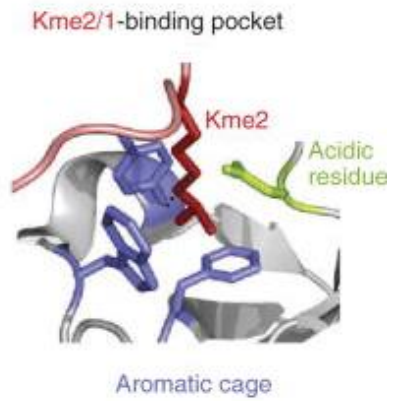


Figure 4 Combinatorial readout of PTMs. (**a**–**c**) Recognition of a target PTM is influenced by adjacent PTMs on the same histone tail (**a**) and the combined action of multiple readers within the same protein (**b**,**c**). (**d**) Multivalent engagement of readers within individual subunits of the complex. The reader-harboring proteins can also contain the catalytic domains (which act as writers and erasers) or scaffolding domains that bridge their host proteins with other subunits of the complex. Readers can recognize PTMs on a single histone tail (*cis* mechanism) or different histone tails (*trans* mechanism).

(Musselman et al. 2012)







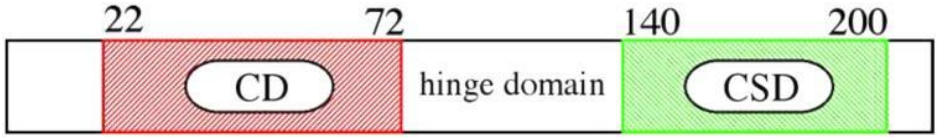
Upon binding, the histone peptide adopts a β-strand conformation and inserts between two β-strands of the chromodomain, completing the five-stranded antiparallel β-barrel. This induced-fit binding mode is stabilized through backbone hydrogen bonds and electrostatic contacts involving up to seven residues preceding trimethylated lysine and one residue following trimethylated lysine of the histone peptide.

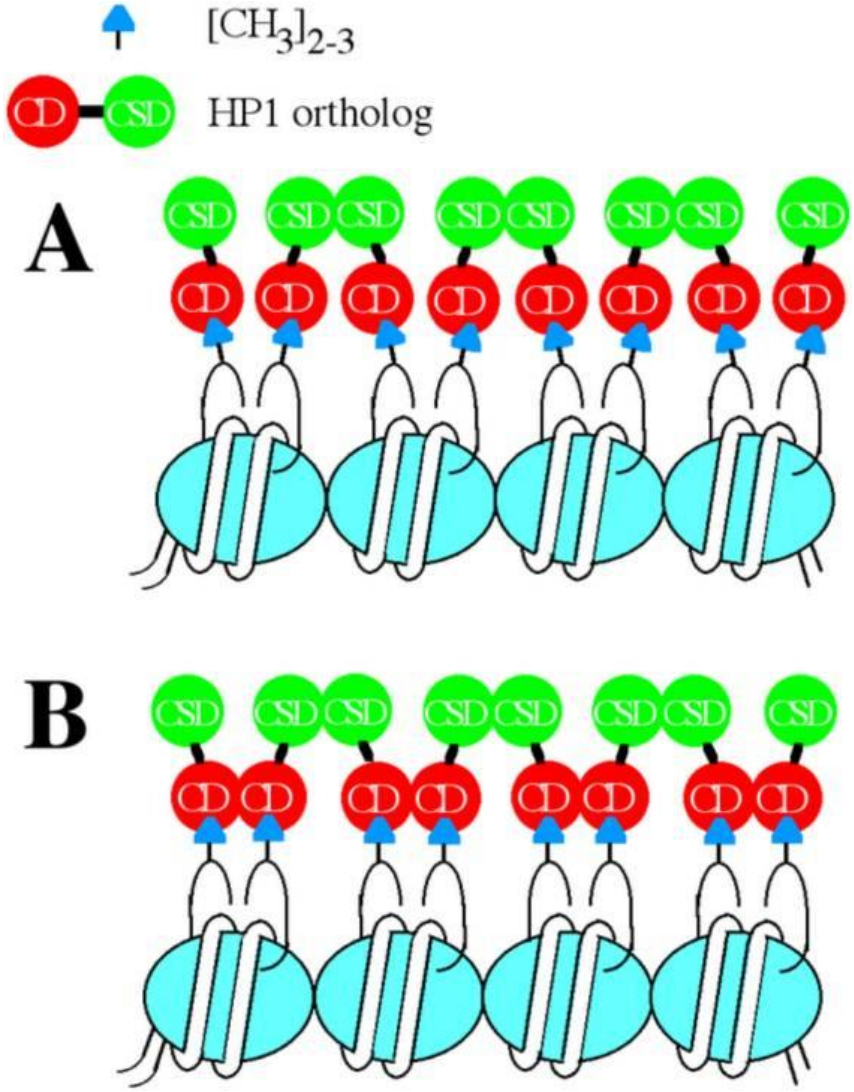
- Molecular mechanisms for the recognition of methyllysine and methylarginine. (a) Structures of the readers in complex with histone peptides methylated at lysine residues. The aromatic cage residues and the histone peptides are colored blue and red, respectively. (b–d) The binding sites for trimethylated lysine (Kme3) (b), dimethylated lysine (Kme2) (c). PDB codes: 3B95, 4DOW, 1KNE, 2PQW, 2G6Q, 2X4X, 2IG0, 3IIW, 2F6J, 4A7J and 4A4E.

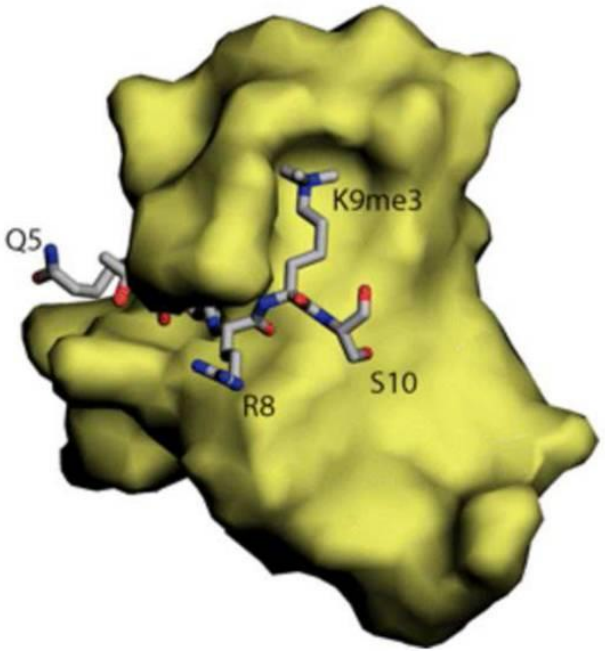
(Musselman et al. 2012)

HP1a

- Обнаружили в политенных хромосомах Drosophila
- Узнает H3K9me2/3
- Есть chromo domain (и chromo shadow), может димеризоваться – мультивалентное узнавание

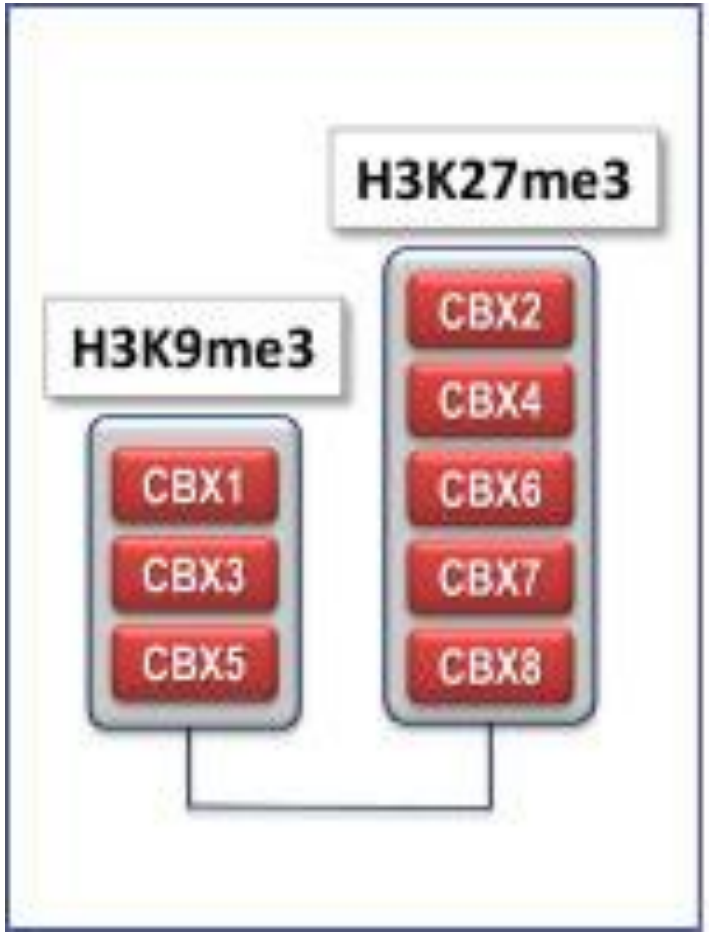






(Eissenberg and Elgin 2014)

- Polycomb-like CBX members (CBX2, CBX4, CBX6, CBX7, and CBX8)
- HP1-like CBX members CBX1, CBX3, and CBX5



BOX 1 Methodologies for the identification of histone readers

Novel techniques are being developed for rapid high-throughput and discovery-based screening of novel PTM readers and characterization of the effect of adjacent PTMs.

Histone microarrays[146–149]. Histone peptides carrying various PTMs are synthesized on cellulose support (SPOT and CelluSpot) or attached to streptavidin-coated platforms by a biotin tag. GST- or His-tagged protein is incubated with the screen, then the bound fraction is detected by western blot or fluorescence. Binding to a single PTM or combinations of PTMs can be assayed in a high-throughput manner; however, this technique requires purification of the protein of interest, and initial generation of the screen can be time consuming.

Bead-based approach[150]. This approach utilizes a PTM-randomized combinatorial library of histone peptides. Proteins incubated with the library are captured on beads and their target peptides are analyzed by MS. This method is very powerful in detecting synergistic and inhibitory combinations of known and yet-to-be-discovered PTMs, but it has the same limitations as histone microarrays.

Stable isotope labeling by amino acids in cell culture (SILAC) technology[151,152]. Nuclear extracts from cells grown in SILAC medium are incubated with PTM-bearing immobilized histone peptides or recombinant nucleosomes. Bound proteins or protein complexes are pulled down and analyzed by quantitative MS. This approach provides an unbiased way for the identification of new PTM readers; however, the particular protein and domain responsible for the targeting may not be easily identified.

(Musselman et al. 2012)

› Mol Cell Proteomics. 2013 Oct;12(10):2750-60. doi: 10.1074/mcp.O112.025015. Epub 2013 Jul 10.

Identification of methyllysine peptides binding to chromobox protein homolog 6 chromodomain in the human proteome

Nan Li [1], Richard S L Stein, Wei He, Elizabeth Komives, Wei Wang

Affiliations – collapse

Affiliation

1 Department of Chemistry and Biochemistry, 9500 Gilman Drive, University of California, San Diego, La Jolla, California 92093-0359;

PMID: 23842000 PMCID: PMC3790288 DOI: 10.1074/mcp.O112.025015

Free PMC article

- we first characterized the domain–peptide interface using a mutation matrix generated from molecular dynamics and virtual mutagenesis, which revealed residues critical for binding and suggested possible ways to engineer the recognition interface.

(Li et al. 2013)

Что делали в статье

- Взяли комплекс CBX6-H3K27me3, молдинамика,
- Виртуальный мутагенез для оценки наиболее важных остатков в CBX6 (взяли 21), также мутировали остатки H3 (n-9, n+1)
- MM-GBSA для оценки свободной энергии связывания комплекса (для каждого мутантного) (траектории с 1 до 2 нс, 100 снэпшотов); декомпозиция энергетического вклада (ВДВ, электростатика, сольватация)

$$\Delta G_{bind} = \langle G_{complex} \rangle - (\langle G_{protein} \rangle + \langle G_{peptide} \rangle)$$

$$G_{(complex/protein/peptide)} = E_{elec} + E_{vdw} + G_{solv}$$

$$G_{solv} = G_{polar} + G_{nonpolar}$$

- Microarray (иммобилизованные пептиды + Cy3 marker, anti-GST mouse monoclonal IgG antibody, secondary anti-mouse IgG Dylight-488 conjugated antibody)
- Протеомный скрининг (для нахождения пептидов, которые CBX6 тоже будет узнавать)

RESULTS

1. a survey of the residue contributions in Table I shows that van der Waals interactions are the main force in the binding between CBX6 and H3, accounting for 73% of the energy that is responsible for the complex binding

TABLE I
Binding free energies of selected residues in the CBX6-H3K27me3 complex

Residue	Contribution (kcal/mol)	Residue	Contribution (kcal/mol)
Trp32	−4.77 ± 0.52	Glu46	−0.99 ± 0.28
Phe11	−3.78 ± 0.81	Arg9	−0.79 ± 0.94
Leu49	−3.62 ± 0.47	Glu58	−0.61 ± 0.04
Asn47	−3.40 ± 0.73	Val30	−0.51 ± 0.09
Asp50	−2.88 ± 0.52	Glu45	−0.45 ± 0.02
Val10	−2.77 ± 0.39	H3Q19	+0.08 ± 0.67
Arg52	−2.41 ± 1.14	H3L20	−0.72 ± 0.65
Leu53	−2.15 ± 0.30	H3A21	−0.45 ± 0.25
Thr41	−2.04 ± 0.32	H3T22	−1.53 ± 0.43
Ala13	−1.98 ± 0.43	H3K23	−6.20 ± 0.77
Trp35	−1.98 ± 0.60	H3A24	−4.64 ± 0.51
Glu43	−1.88 ± 0.74	H3A25	−6.05 ± 0.55
Ala12	−1.66 ± 0.24	H3R26	−7.22 ± 0.77
Glu14	−1.29 ± 0.37	H3K27me3	−11.79 ± 0.93
Ile48	−1.22 ± 0.22	H3S28	−1.73 ± 0.71
Tyr39	−1.04 ± 0.30	H3A29	−1.14 ± 0.53

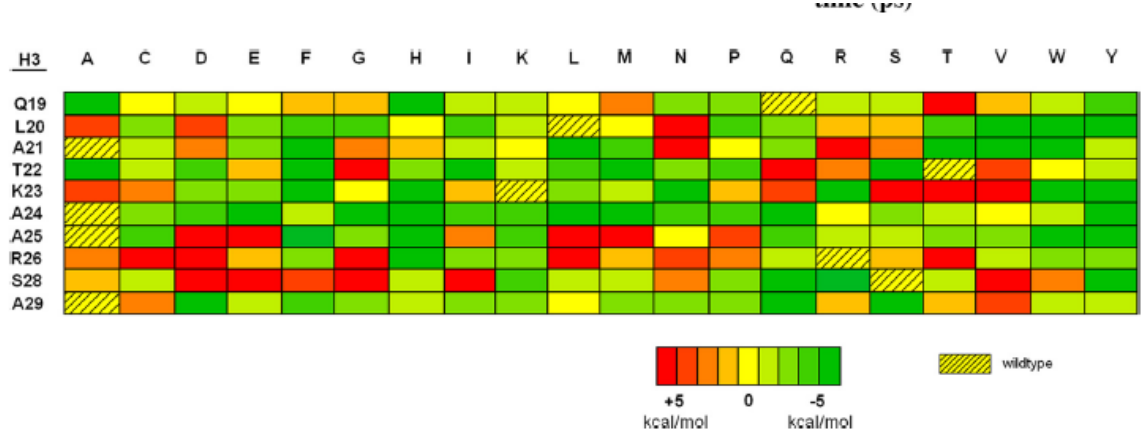


FIG. 2. **Mutation matrix for the 10 residues in H3 apart from K27 that occur in the crystal structure used for the simulation.** Red cells indicate mutations relatively unfavorable for binding to CBX6 protein, and green cells show mutations that are favorable relative to wild-type H3.

(Li et al. 2013)

RESULTS

- A total of 50 non-histone peptides were identified as putative binders for the CBX6 chromodomain with significant fluorescence intensities above a statistical cutoff of $p < 0.05$ (from the background Gaussian distribution).

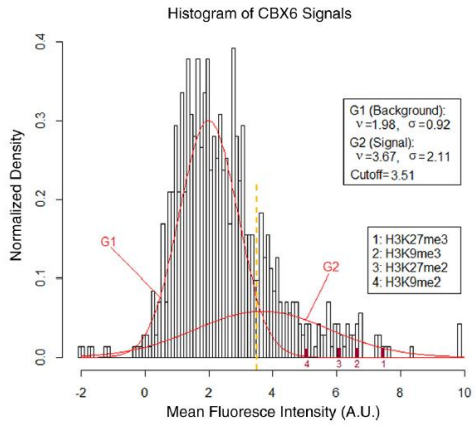


FIG. 4. **Gaussian mixture fitting of signal intensities from peptide microarray for the CBX6-GST fusion protein.**

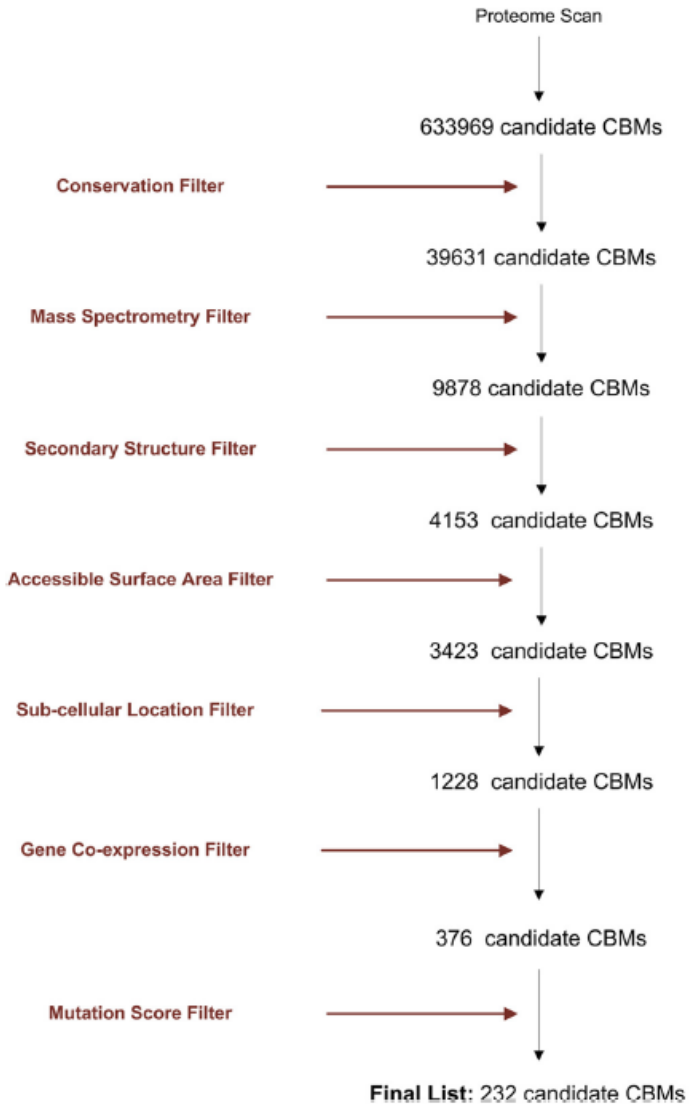


FIG. 3. **Flowchart of the proteomic scan of methyllysine peptides binding to the CBX6 chromodomain.**

(Li et al. 2013)

Science Advances Contents News Careers Journals

SHARE

RESEARCH ARTICLE | BIOCHEMISTRY

Deciphering and engineering chromodomain-methyllysine peptide recognition

Ryan Hard[1,*], Nan Li[1,*], Wei He[1,*,†], Brian Ross[2,3], Gary C. H. Mo[2], Qin Peng[4], Richard S. L. Stein[1,‡], Elizabeth Komives[1], Yingxiao Wang[4], Jin Zhang[2,5] and Wei Wang[1,6,§]

[1]Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, CA 92093, USA.
[2]Department of Pharmacology, University of California, San Diego, La Jolla, CA, 92093, USA.
[3]Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD 21218, USA.
[4]Department of Bioengineering, University of California, San Diego, La Jolla, CA 92093, USA.
[5]Department of Pharmacology and Molecular Sciences, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA.
[6]Department of Cellular and Molecular Medicine, University of California, San Diego, La Jolla, CA 92093, USA.

↵§Corresponding author. Email: wei-wang@ucsd.edu

↵* These authors contributed equally to this work.

↵† Present address: Polypeptide Group, 365 Maple Ave., Torrance, CA 90503, USA.

↵‡ Present address: Microsoft Corporation, 14855 NE 36th Street, Redmond, WA 98052, USA.

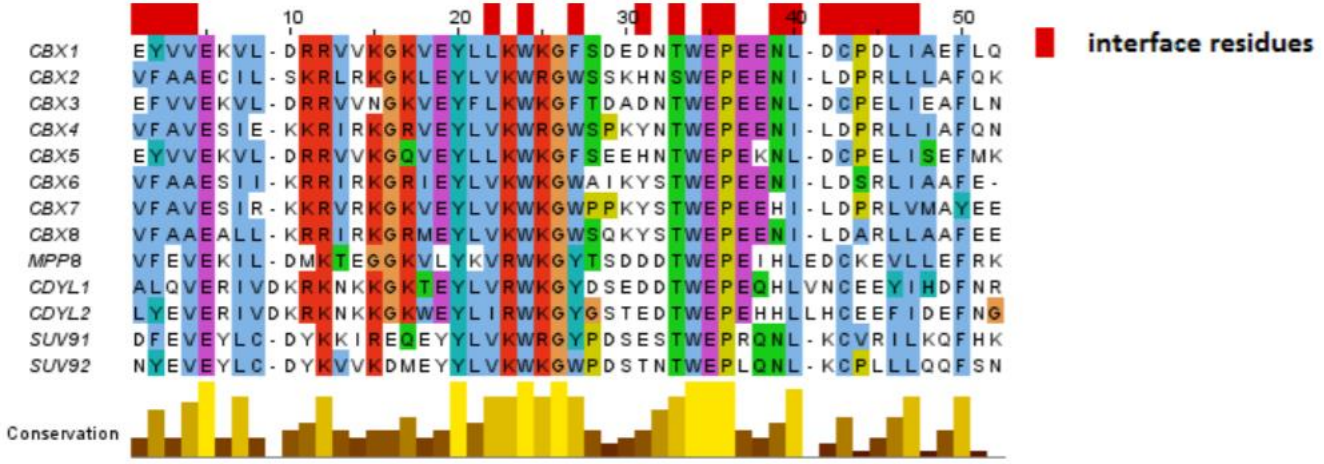
– Hide authors and affiliations

Science Advances 07 Nov 2018:
Vol. 4, no. 11, eaau1447
DOI: 10.1126/sciadv.aau1447

(Hard et al. 2018)

29 chromodomains and 467 peptides

- 9 chromodomains encoded in the human proteome
- 153 nine–amino acid–long histone peptides with a single trimethylated lysine
- 72 H3 histone peptides with multiple modifications, which includes up to three possible combinations of modifications adjacent to K4, K9, K27, K36, K56, and K79 of the H3 histone protein
- 232 non-histone peptides
- 10 proline-rich peptides without any modifications

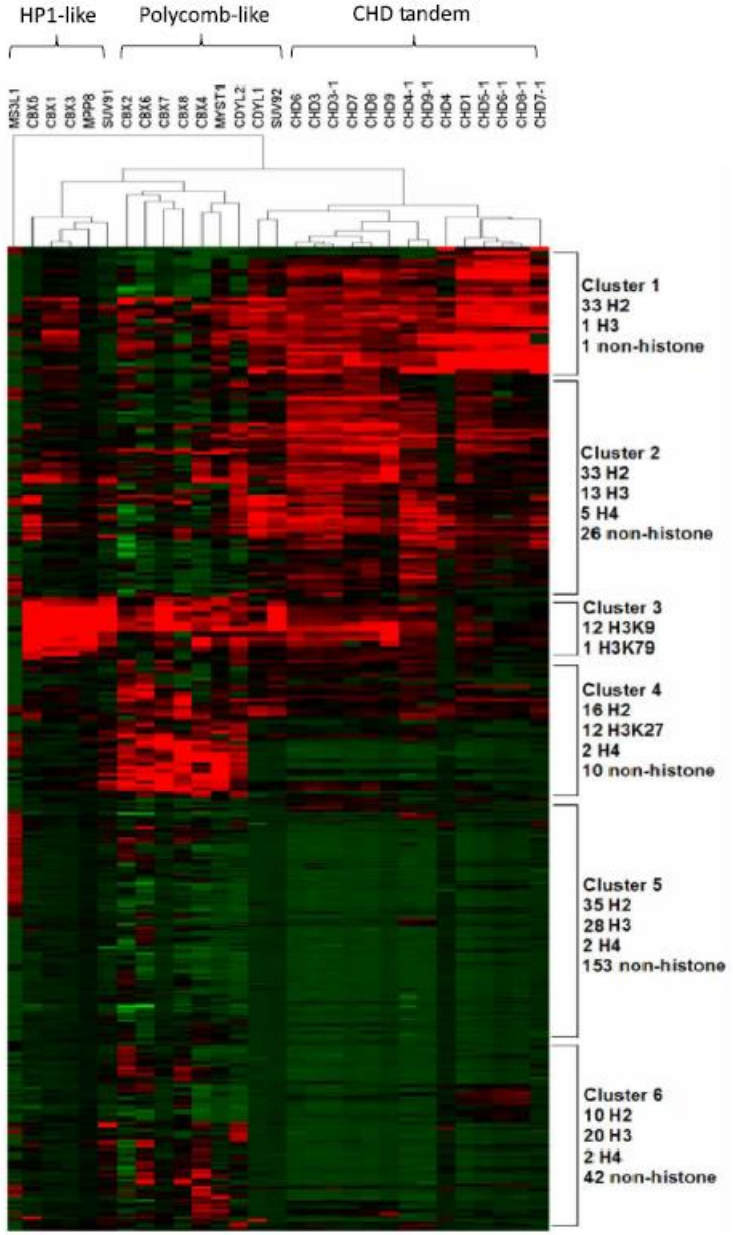


(Hard et al. 2018)

Results

- (i) a single binding mode where one single chromodomain forms a binding pocket to accommodate the histone peptide (such as CBX proteins), HP1-like proteins and Polycomb-like proteins (from *Drosophila melanogaster* HP1 and Pc chromodomains)
- (ii) a potential tandem binding mode in which two chromodomains are linked through a loop and both domains may interact with the peptide [such as chromodomain-helicase DNA binding (CHD) proteins]

Domain	H3K9me3 signal	H3K27me3 signal	K_d H3K9me3	K_d H3K27me3
CBX1	4.20	0.07	5 ± 2[9]	N/B[9]
CBX2	1.57	2.77	>500[9]	185 ± 20[9]
CBX3	7.23	-0.03	15 ± 8[9]	N/B[9]
CBX4	2.39	3.62	70 ± 7[9]	205 ± 20[9]
CBX5	4.45	0.03	30 ± 5[9]	N/B[9]
CBX6	2.55	3.11	>500[8]	330 ± 120[8]
CBX7	4.98	3.73	55 ± 5[9]	110 ± 17[9]
CBX8	1.62	1.97	>500[8]	165 ± 20[8]

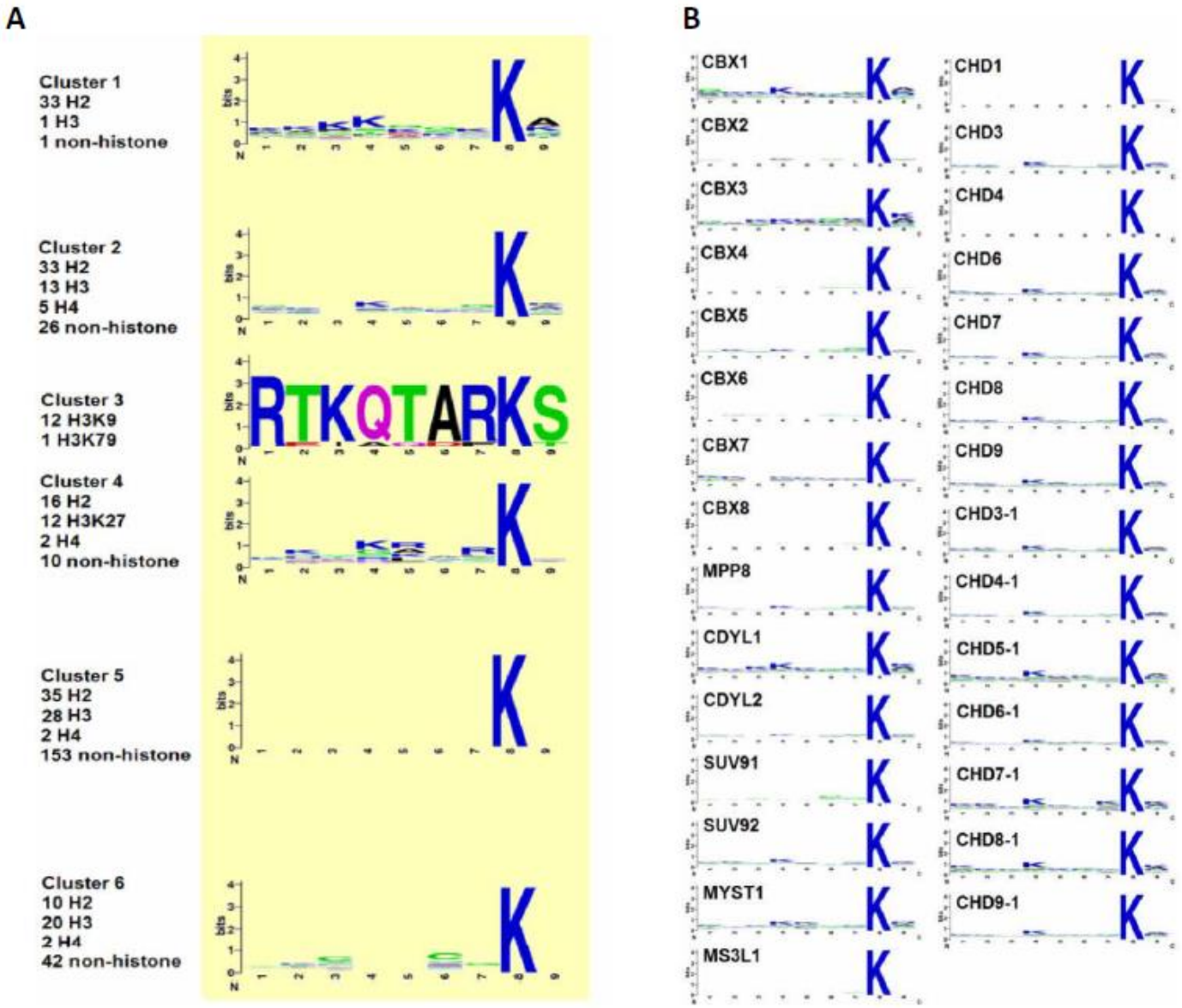


В кластерах и гистоновые петиды и негистоновые – конкуренция

(Hard et al. 2018)

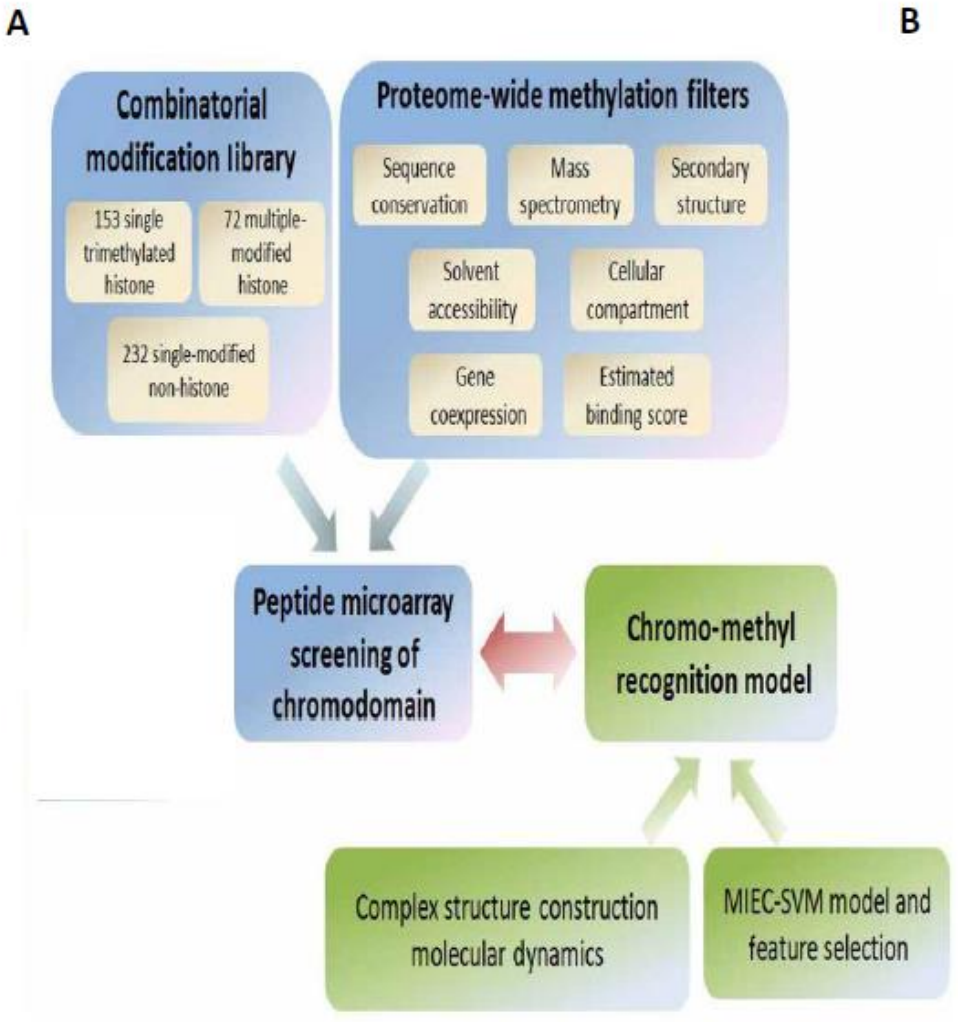
Results

Together, these observations suggest an absence of PTM patterns that govern chromodomains' binding specificity.



(Hard et al. 2018)

we constructed a MIECSVM model to capture the structural and energetic patterns critical for the chromodomain-peptide binding



(Hard et al. 2018)

MIEC-SVM

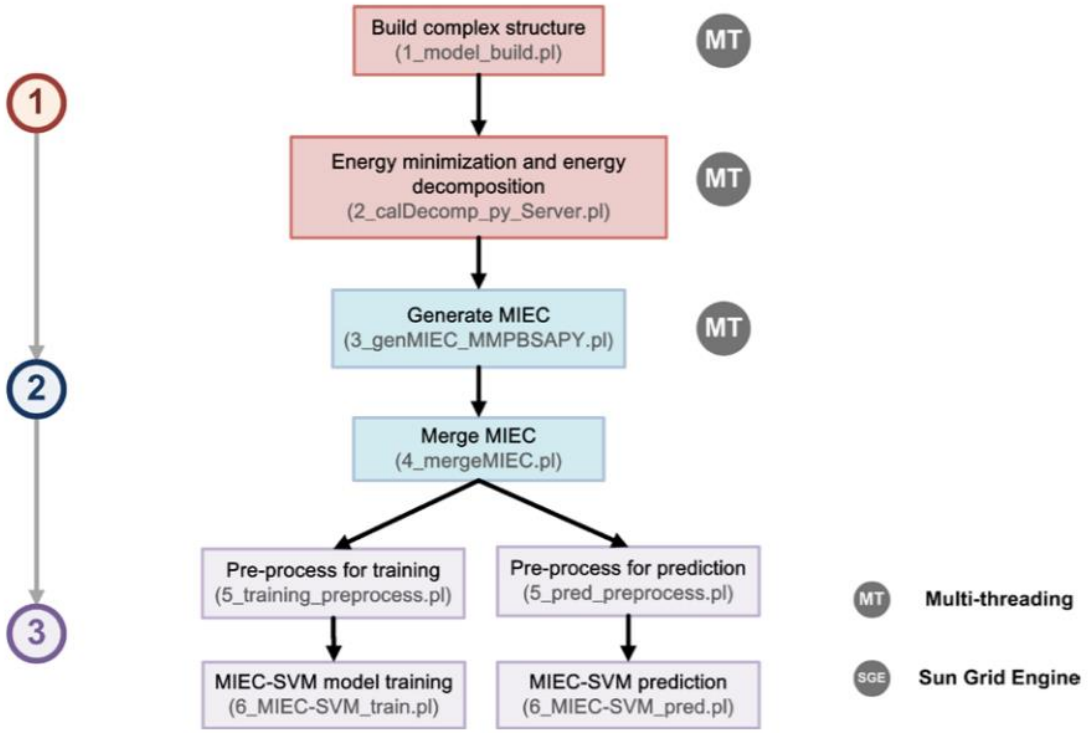


Figure 1. Flow of data for MIEC-SVM pipeline application on servers/workstations.

Part VII. MIEC-SVM overview

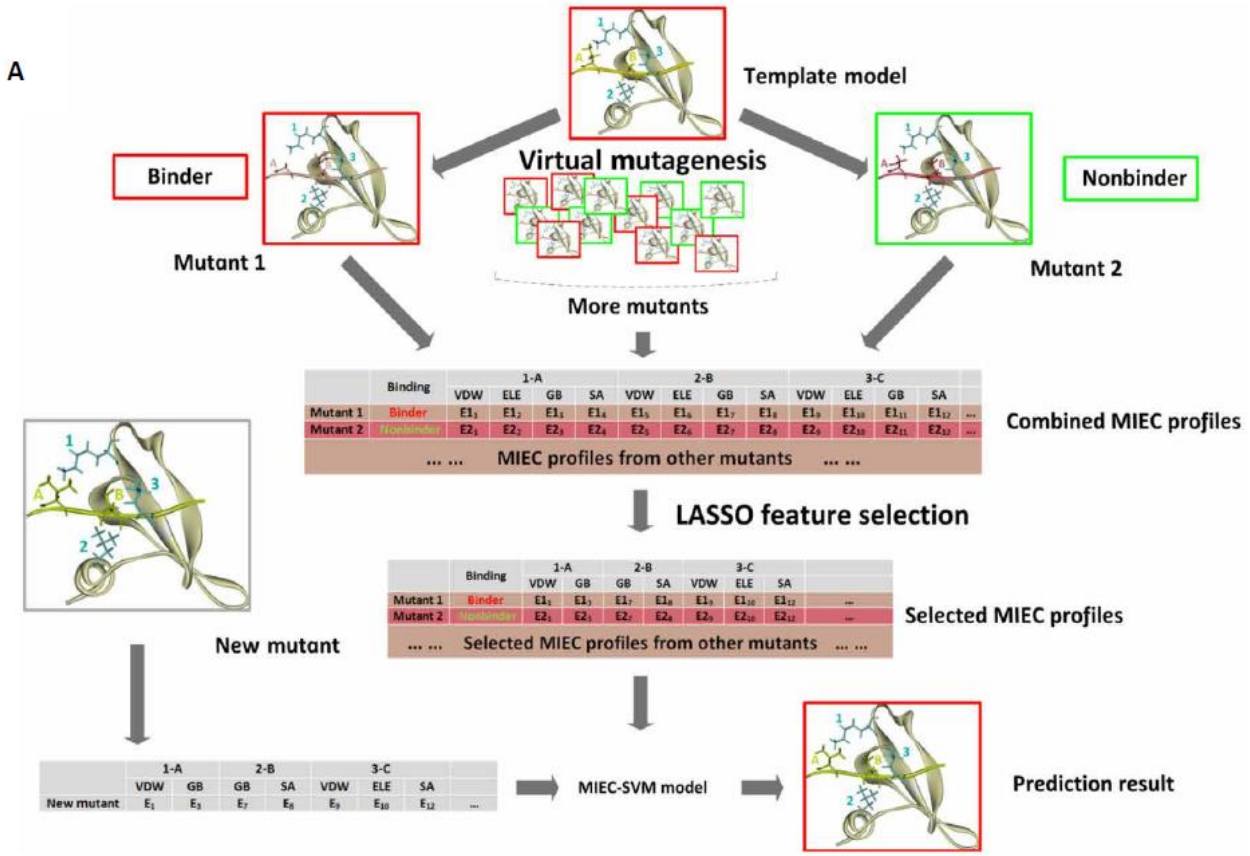
MIEC-SVM is a computational method to predict protein-peptide and protein-ligand interactions. The MIEC (Molecular Interaction Energy Component) profile is used to characterize the interactions through the energetic pattern of residue pairs between the protein and the peptide/ligand. SVM (support vector machine) is a machine learning method that can make a binary classification on the protein-peptide/protein-ligand interactions, based on the MIEC profiles.

The MIEC-SVM pipeline provides an integrated and user-friendly workflow for the construction and application of the MIEC-SVM model. It consists of three sections: model building, MIEC construction, and model training/prediction.

MIEC-SVM manual

- AMBER ff03 force field provides force field parameters and partial charges for all standard residues and the AMBER gaff force field defines atom types and force field parameters for small molecules (ligands). The pipeline provides a lightweight database MolDB (the force field information for non-standard residues and ligands of interest)
- four energy components of MM/GBSA: van de Waals (VDW), electrostatics (ELE), generalized Born (GB), and surface area (SA).

(Hard et al. 2018)



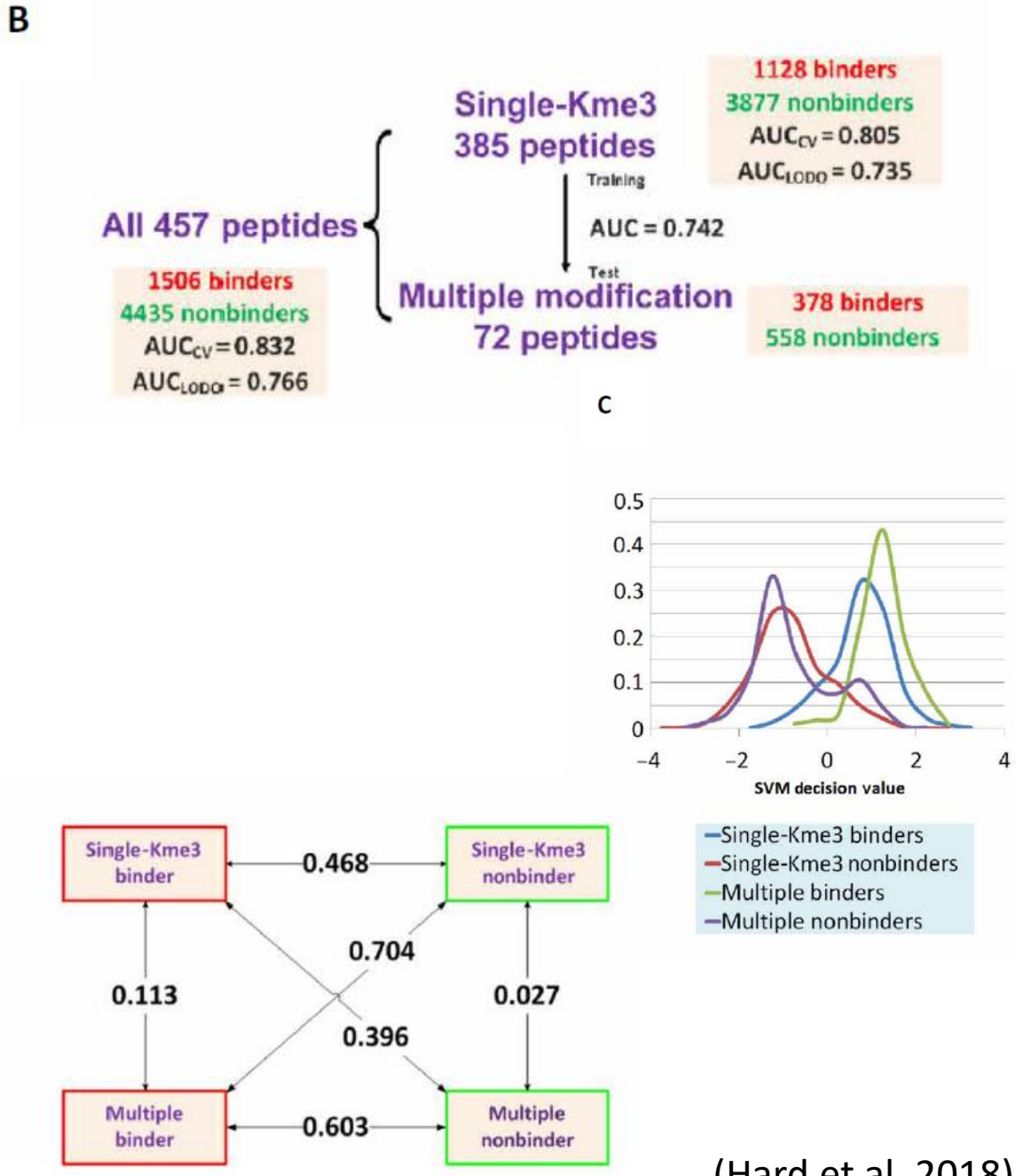
Flowchart of MIEC-SVM that predicts binding specificity between chromodomains and methyllysine peptides. Complex structures between 13 chromodomains and 457 peptides were constructed by computationally mutating peptide sequence from a template complex for each chromodomain (virtual mutagenesis). From the modeled complex structures, MIEC terms between peptide-protein residues at the binding interface were computed. The MIECs and the binding/nonbinding label (obtained from microarray experiments) for each domain-peptide pair were input to a LASSO logistic regression model to select most predictive MIECs (LASSO feature selection). These selected MIEC features were then used to train an SVM model to discriminate binding from nonbinding events. VDW, Van der Waals forces; ELE, electrostatic forces; GB, polar contribution to the desolvation energy; SA, nonpolar contribution to the desolvation energy.

(Hard et al. 2018)

(B)Performance of MIEC-SVM model on three different peptide groups (all peptides, singly modified peptides, and multiply modified peptides). The MIEC-SVM model showed consistent performance regardless of the number of modifications on the peptides, indicating that chromodomain-peptide recognition share the same MIEC features for singly and multiply modified peptides.

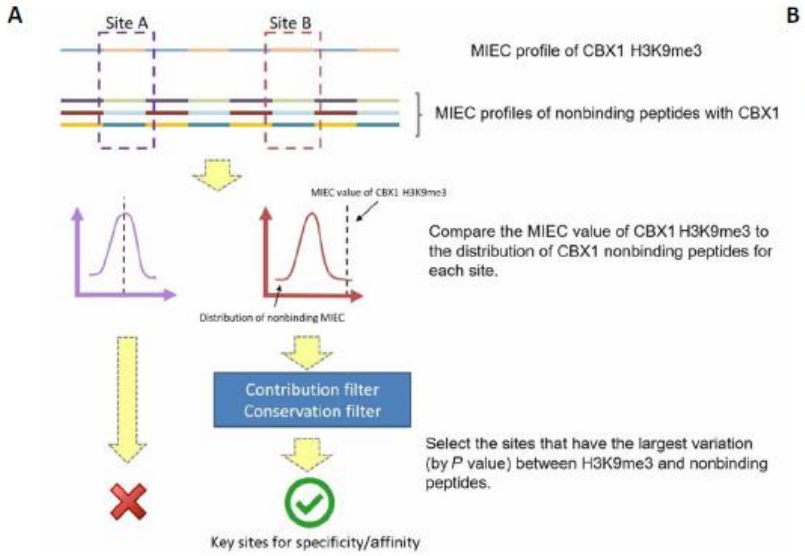
(C) SVM decision value distribution of the four classes of peptides (binders/nonbinders with single or multiple modifications). Binders and nonbinders are well separated regardless of the modification number.

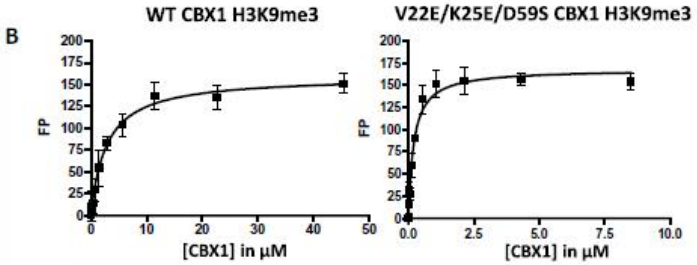
(D) Pairwise Jensen-Shannon (JS) divergences between the SVM decision value distributions of the four classes. The differences between any binder class and nonbinder class (regardless of the PTM number) are large (larger JS divergence value) singly modified binder–singly modified nonbinder, JS = 0.468 ($P < 1.0 \times 10^{-20}$); singly modified binder–multiply modified nonbinder, JS = 0.396 ($P < 1.0 \times 10^{-19}$); multiply modified binder–single modified nonbinder, JS = 0.704 ($P < 1.0 \times 10^{-20}$); and multiply modified binder–multiply modified nonbinder, JS = 0.603 ($P < 1.0 \times 10^{-20}$). In contrast, binder (or nonbinder) peptides are similar to each other regardless of the PTM numbers: JS values of 0.113 for binders ($P = 7.0 \times 10^{-15}$ for statistical similarity) and 0.027 for nonbinders ($P = 6.1 \times 10^{-10}$). All P values were calculated on the basis of the background distributions of JS divergence of randomly selected decision values for the same number of binders or nonbinders as the foreground.



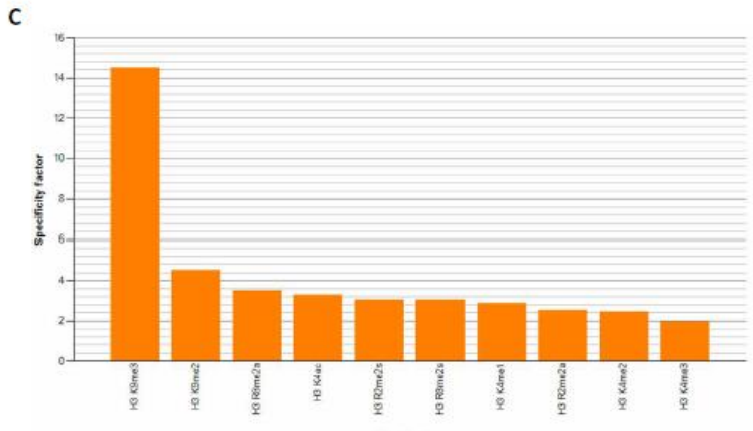
(Hard et al. 2018)

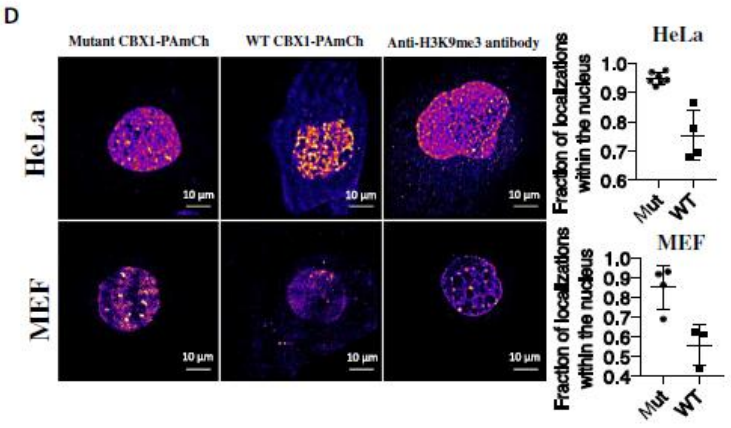
Application of the chromodomain MIEC-SVM model to engineering the CBX1 chromodomain





Mutant	K_d to H3K9me3	K_d to H3K27me3
WT CBX1	2.78 ± 0.60 μM	>25 μM
V22E/K25E/D59S (CBX1)	0.21 ± 0.04 μM	>50 μM
V22E/K25S/D59F (CBX1)	0.32 ± 0.07 μM	>25 μM
WT MPP8	0.50 ± 0.12 μM	>50 μM

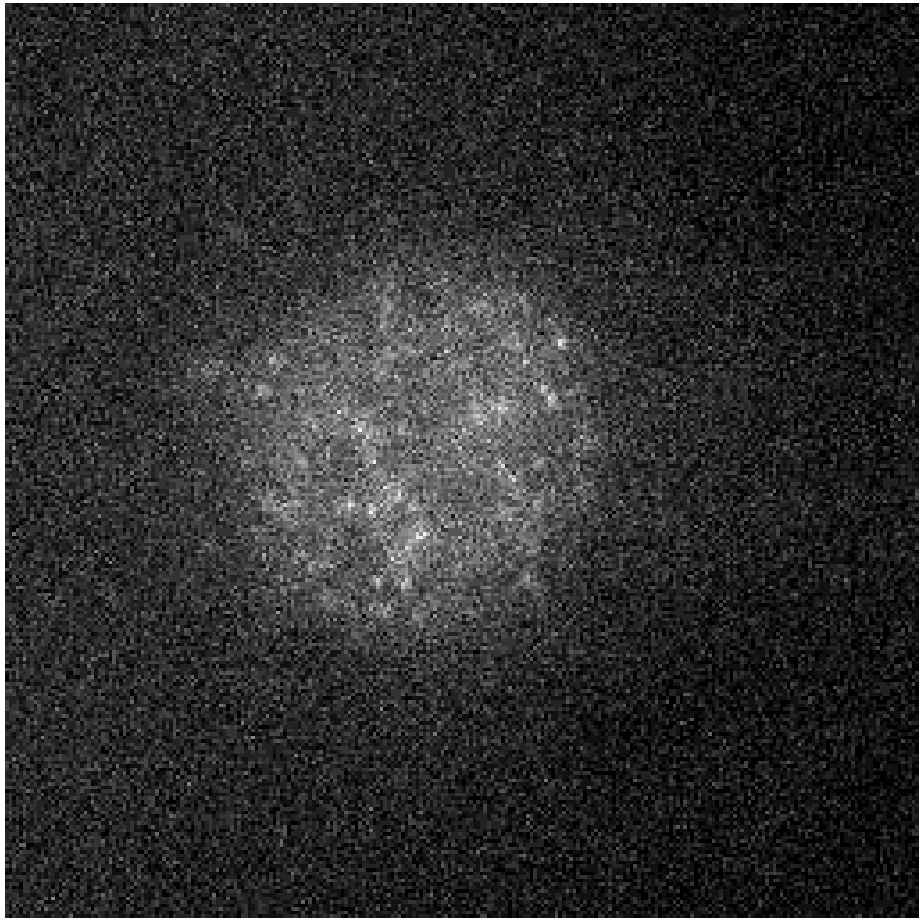




The two mutants selected from the H3K9me3 sorts (V22E/K25S/D59F and V22E/K25E/D59S) showed superior binding to H3K9me3 than the WT domain (K_d values of 0.32 and 0.21 μM versus 2.78 μM of the WT domain), yet showed very weak binding to the similar H3K27me3 peptide. To our knowledge, the V22E/K25E/D59S mutant has the strongest binding affinity of any human chromodomain toward H3K9me3, even over that of MPP8.

(Hard et al. 2018)

V22E/K25E/D59S CBX1-PAmCherry in MEF cells



PALM imaging:

the mutant showed a very high level of nuclear localization (approximately 85% in MEF and 95% in HeLa cells), which is impressive given the lack of a nuclear localization sequence (NLS) tag in the construct.

Measured that an average of 63% of localizations of the CBX1 mutant coclustered with the H3K9me3 antibody, and 94% of H3K9me3 antibody localizations coclustered with mutant CBX1 (V22E/K25E/D59S). These data are consistent with the peptide array data and confirm the CBX1 mutant's specificity for the H3K9me3 mark within the cellular environment.

(Hard et al. 2018)

CONCLUSION

- We found great degeneracy of chromodomain proteins binding to modified histone peptides in that one chromodomain can bind to diverse modifications at different histone sites, while one histone modification pattern can be bound by different reader proteins.
- We found this recognition degeneracy is not associated with amino acid sequence or PTM patterns, but rather is rooted in the same physiochemical properties of the binding interface that are defined by the PTMs.

(Hard et al. 2018)