

Лекция 11. Структура нуклеиновых кислот и хроматин

Курс: Структурная Биоинформатика и моделирование
лекарств (ВШЭ)

Головин А.В.¹

¹МГУ им М.В. Ломоносова, Факультет Биоинженерии и Биоинформатики

Москва, 2017

Содержание:

Введение

Химическое строение НК

Вторичная структура РНК

Механические модели ДНК

Мезомоделирование ДНК

Хроматин

Нуклеиновые кислоты.

Нуклеиновые кислоты (НК) – это высокомолекулярные линейные полярные биополимеры, **полинуклеотиды**, которые построены из **нуклеотидных остатков**.

2 типа нуклеиновых кислот:

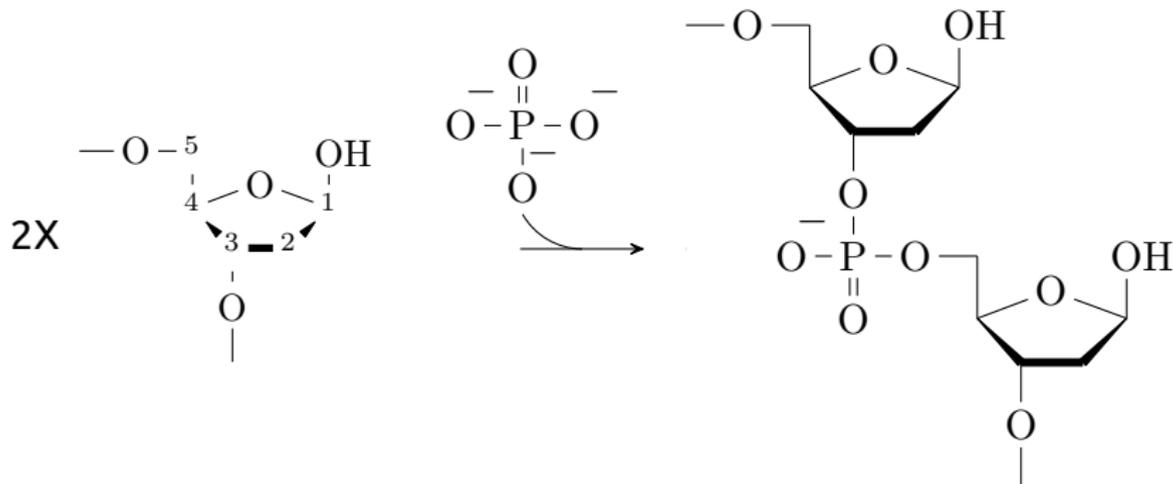
- дезоксирибонуклеиновая кислота – ДНК
- рибонуклеиновая кислота – РНК

ДНК

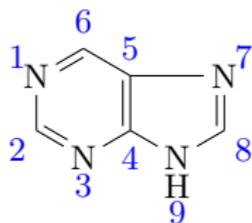
Расшифровка аббревиатуры ДНК:

- ДНК – открытие и выделение **“нуклеина”** из ядер (нуклеус) лейкоцитов Ф. Мишером 1869 г.
- ДНК – линейный сополимер на основе ортофосфорной **кислоты**.
- ДНК – линейный сополимер ортофосфорной кислоты и **дезоксирибозы**

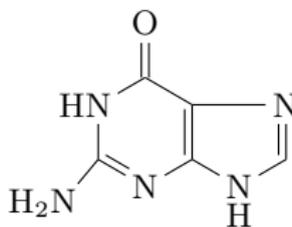
Дезоксирибоза и ортофосфорная кислота



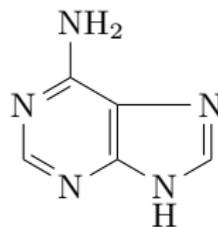
Основания



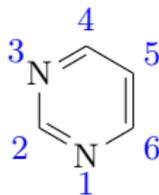
Пурин



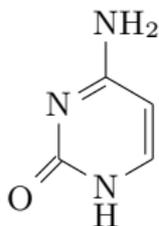
Гуанин



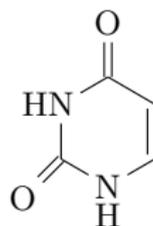
Аденин



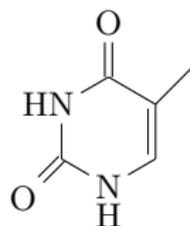
Пиримидин



Цитозин

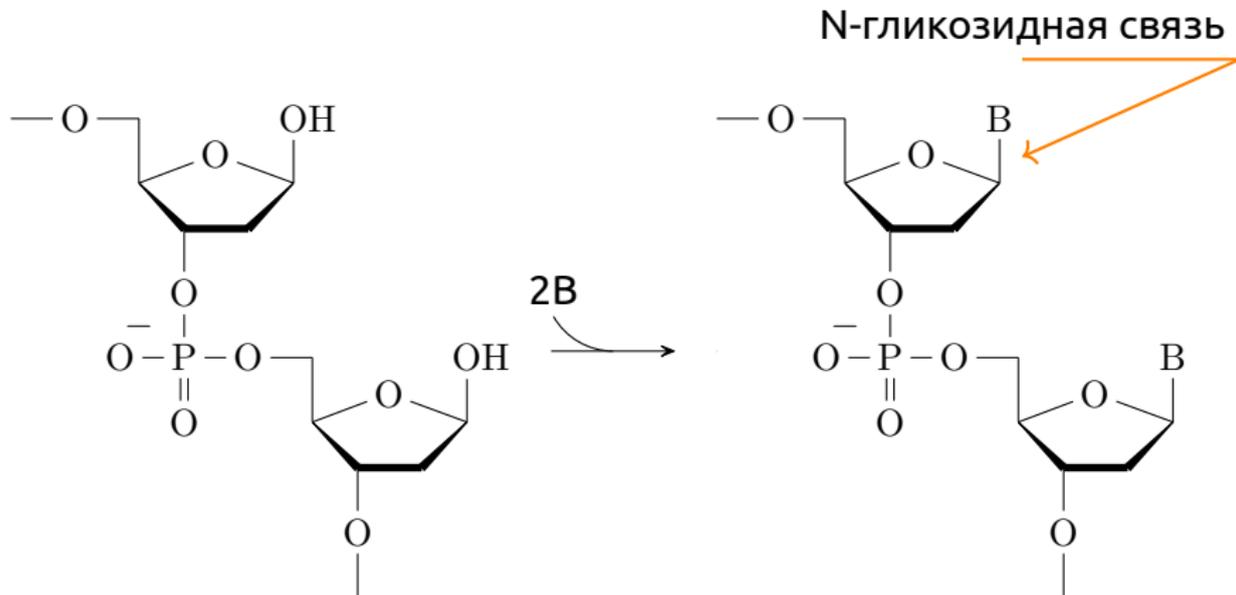


Урацил



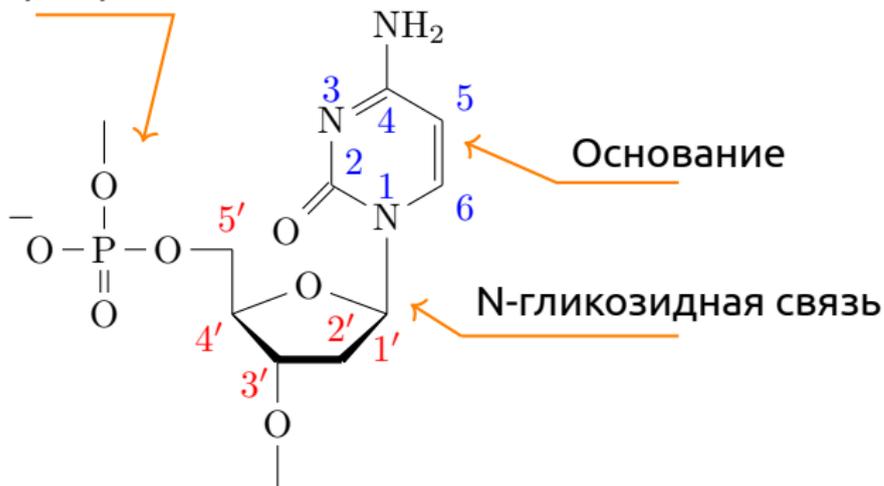
Тимин

Основания и остов



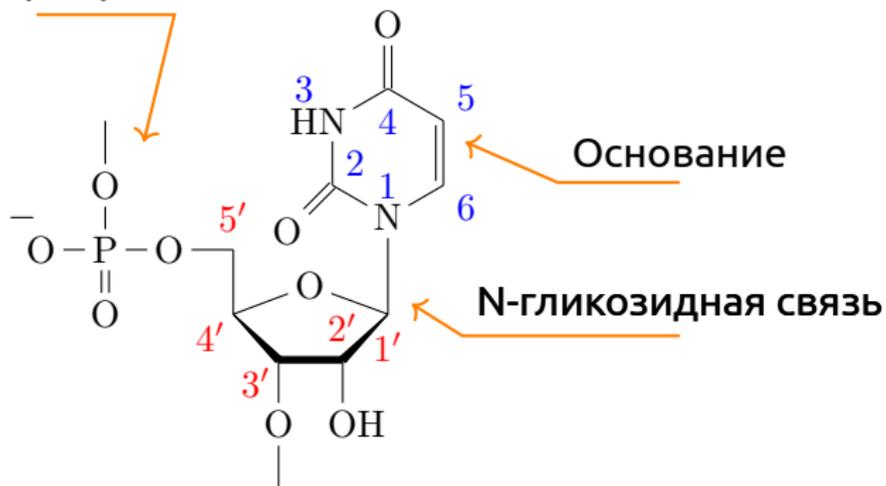
Нуклеотид ДНК

Ортофосфат

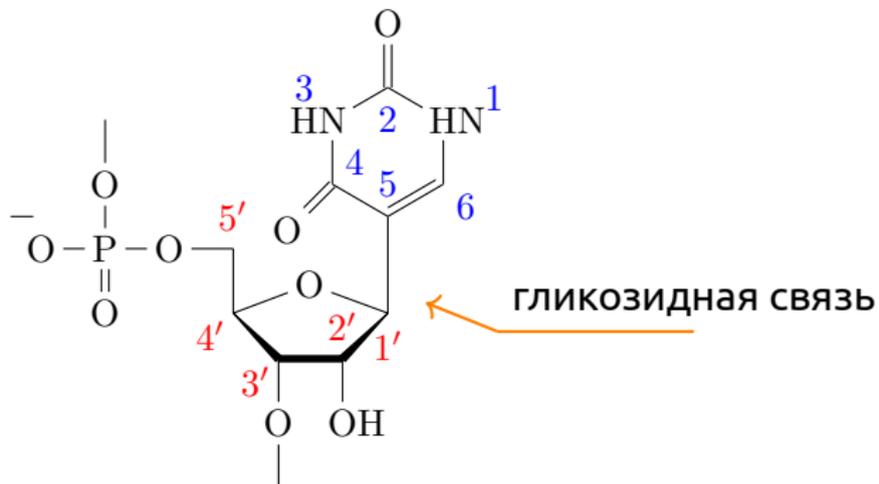


Нуклеотид РНК

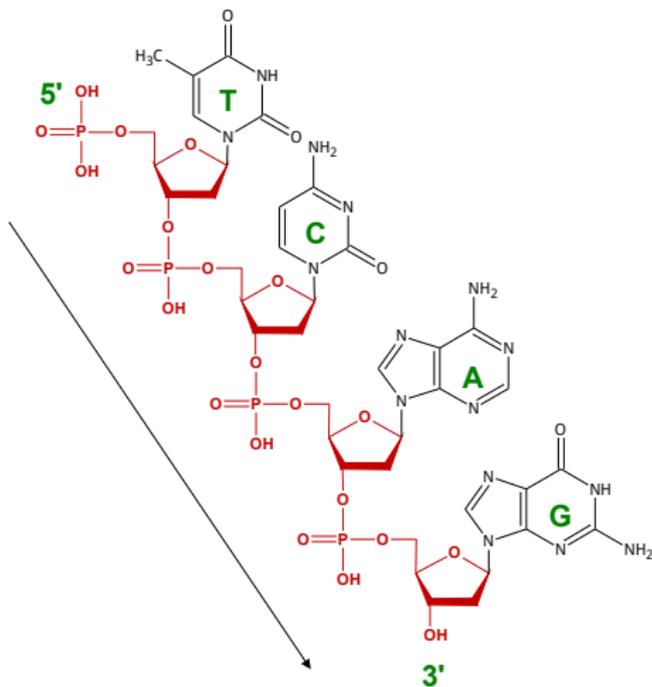
Ортофосфат



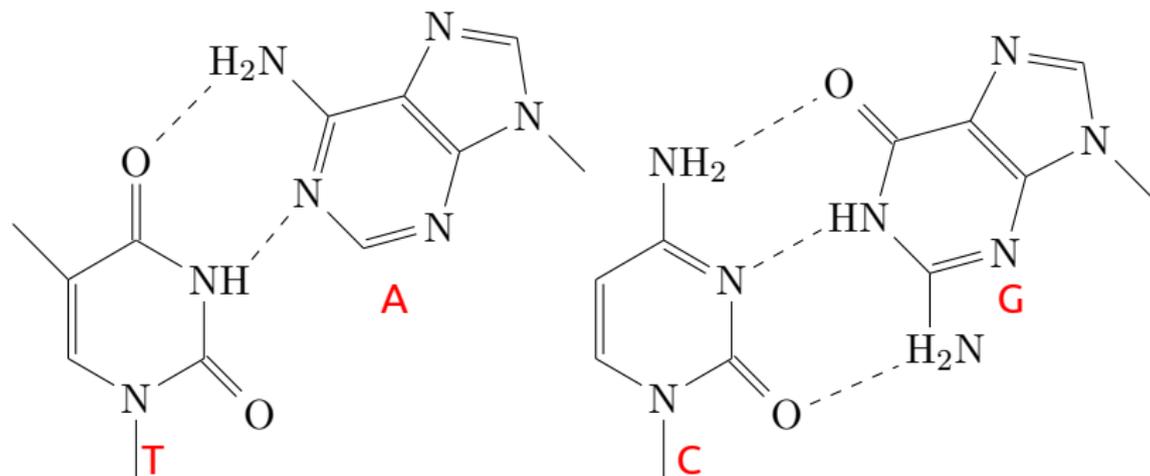
Псевдоуридин



Цепь ДНК

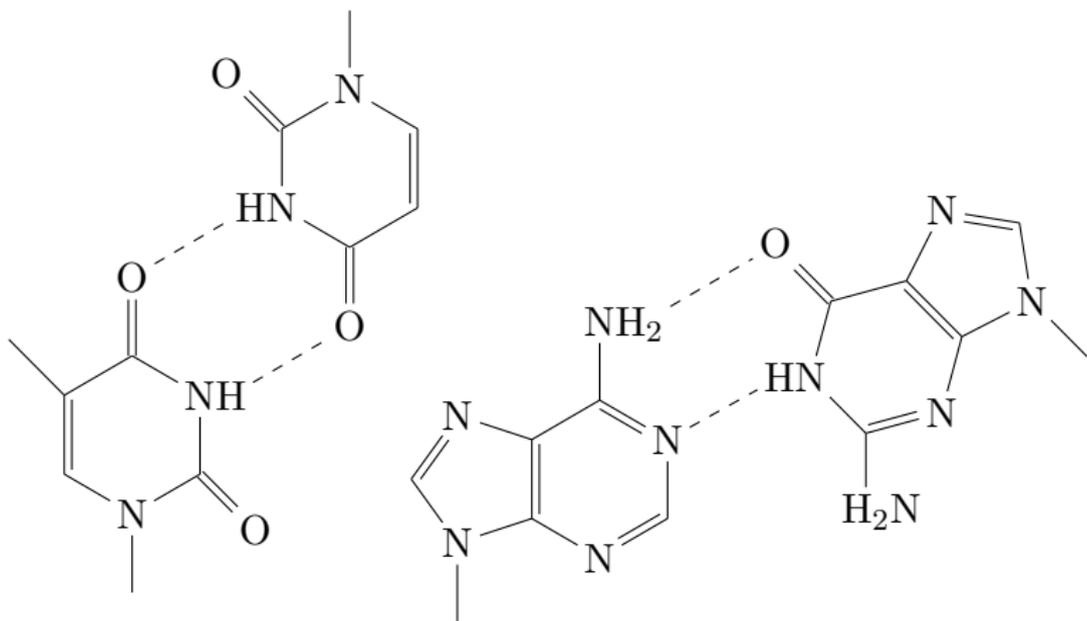


Канонические взаимодействия

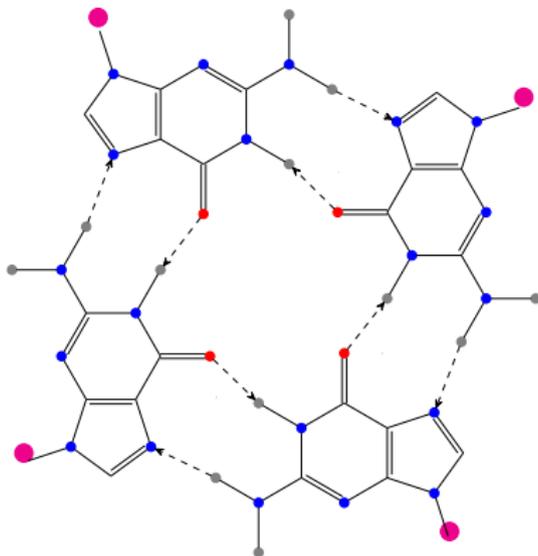


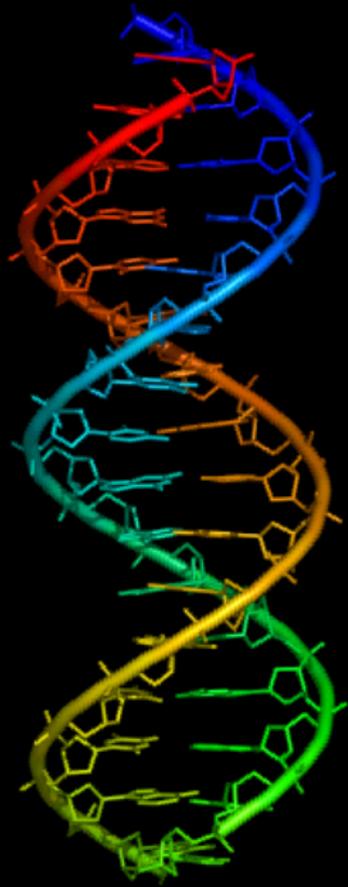
п

Неканонические взаимодействия



Неканонические взаимодействия





Структура ДНК

Структура ДНК впервые была предложена Watson и Crick в 1953 на основе результатов РСА низкого разрешения

Основные свойства:

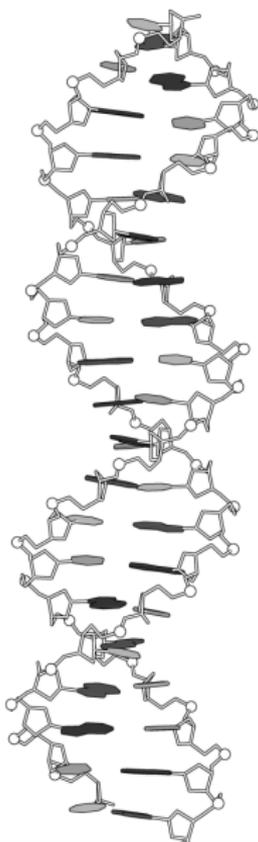
- Две антипараллельные цепи
- ДНК это двойная спираль
- Имеет две оси симметрии

Взаимодействия в ДНК

Два типа взаимодействий гетероциклических оснований в ДНК.

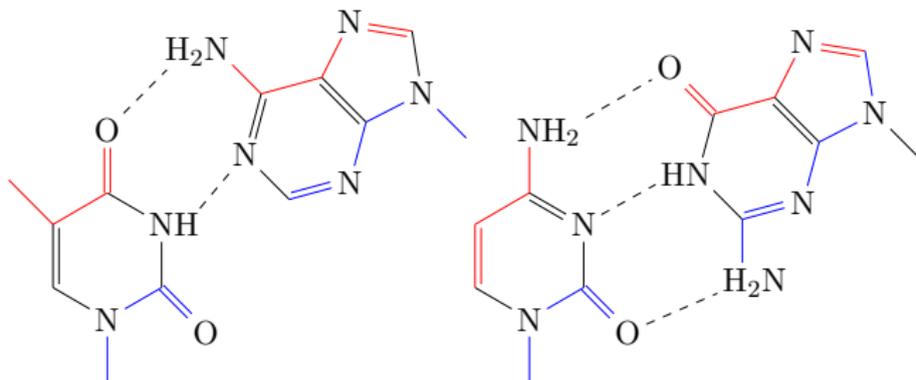
- Копланарные взаимодействия (в одной плоскости). В основном реализуются как водородные связи.
- Стопочные взаимодействия основаны на Ван-дер-Ваальсовых взаимодействиях.

Спираль ДНК

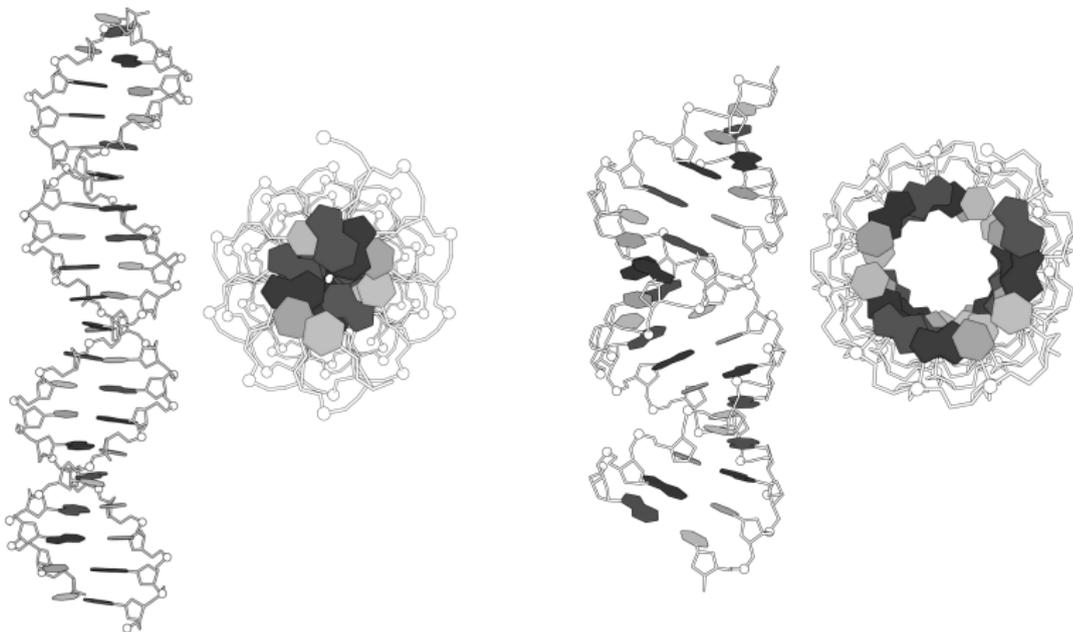


Атомы принадлежащие большой и малой бороздкам

	Большая бороздка	Малая бороздка
Adenine	C6, N6, C5, N7	C2, N3, C4, N9
Guanine	C6, O6, C5, N7, C8	C2, N2, N3, C4, N9
Cytosine	C6, C5, C4, N4	O2, N1, C2
Thymine	C6, C5, C4, O4, C5M	O2, C2, N1, C6

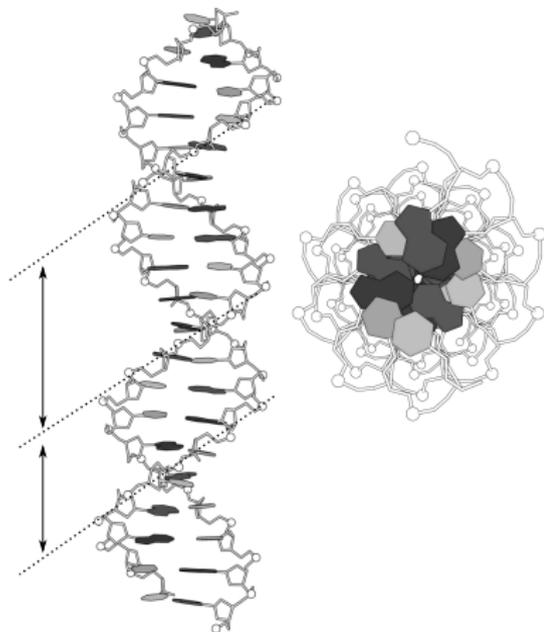


Регулярные формы спирали ДНК

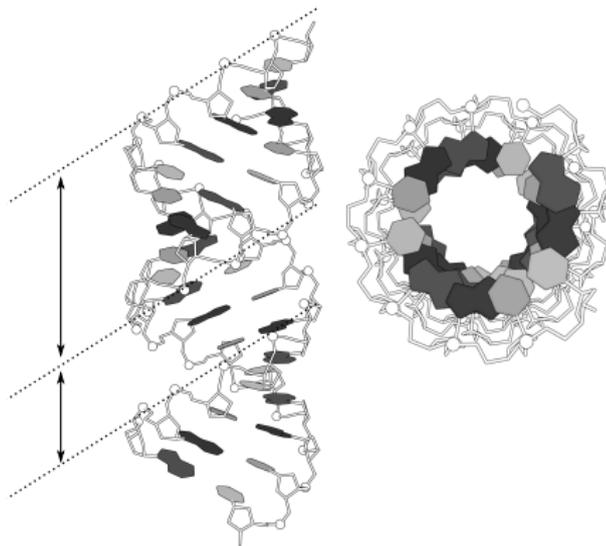


Регулярные формы спирали ДНК

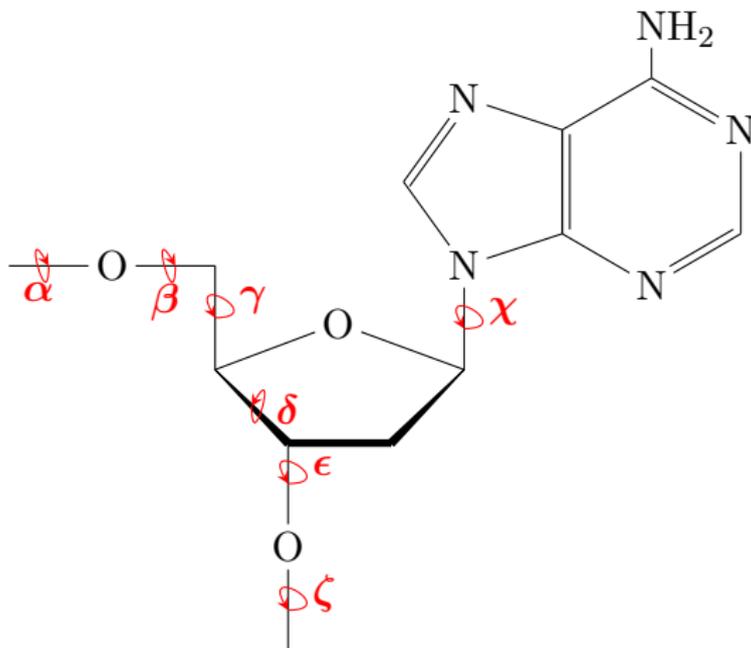
B-DNA



A-DNA



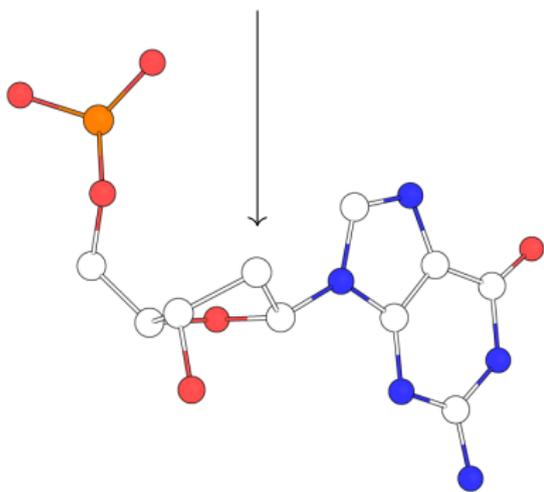
Торсионные углы НК



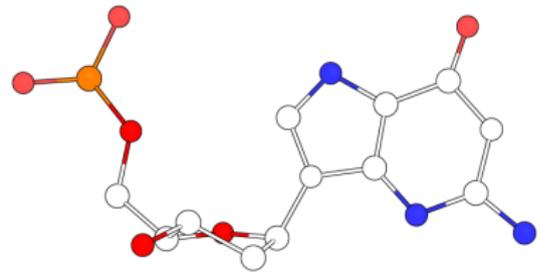
Форма	α	β	γ	δ	ϵ	ζ	χ
А-ДНК	62	173	52	88/3	178	-50	-160
В-ДНК	63	171	54	123/131	155	-90	-117

Конформация дезоксирибозы

B-DNA
C2'-endo



A-DNA



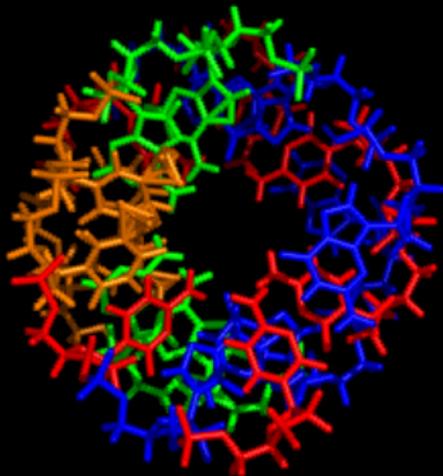
C2'-exo

Условия существования различных форм

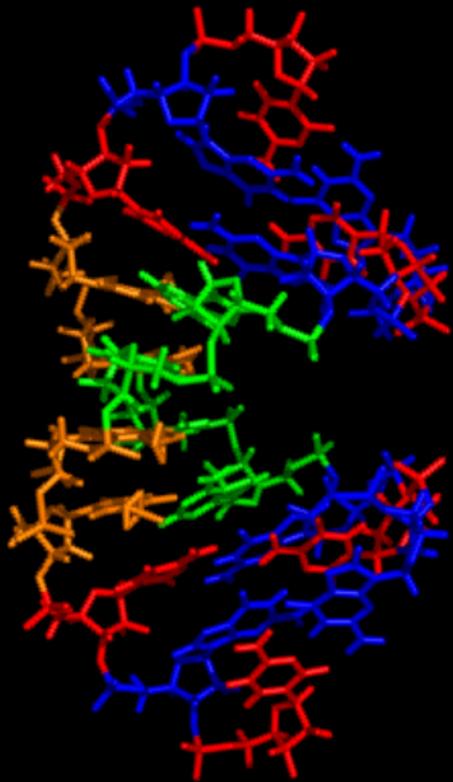
Разные формы ДНК переходят друг в друга при изменении условий внешней среды:

- В-форма стабильна при нормальных физиологических условиях
- Дегидратация, понижение относительной влажности до 0.75 инициирует переход $B \Rightarrow A$

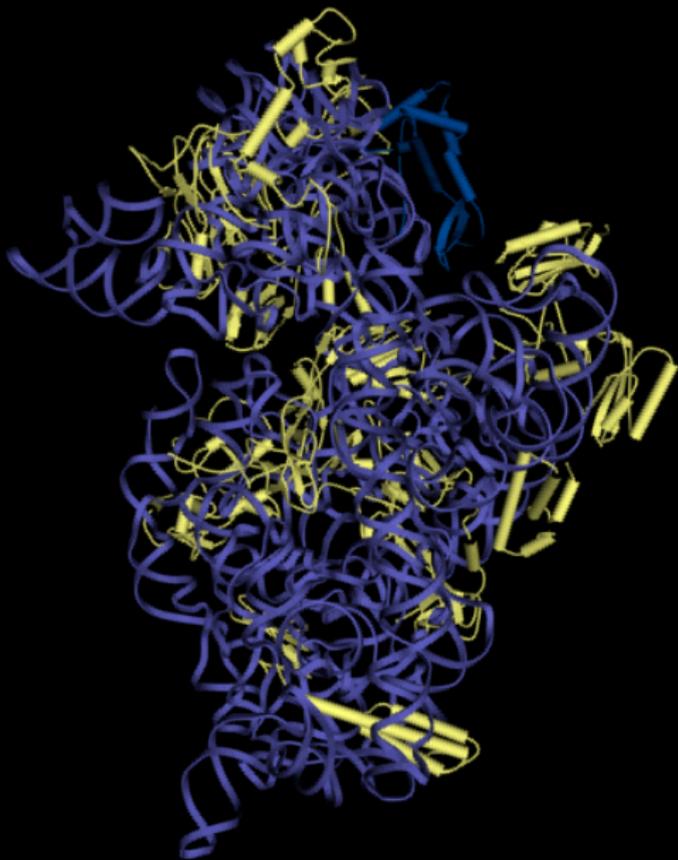
Пример: смеси вода-этанол(метанол) при росте доли спирта > 0.75 , переход $B \Rightarrow A$



Переход ДНК из А в В форму в солевом растворе



Переход ДНК из А в В форму в солевом растворе



Переход ДНК из А в В форму в солевом растворе

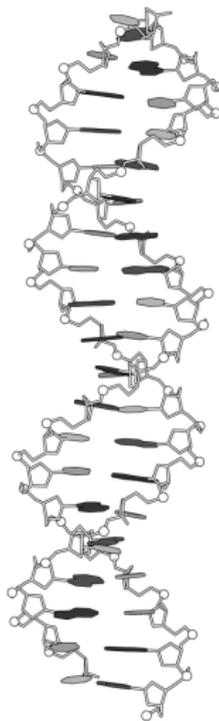
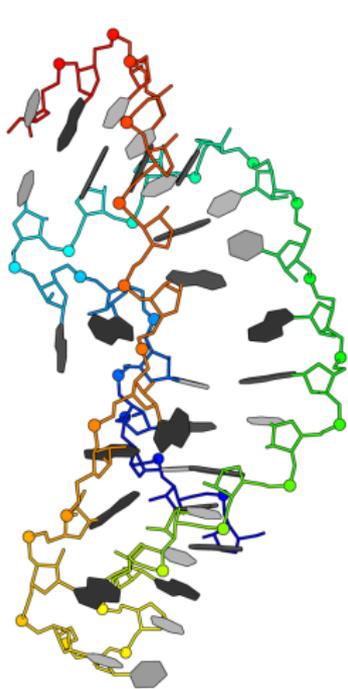
Структура РНК

Основные свойства.

- Одно-цепочечная молекула
- В клетке найдено множество видов РНК и каждый из них имеет специфичную функцию.

Основные типы: рРНК, мРНК, тРНК.

РНК и ДНК



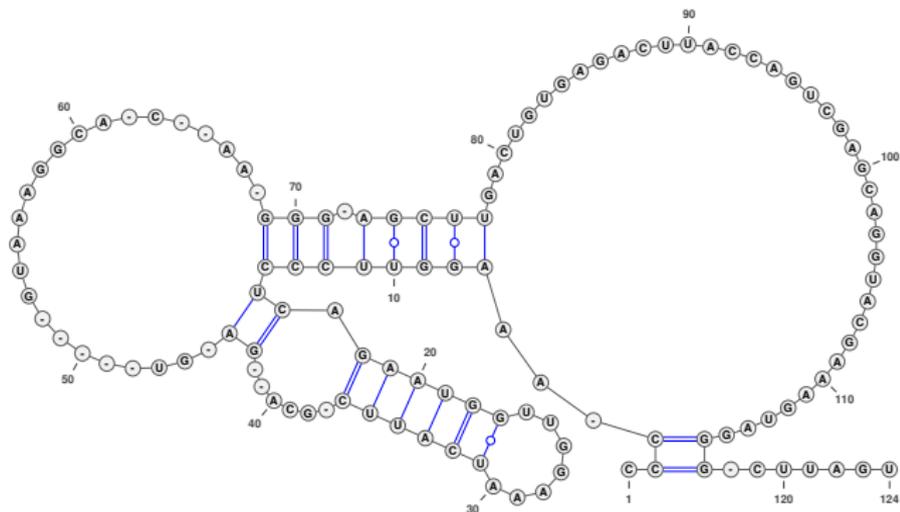
Почему важна структура РНК?

Структура РНК определяет её функцию:

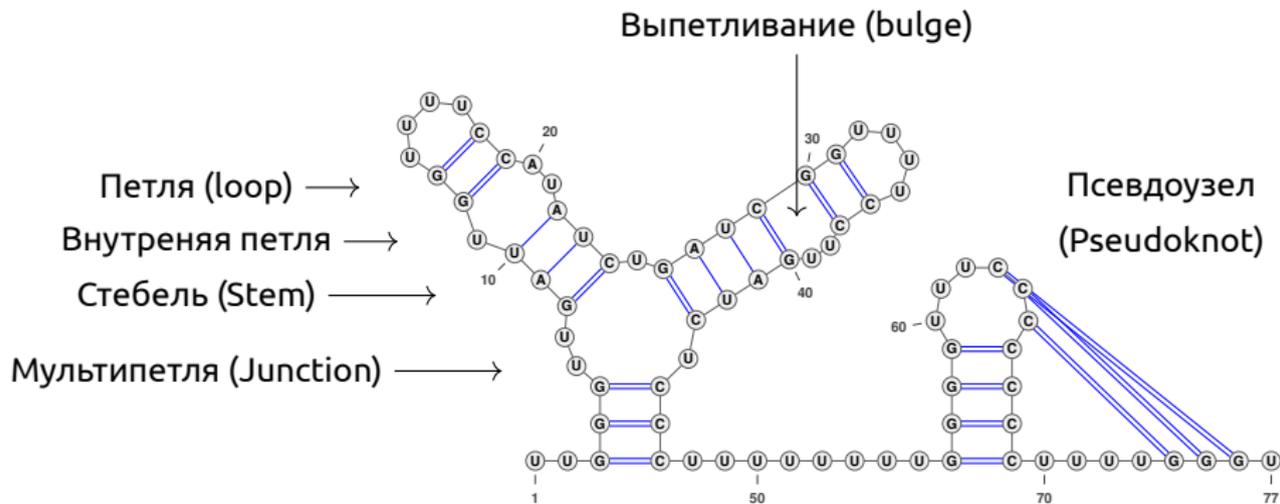
- Регуляторная
- Структурная
- Каталитическая (рибозимы)

Некоторые вирусы имеют РНК геном (HIV, грипп).

Вторичная структура РНК

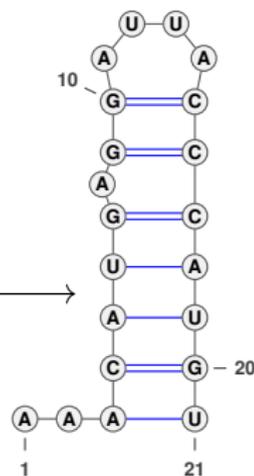
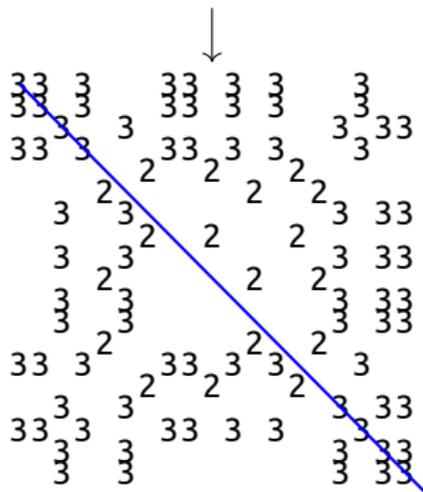


Вторичная структура РНК

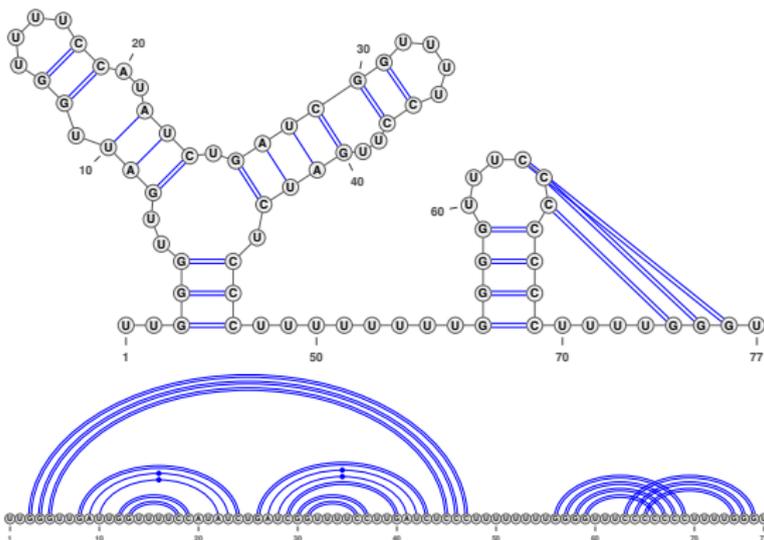


Возможность предсказания вторичной структуры

GGCGACUGGUGAGUACGCC



Вторичная структура РНК



UUGGGUUGAUUGGUUUUCCAUUUCUGAUCGGUUUCCUUGAUCUCCUUUUUUUUGGGUUUCCCCCUUUUGGGU
 --(((---(((---))--)))-((((---))--)))--))-----(((---[[[]]))----]])-

*VARNA: Visualization Applet for RNA secondary structure

Расчёт энергии структуры по алгоритму Зукера

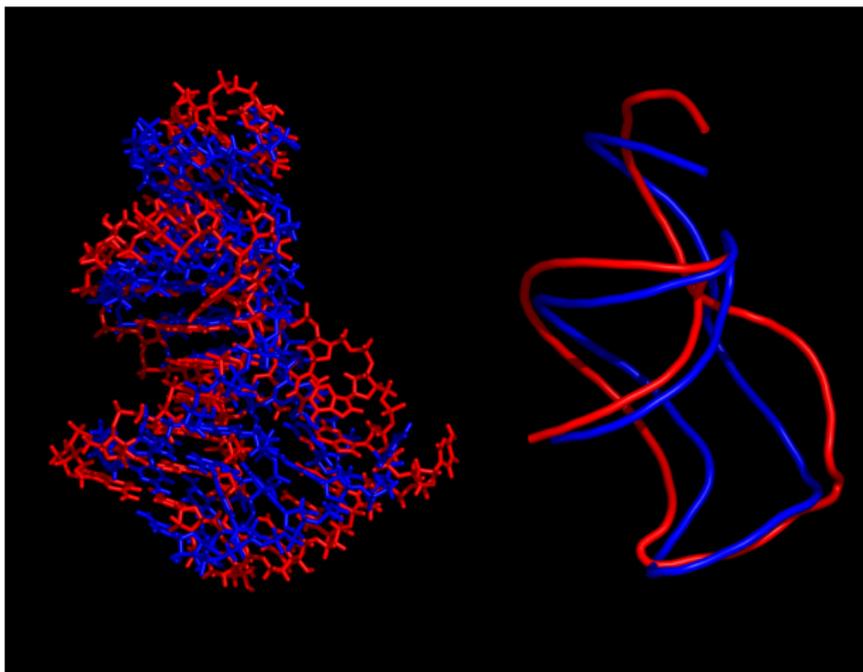


$$\Delta G = 3.2 \text{ Ккал/моль}$$

РНК, которые трудно предсказать алгоритмом Зукера

- РНК связанная с белками.
- Длинные РНК.
- Псевдоузлы.

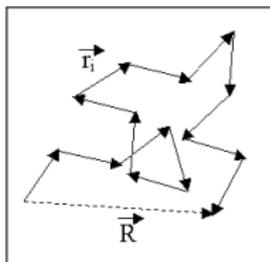
Моделирование структуры



Freely Jointed Chain

Имеется полимер из N мономеров, тогда не упакованная длина
: $L = Nl$

- сегменты между собой не взаимодействуют
- полимер флуктуирует и его форма определяется простым распределением
- Если полимер образует глобулу:

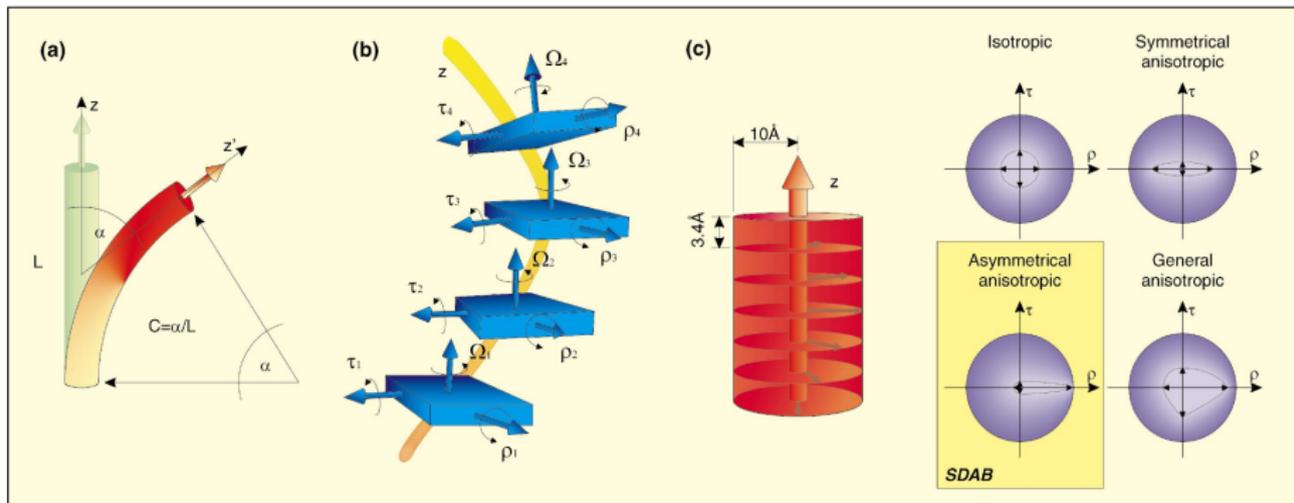


$$\langle R^2 \rangle = Nl^2 = Ll; \quad P(\vec{R}) = \left(\frac{3}{2\pi Nl^2} \right)^{3/2} e^{-\frac{3\vec{R}^2}{2Nl^2}}$$

$$\sqrt{\langle \vec{R}^2 \rangle} = \sqrt{N}l = \sqrt{Ll}$$

Rod model

ДНК представляется как стержень:



$$\Delta G = \frac{1}{2} E I L \alpha^2$$

E - коэффициент Юнга

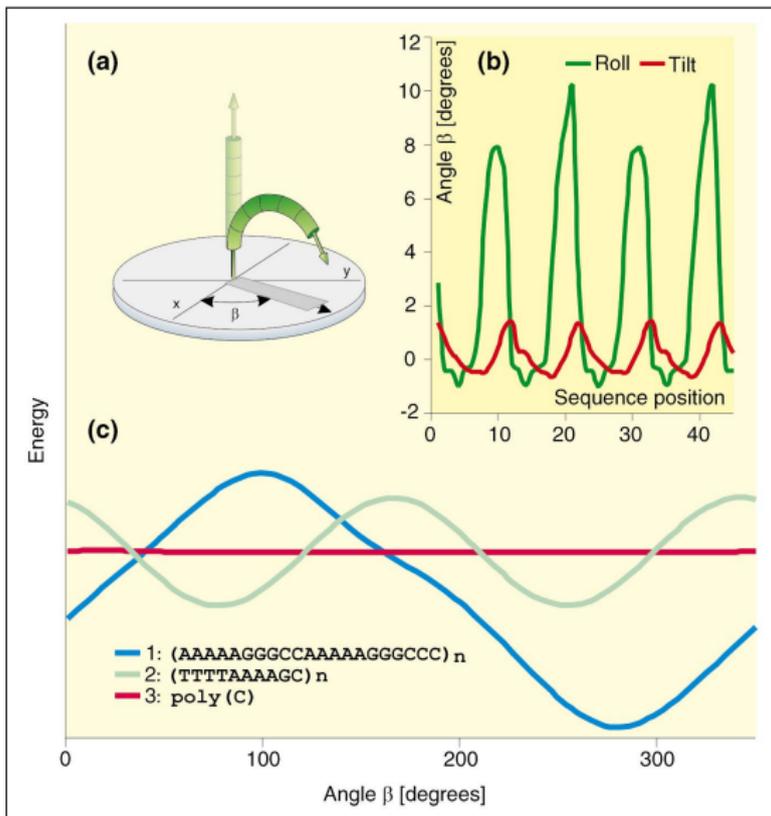
I - момент инерции, для цилиндра радиусом r , $I = \frac{\pi r^4}{4}$

Rod model, последовательность контекста

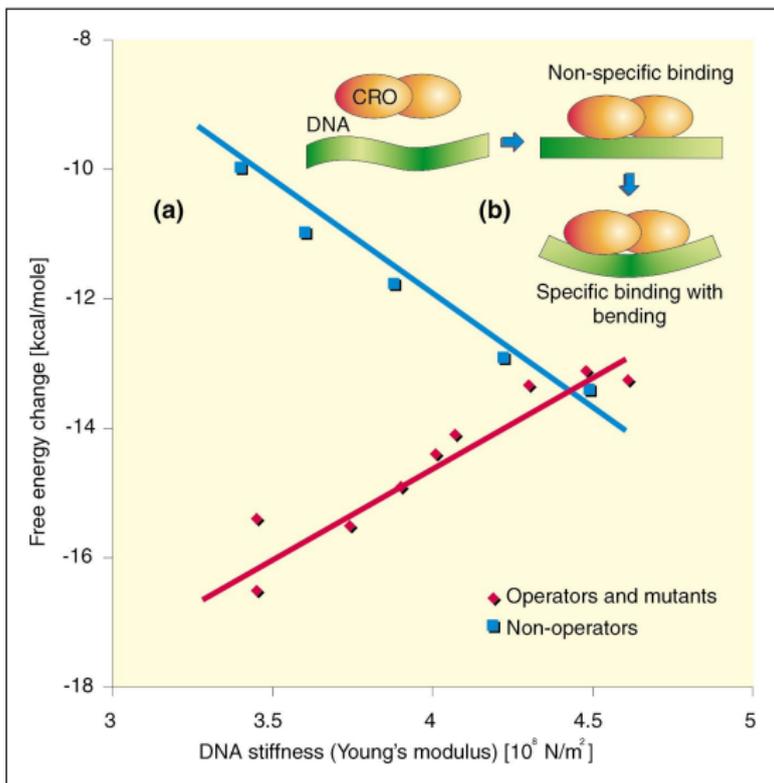
Table I. Sequence-dependent Young's modulus values calculated for various bending models^{26,34,40}

Trinucleotide	Relative bendability (arbitrary units)		Young's modulus (10^9Nm^{-2})	
	DNaseI scale	Consensus scale	DNaseI scale	Consensus scale
AAA/TTT	0.1	0.05	2.307	2.23
AAC/GTT	1.6	2.65	2.016	1.71
AAG/CTT	4.2	4.70	1.523	1.30
AAT/ATT	0.0	0.35	2.327	2.17
ACA/TGT	5.8	5.50	1.225	1.15
ACC/GGT	5.2	5.30	1.336	1.18
ACG/CGT	5.2	5.30	1.336	1.18
ACT/AGT	2.0	7.80	1.94	1.46
AGA/TCT	6.5	4.90	1.096	1.26
AGC/GCT	6.3	6.90	1.133	0.87
AGG/CCT	4.7	5.05	1.429	1.23
ATA/TAT	9.7	6.25	0.519	0.99
ATC/GAT	3.6	4.45	1.636	1.35
ATG/CAT	8.7	7.70	0.697	0.71
CAA/TTG	6.2	4.75	1.151	1.29
CAC/GTG	6.8	6.65	1.041	0.92
CAG/CTG	9.6	6.90	0.536	0.87
CCA/TGG	0.7	3.05	2.19	1.63
CCC/GGG	5.7	5.85	1.244	1.08
CCG/CGG	3.0	3.85	1.749	1.47
CGA/TCG	5.8	7.05	1.225	0.84
CGC/GCG	4.3	5.90	1.504	1.07
CTA/TAG	7.8	5.00	0.859	1.24
CTC/GAG	6.6	6.00	1.078	1.05
GAA/TTC	5.1	4.05	1.355	1.43
GAC/GTC	5.6	5.50	1.262	1.14
GCA/TGC	7.5	6.75	0.914	0.90
GCC/GGC	8.2	9.10	0.787	0.45
GGA/TCC	6.2	5.00	1.151	1.24
GTA/TAC	6.4	5.05	1.115	1.23
TAA/TTA	7.3	4.65	0.95	1.31
TCA/TGA	10.0	7.70	0.465	0.71

Rod model, гибкость



Rod model, биология



Worm Like Chain model

Суть модели это непрерывное описание цепи для решения ряда ограничений:

- Энтропийная эластичность ДНК цепи состоит из малых девиаций по оси молекулы из-за температуры
- Направление цепи коррелирует с длиной цепи, называемой "the persistence length".

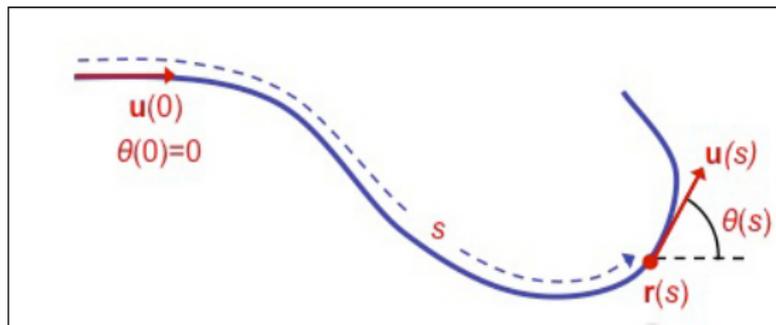
Для ДНК в 10 mM растворе NaCl, $P_{DNA} = 150$ п.о. или 550 nm.

- Силы порядка $\frac{k_B T}{P}$ нужны для выравнивания и направления единиц эластичности вдоль оси полимера.

Worm Like Chain model

Фиксированная величина длины контура, L и некая стандартная величина "сгибаемости" C .

$$U = C \int_{s=0}^{s=L} \left| \frac{\partial t}{\partial s} \right|^2$$



Подытожим:

- Для сил действующих в диапазоне <100 fN и >5 pN модель FJC работает хорошо.

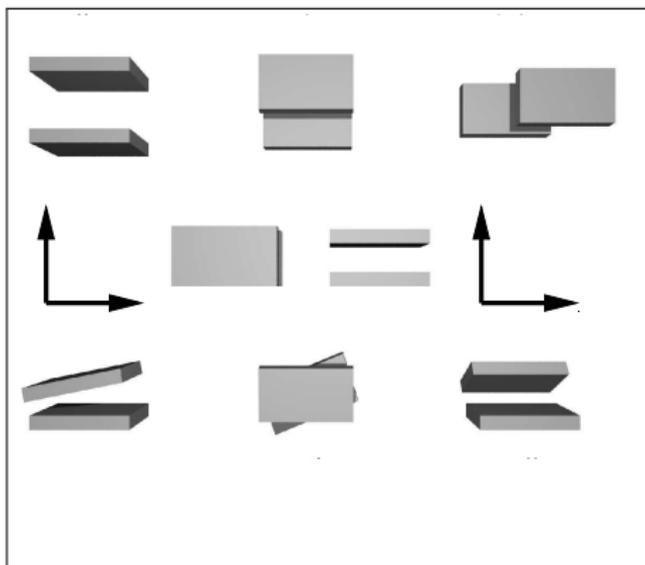
$$f = \frac{k_B T}{b} \frac{1}{1 - z/L}$$

- WLC модель работает во всех диапазонах.

$$f = \frac{k_B T}{P} \left[\frac{1}{4(1 - z/L)^2} - \frac{1}{4} + \frac{z}{L} \right]$$

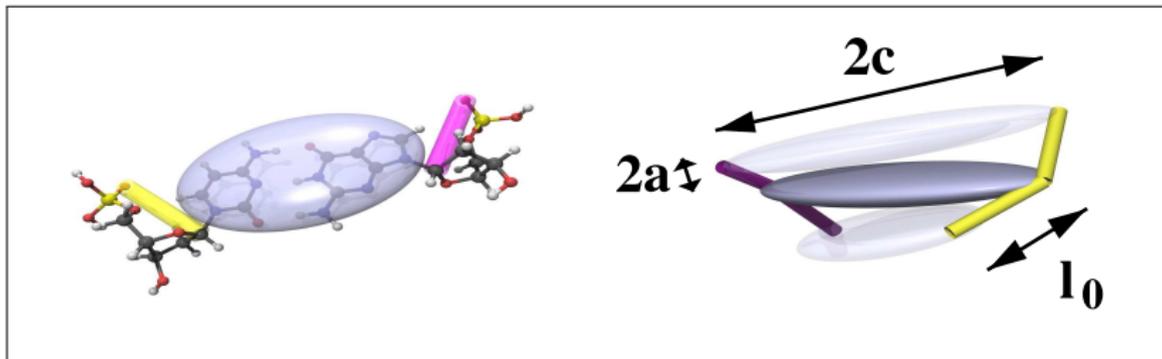
Моделирование ДНК на уровне пар оснований

Рассмотрим типы движений пар оснований друг относительно друга.



Rise (Ri), Slide (Sl), Shift (Sh), Twist (Tw), Roll (Ro), and Tilt (Ti)

Потенциал

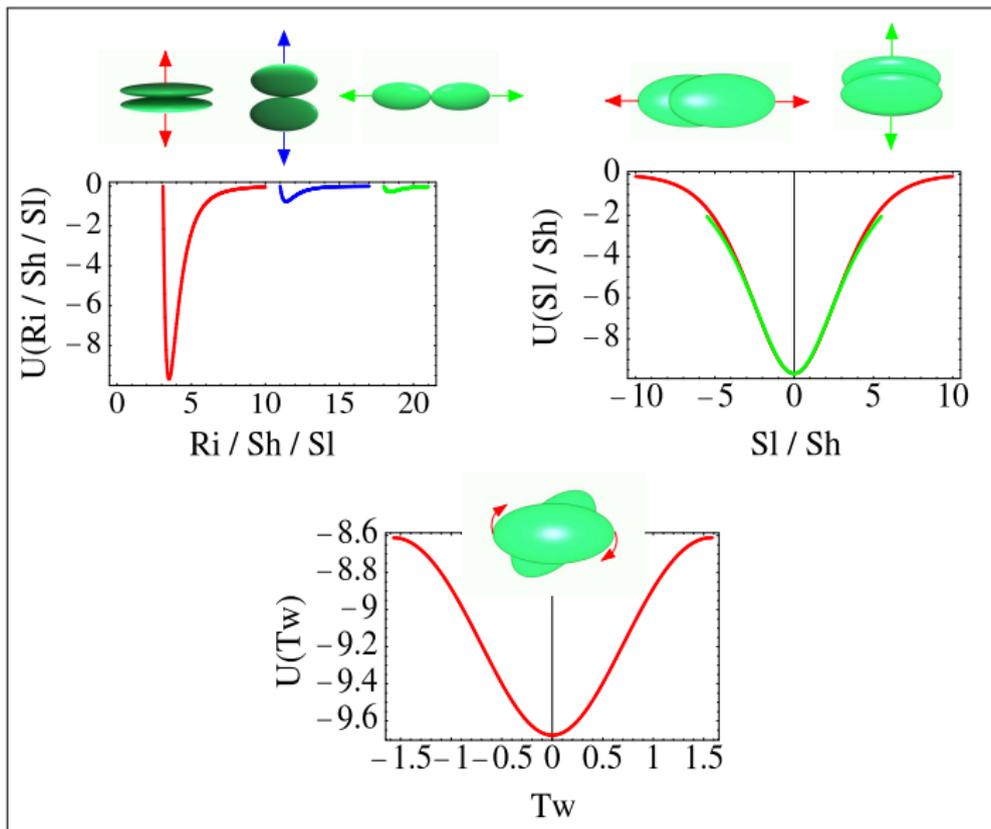


$$U(A_1, A_2, r_{12}) = U_r \eta_{12} \chi_{12}$$

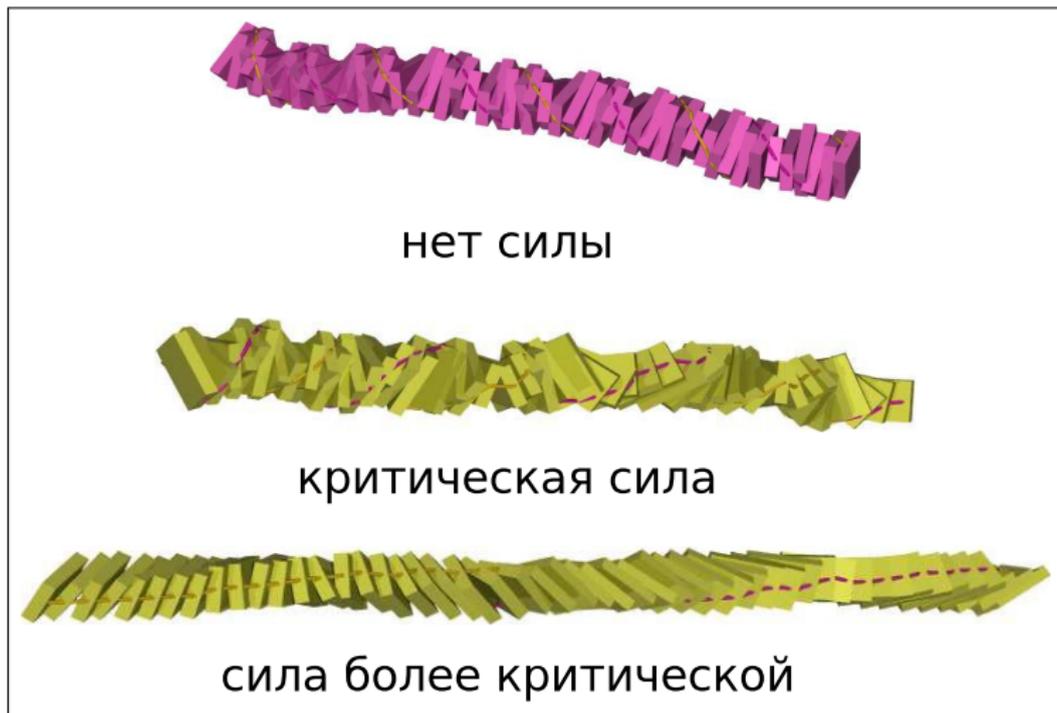
$$U_r = 4\epsilon \left(\left(\frac{\sigma}{h + \gamma\sigma} \right)^{12} - \left(\frac{\sigma}{h + \gamma\sigma} \right)^6 \right)$$

η_{12} и χ_{12} отвечают за силу взаимодействия в зависимости от ориентации эллипсов.

Вид зависимостей

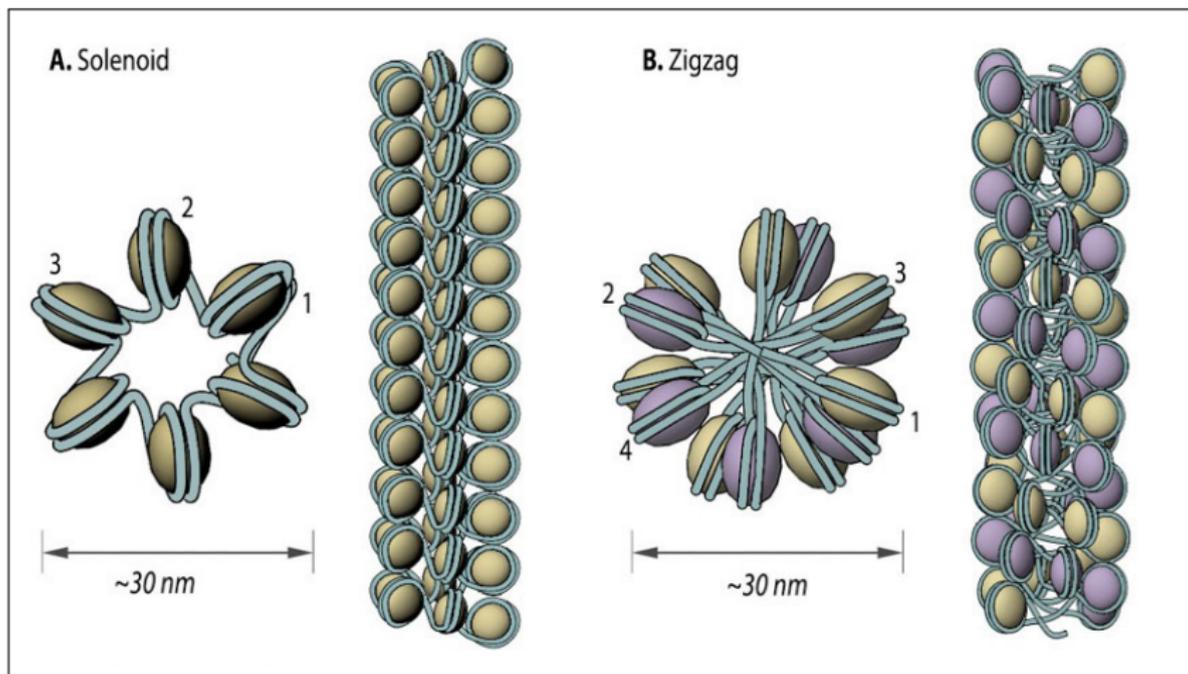


Растягивание

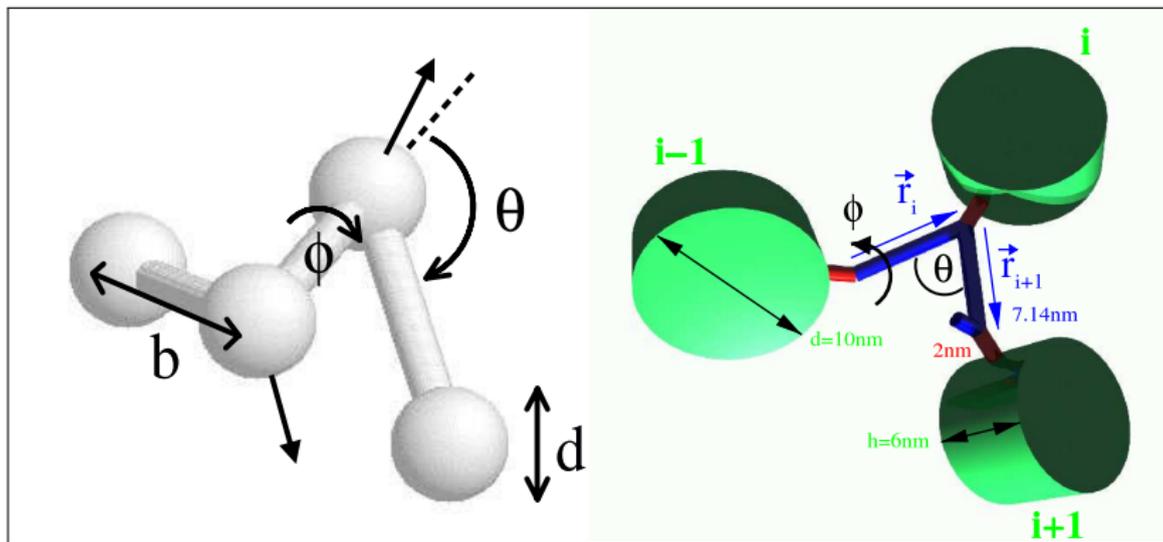


Значения сил совпадают с экспериментальными. Локальные деформации реалистичны.

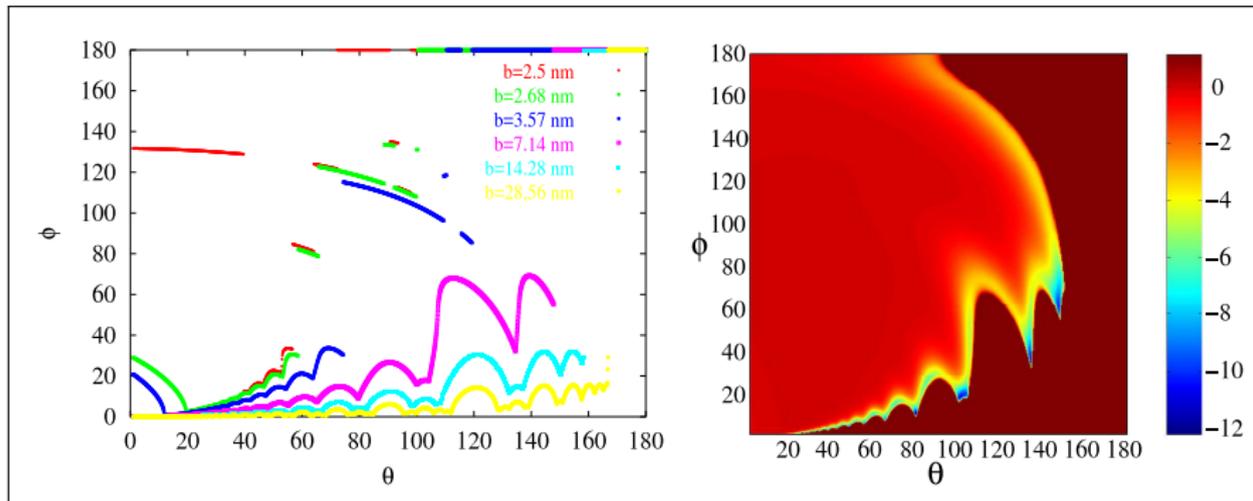
Хроматин



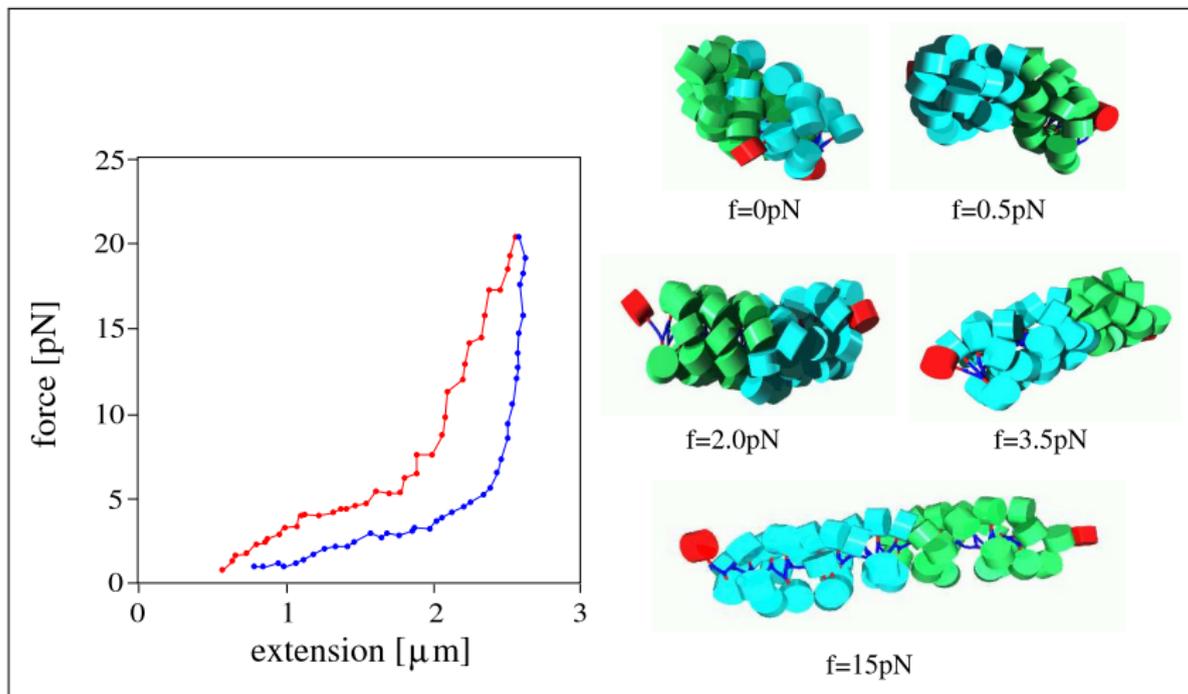
Хроматин, модель блока



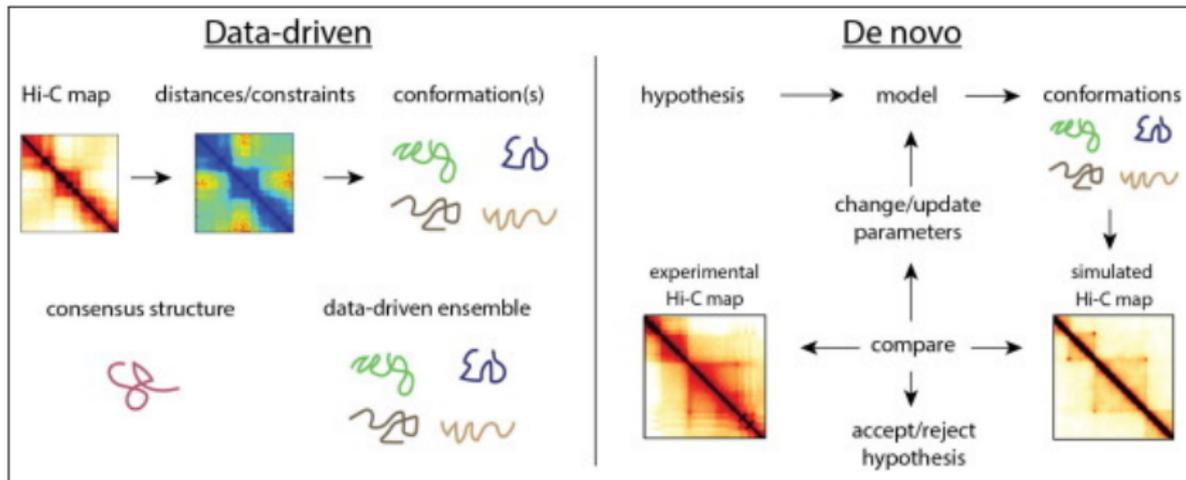
Хроматин, запрещённые зоны



Хроматин, растягивание

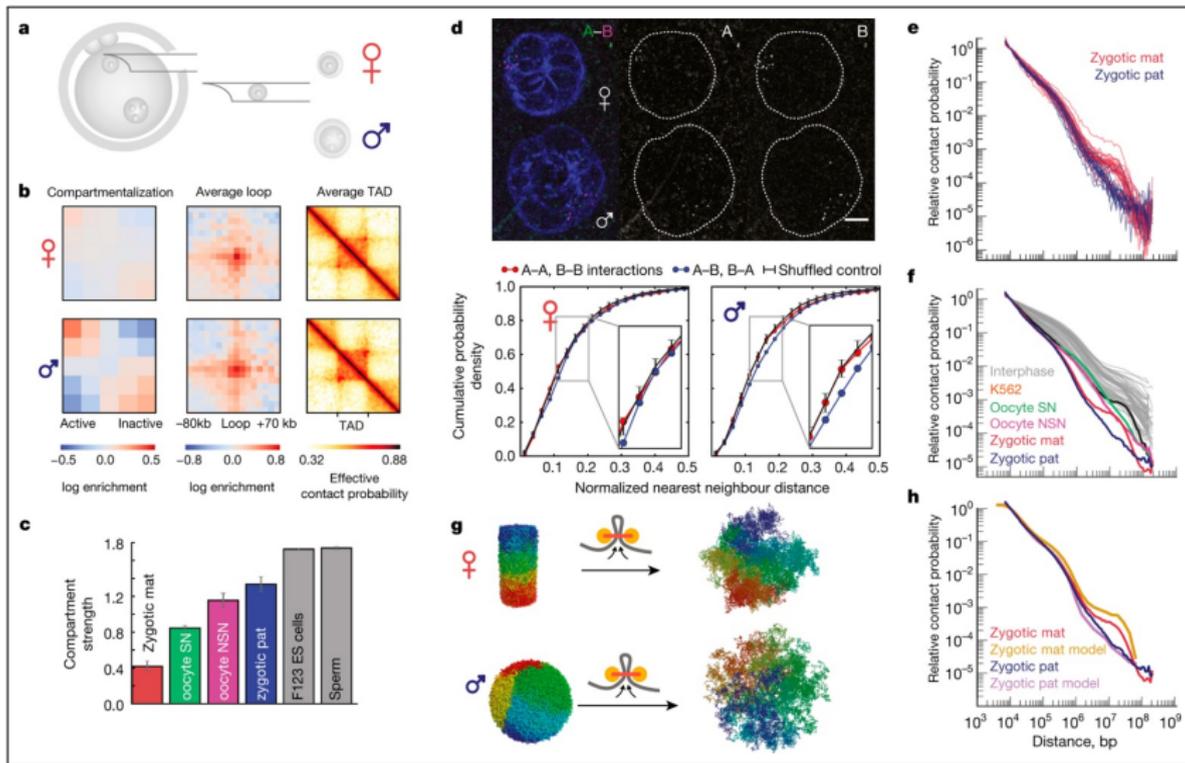


HiC

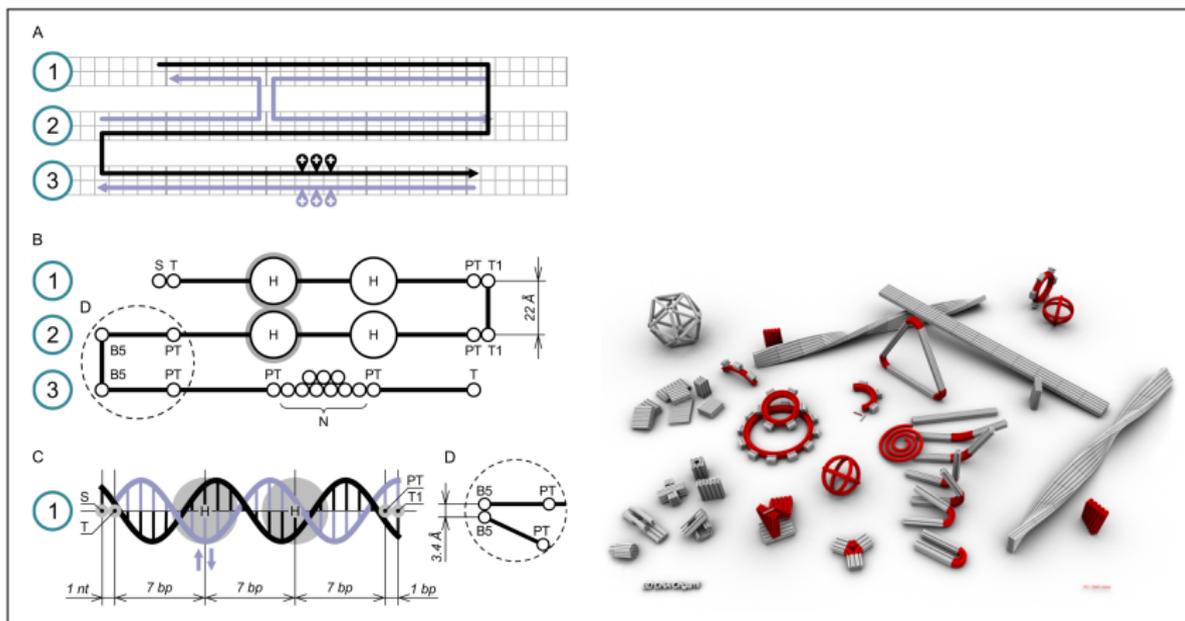




chromatin architecture in haploid nuclei

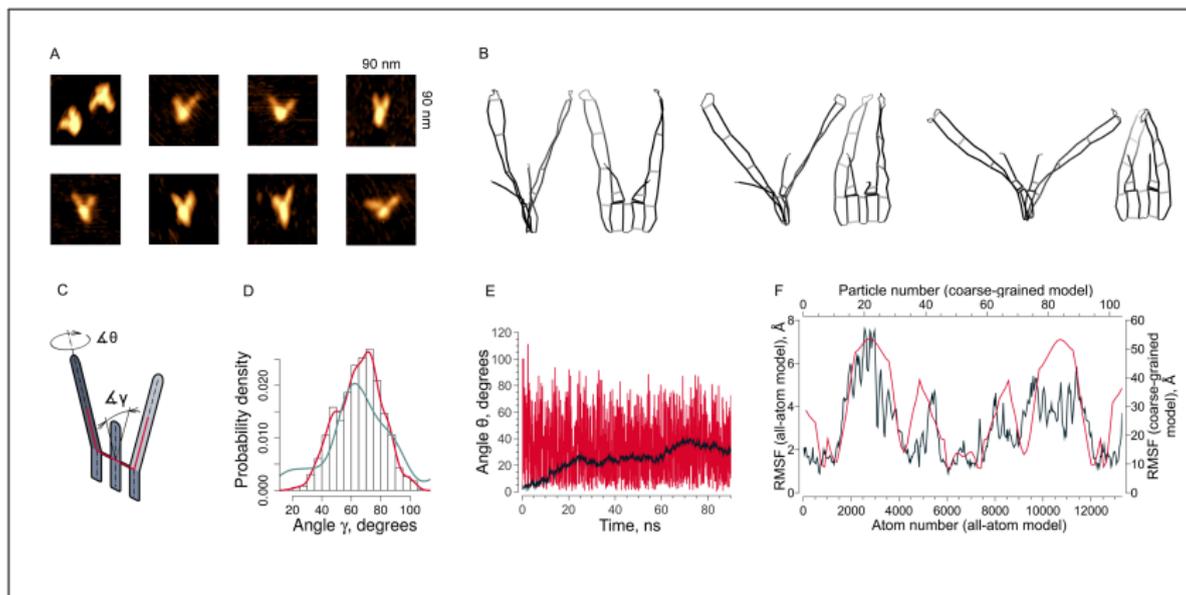


ДНК-оригами



doi: 10.1093/nar/gkx1262

Моделирование ДНК-оригами



10.1093/nar/gkx1262