

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ  
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ  
«МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ  
имени М.В.ЛОМОНОСОВА»  
ФАКУЛЬТЕТ ФУНДАМЕНТАЛЬНОЙ ФИЗИКО-ХИМИЧЕСКОЙ ИНЖЕНЕРИИ**

Биологический факультет МГУ имени М.В. Ломоносова  
Кафедра биоинженерии  
Группа интегративной биологии

**БАКАЛАВРСКАЯ РАБОТА**

**«Анализ взаимодействий пионерных транскрипционных факторов с  
белками ядра клеток человека с помощью методов искусственного  
интеллекта»**

Выполнила студентка  
4 курса 401 группы  
Хасанова Ума Наурузовна  
Научный руководитель  
(д.ф.-м.н., профессор, чл.-корр. РАН)  
Шайтан Алексей Константинович  
Научный консультант  
(м.н.с.)  
Грибкова Анна Кирилловна

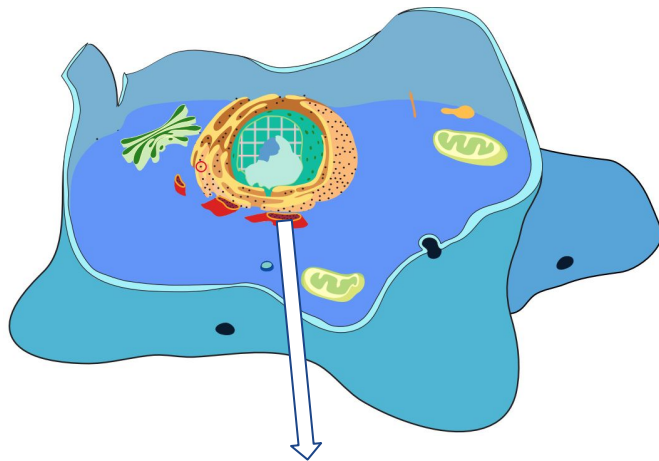


Москва  
2025

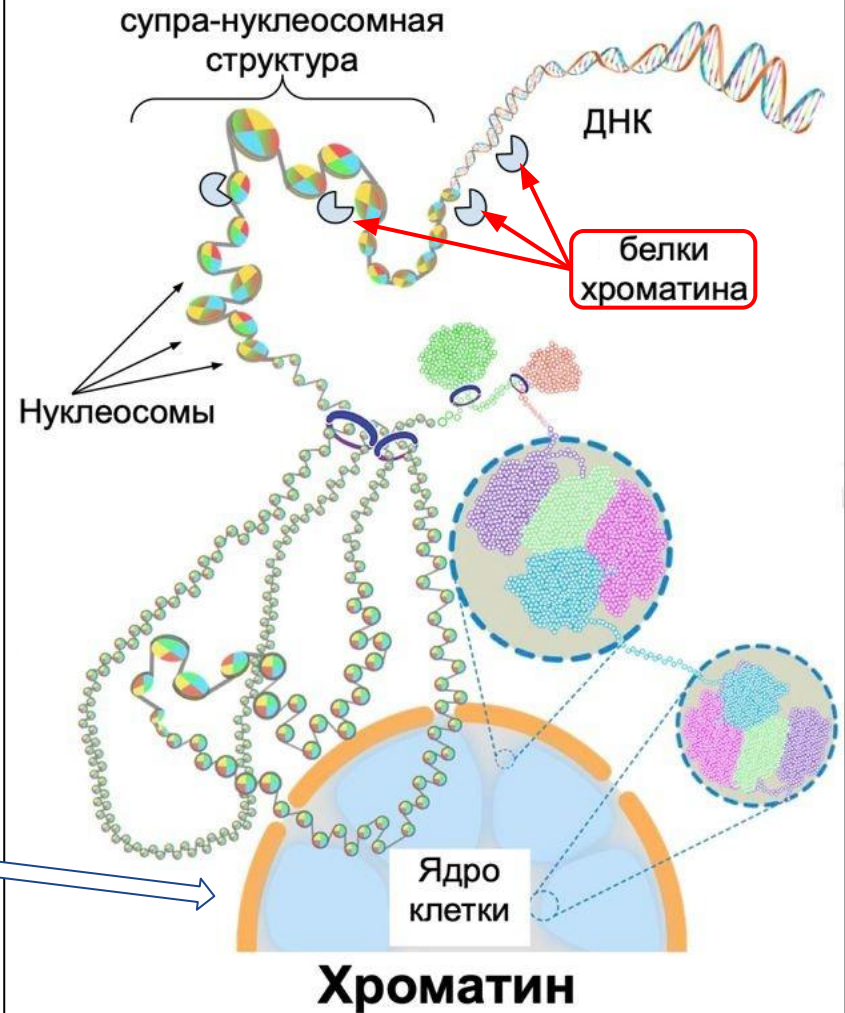


# Введение. Организация ДНК в клетке

## Клетка человека



## Ядро

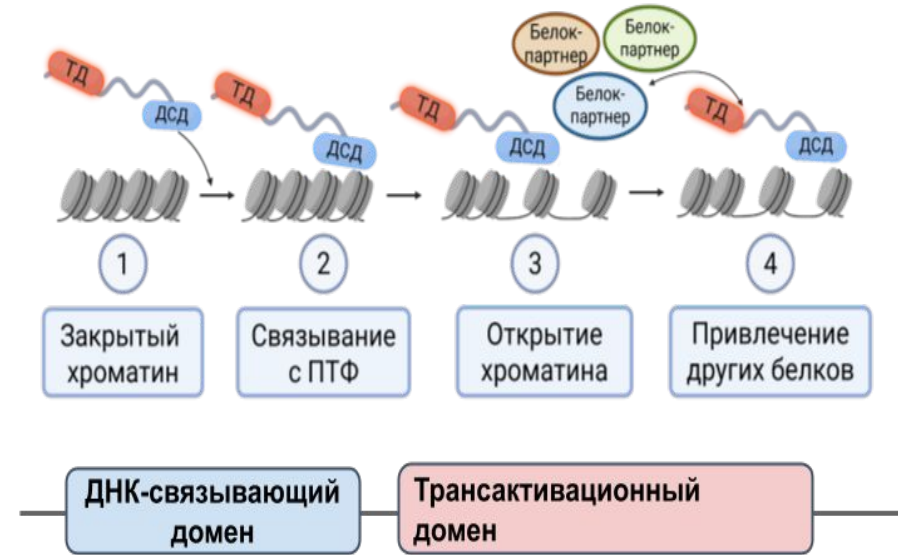


Шайтан А.К., диссертация д.ф-м.н. 2021  
(с изменениями)

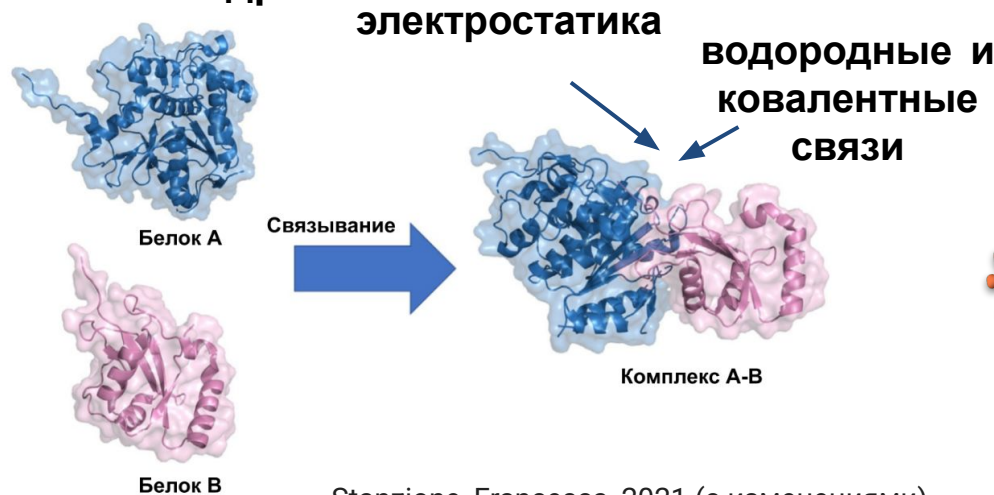
# Введение. Пионерные транскрипционные факторы

## Пионерные транскрипционные факторы (ПТФ) - особые белки, которые:

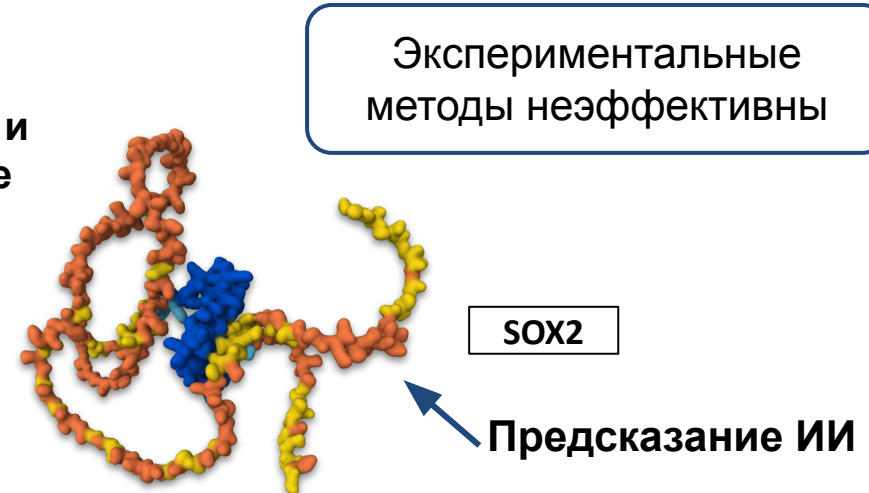
- связываются с “закрытой” ДНК, открывая доступ другим белкам
- **SOX2, KLF4, OCT4** способны перепрограммировать клетки в стволовые (“коктейль Яманаки”, Нобелевская премия по медицине, 2012)
- перспективны для лечения болезней (рак, регенеративная медицина)



## Как именно ПТФ взаимодействуют с белками ядра?



Stanzione, Francesca, 2021 (с изменениями)



AlphaFoldDB

# Цель и задачи работы

**Цель работы:** предсказание и анализ белок-белковых взаимодействий между пионерными транскрипционными факторами (ПТФ) и ядерными белками человека с использованием методов генеративного ИИ.

## **Задачи:**

1. Формирование репрезентативного набора ядерных белков и предсказание комплексов ПТФ с ядерными белками;
2. Структурный анализ предсказанных белковых комплексов;
3. Функциональная характеристика и анализ представленности в экспериментальных базах данных белков-партнеров.

# Методы. Общая схема исследования

## ЭТАП

## ОПИСАНИЕ

## РЕЗУЛЬТАТ

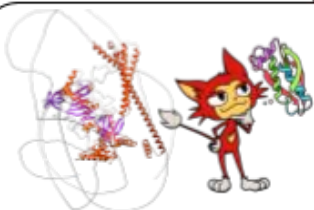


**ОТБОР И  
ФИЛЬТРАЦИЯ  
ЯДЕРНЫХ БЕЛКОВ**

Сбор по базам данных

Фильтрация по длине белков,  
локализации, гомологии

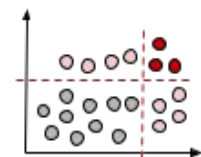
**3557  
белков  
ядра**



**ПОСТРОЕНИЕ  
КОМПЛЕКСОВ**

Запуск AlphaFold2-Multimer с  
тремя циклами переработки и  
использованием шаблонов из  
PDB

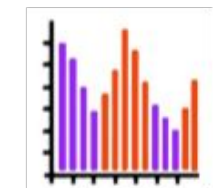
**10632  
комплекса**



**ОЦЕНКА КАЧЕСТВА  
СТРУКТУР**  
pDockQ, IPSAE, iPTM

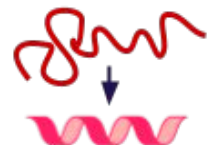
Разделение на комплексы  
среднего и высокого качества

**1453 и 267  
комплекса**



**АНАЛИЗ БЕЛОК-  
БЕЛКОВЫХ  
КОНТАКТОВ**

Число контактов, типы  
контактирующих аминокислот



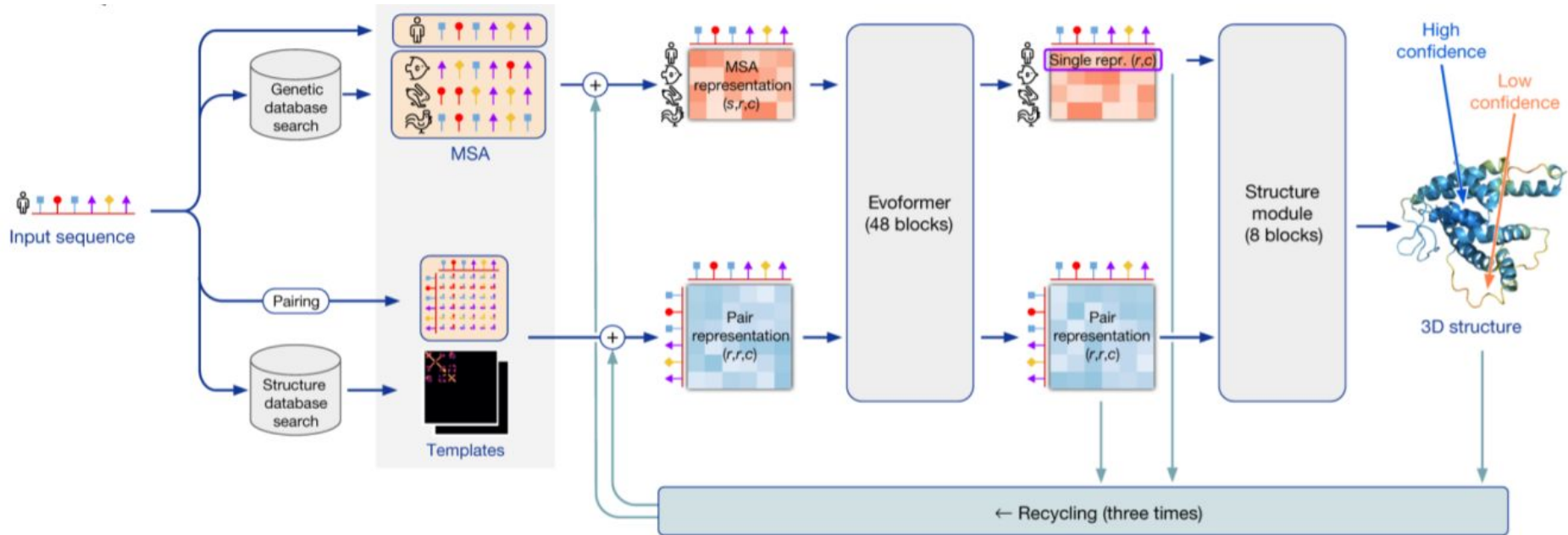
**АНАЛИЗ  
КОНФОРМАЦИОННЫХ  
ПЕРЕХОДОВ**

Исследование динамики  
неупорядоченных регионов,  
идентификация мотивов  
связывания

# Методы. AlphaFold2

Нобелевская премия по химии, 2024

Jumper et al., 2021



**Вход → Эвоформер → Структурный модуль → Выход**

Аминокислотная  
последовательность  
белка

Анализ коэволюции  
аминокислот и  
парных  
представлений

3D-координаты  
атомов с учетом  
геометрических  
ограничений (длины  
связей, углы)

3D-структура с  
оценкой  
достоверности  
предсказания (pLDDT,  
pTM, ipTM)



# Методы. Метрики качества

**ipTM** - оценка достоверности предсказания интерфейса взаимодействия

$$pTM_{ij} = \sum_{b=1}^{64} p_{ij}^b \left( \frac{1}{1 + \left( \frac{\Delta_b}{d_0} \right)^2} \right) \quad ipTM(A \rightarrow B) = \max_{i \in A} \left[ \text{mean}_{i \in B} (pTM_{ij}) \right]$$

$$ipTM = \max[ipTM(A \rightarrow B), ipTM(B \rightarrow A)]$$

**ipSAE** - усовершенствованная версия ipTM

$$ipSAE(A \rightarrow B) = \max_{i \in A} \left[ \text{mean}_{\substack{j \in B \\ PAE_{ij} < cutoff}} \left( \frac{1}{1 + \left( \frac{PAE_{ij}}{d_0} \right)^2} \right) \right] \quad PAE_{ij} = \sum_{b=1}^{64} p_{ij}^b \Delta_b$$

$$ipSAE(A, B) = \max[ipSAE(A \rightarrow B), ipSAE(B \rightarrow A)]$$

**pDockQ** - аппроксимированная версия DockQ

$$pDockQ = \frac{L}{1 + e^{-k(x-x_0)}} + b$$

**Среднее качество:**

$$pDockQ(A, B) \geq 0.23$$

$$ipTM(A, B) \geq 0.6$$

$$ipSAE(A, B) \geq 0.3$$

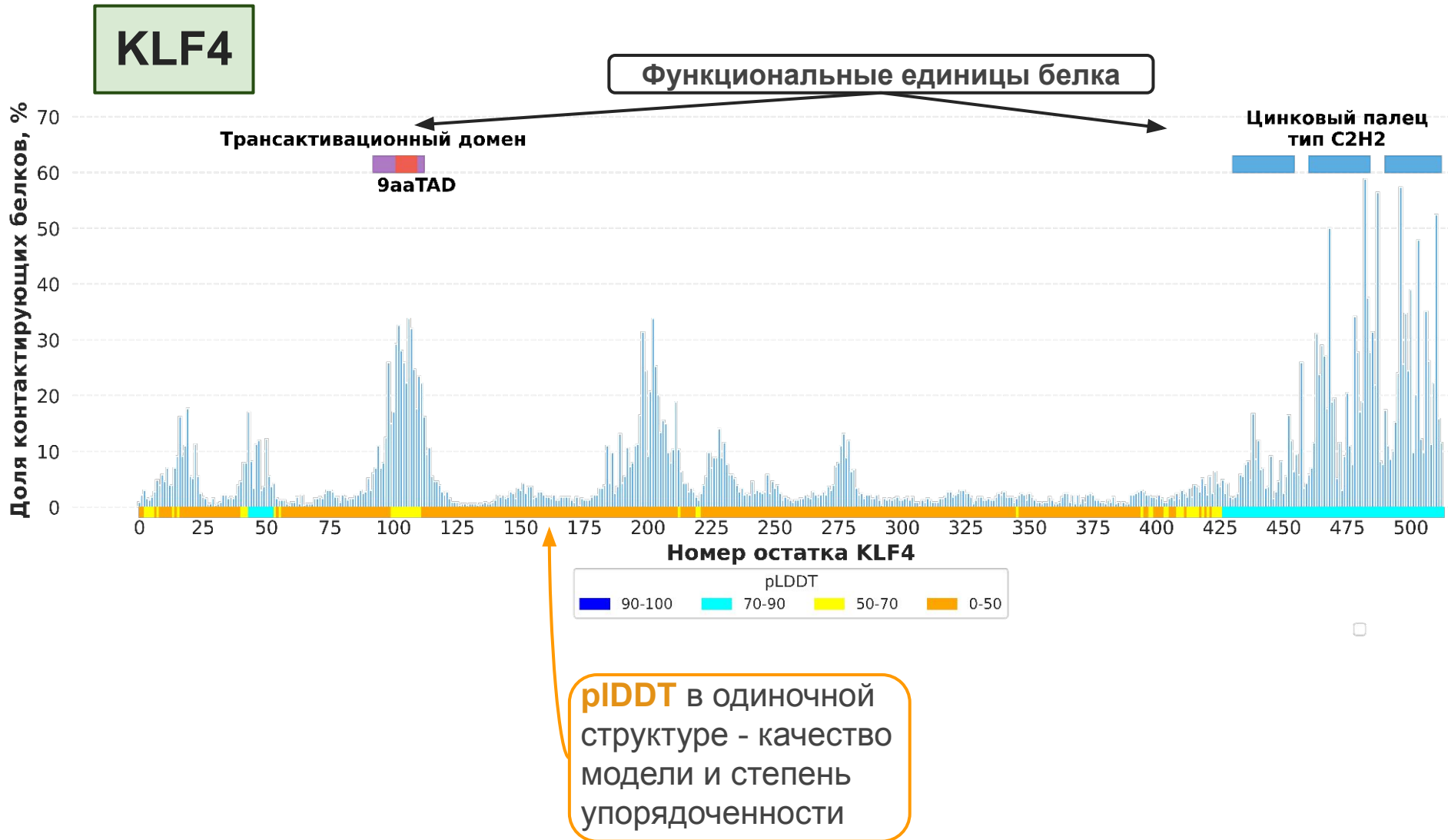
**Высокое качество:**

$$pDockQ(A, B) \geq 0.5$$

$$ipTM(A, B) \geq 0.8$$

$$ipSAE(A, B) \geq 0.5$$

# Результаты. Структурный анализ белок-белковых контактов





# Результаты. Структурный анализ белок-белковых контактов

**KLF4**

Функциональные единицы белка

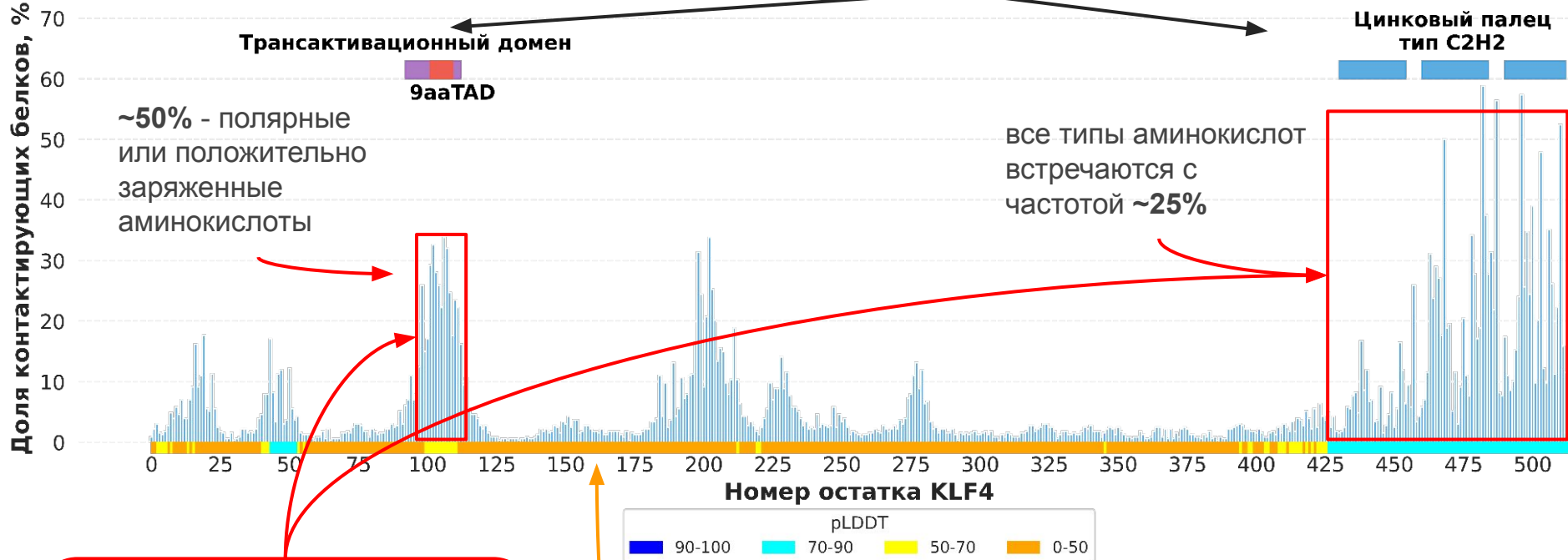
Трансактивационный домен

9aaTAD

Цинковый палец  
тип C2H2

~50% - полярные  
или положительно  
заряженные  
аминокислоты

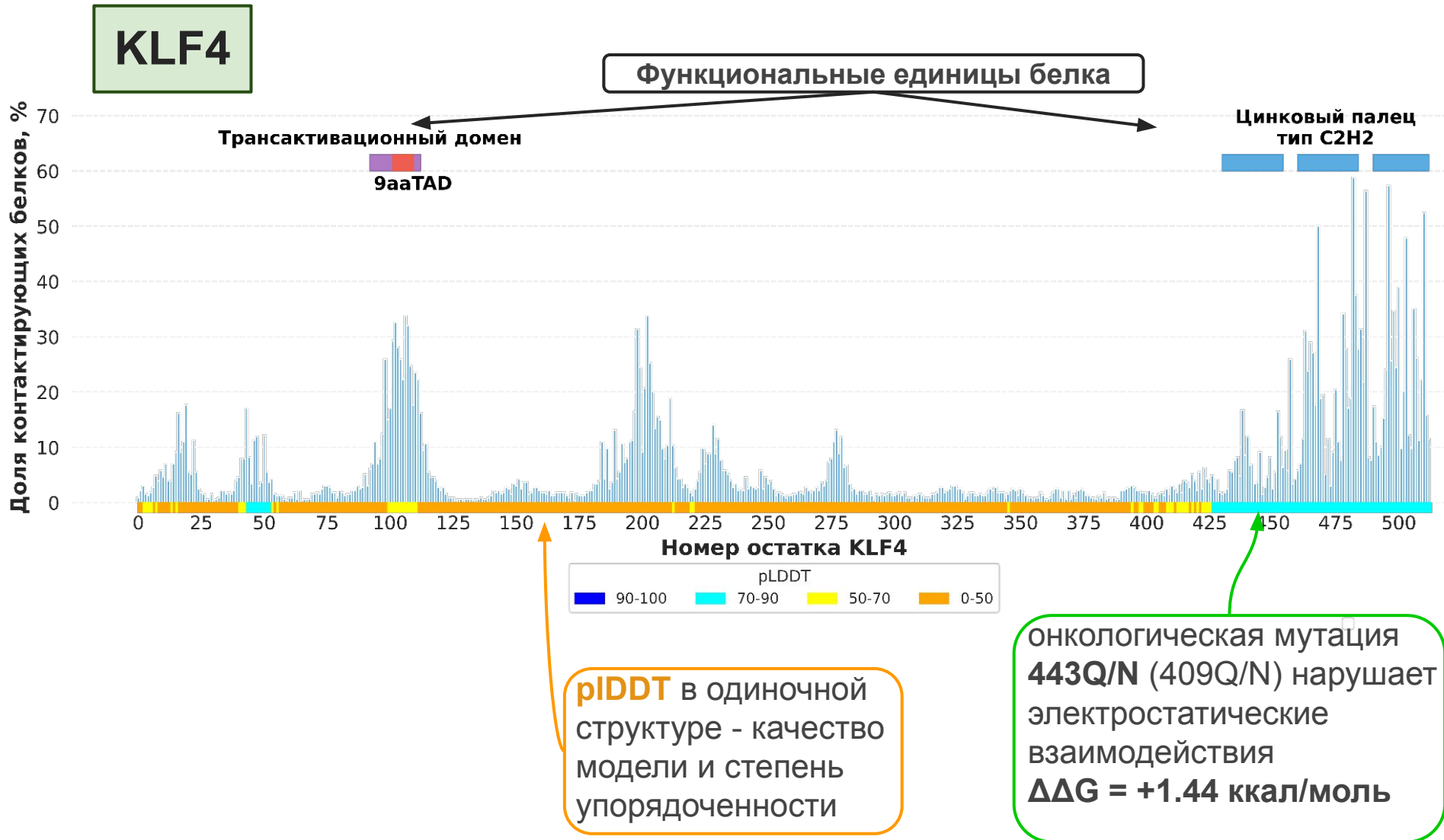
все типы аминокислот  
встречаются с  
частотой ~25%



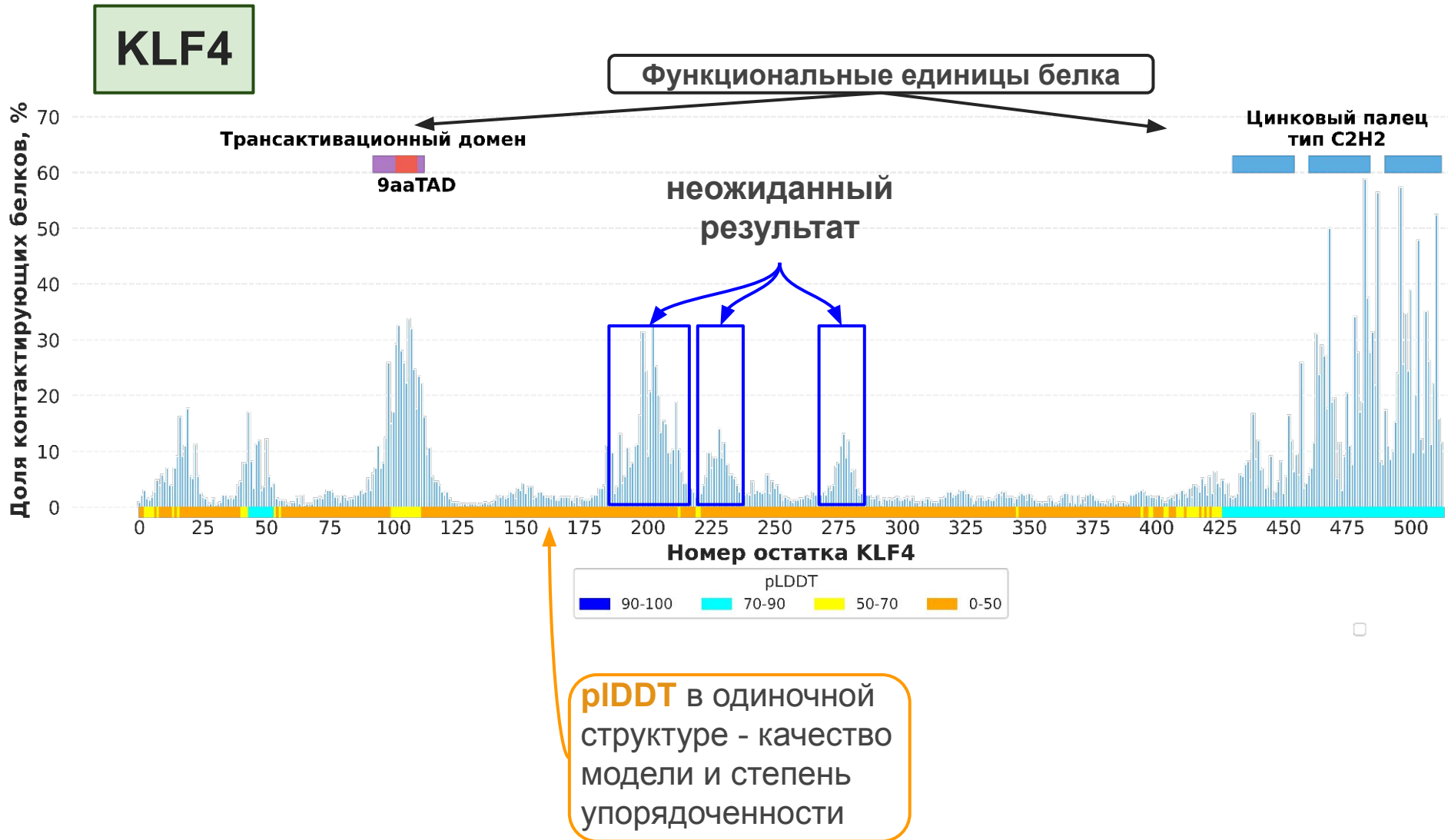
большая часть контактов  
приходится на  
структурированные участки  
белка, соответствующие  
функциональным единицам

**pLDDT** в одиночной  
структуре - качество  
модели и степень  
упорядоченности

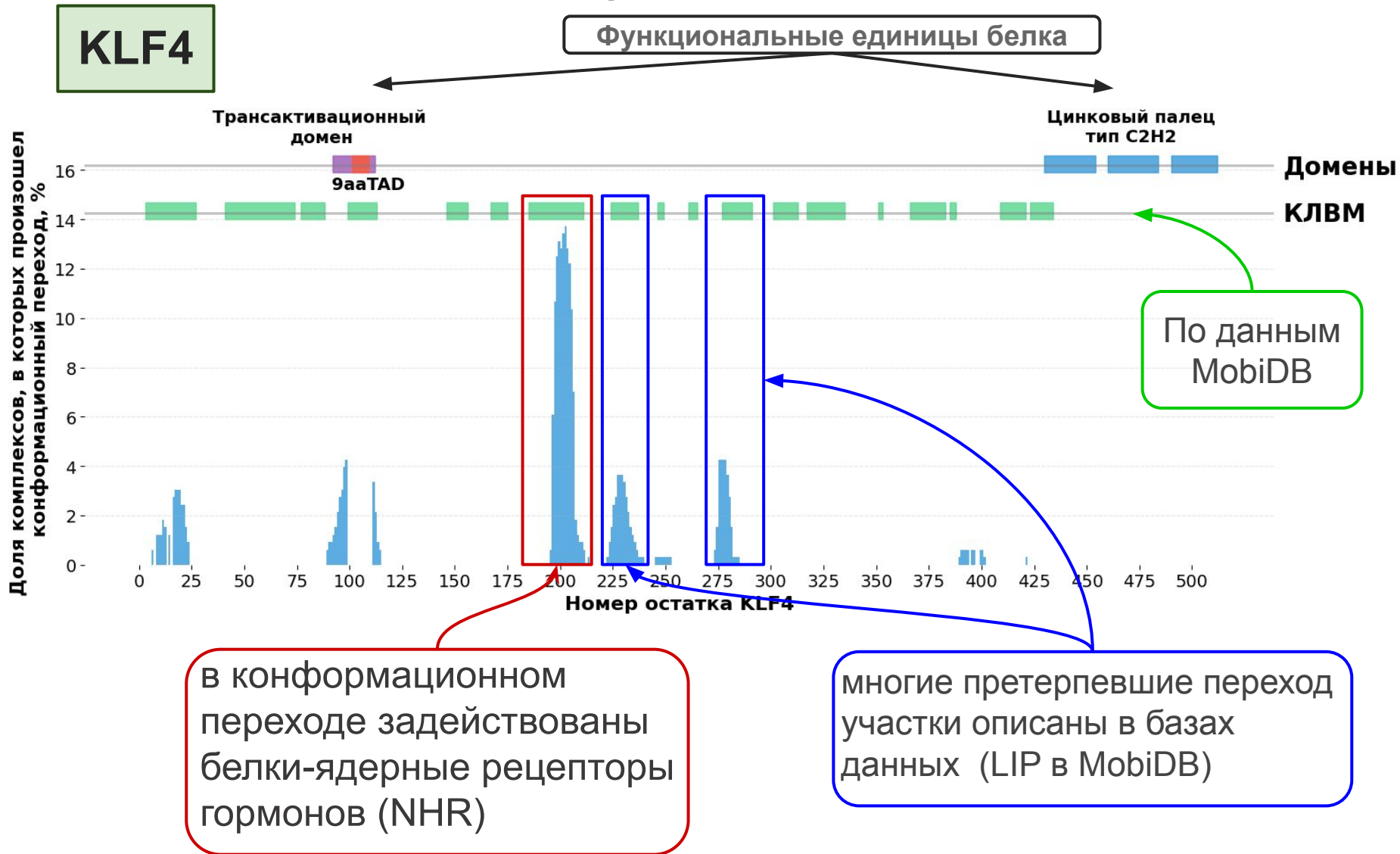
# Результаты. Структурный анализ белок-белковых контактов



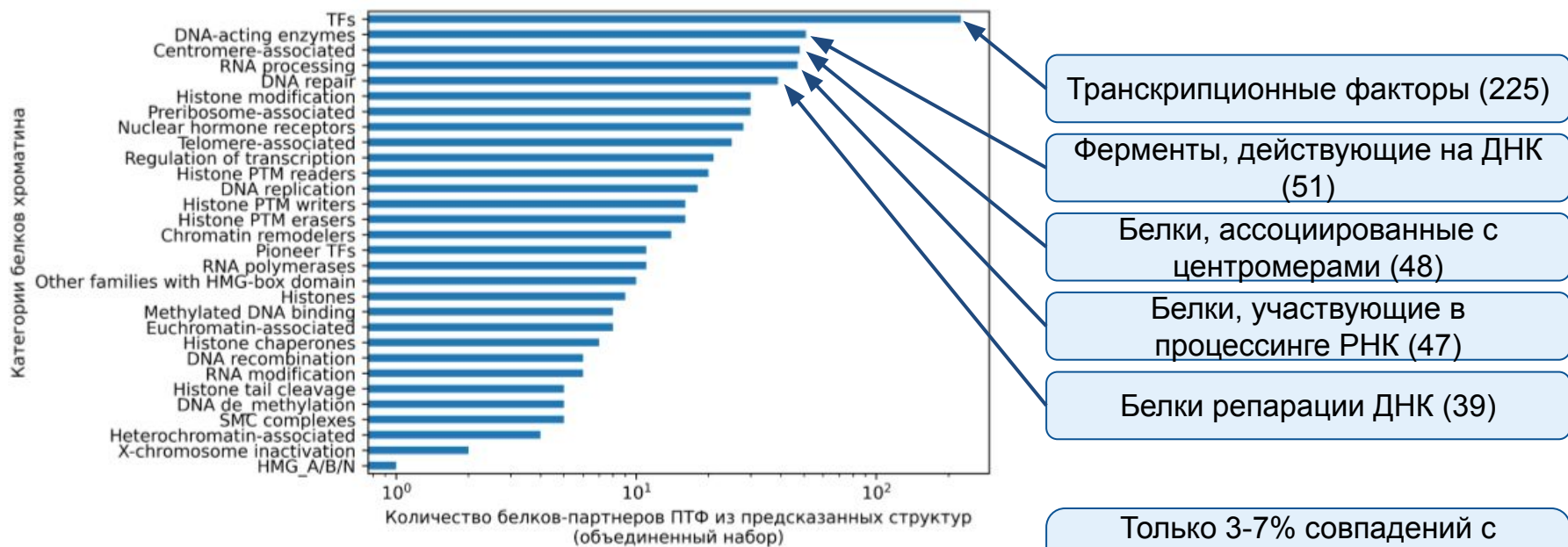
# Результаты. Структурный анализ белок-белковых контактов



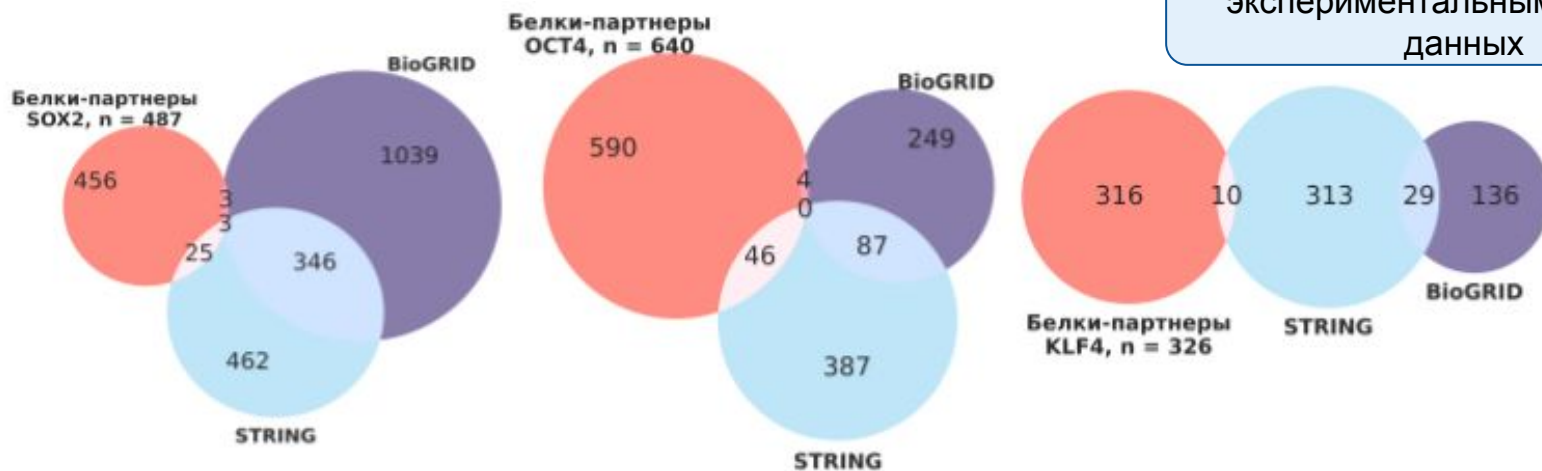
# Результаты. Анализ конформационных переходов



# Результаты. Функциональная принадлежность белков-партнеров. Сравнение с базами данных



Только 3-7% совпадений с экспериментальными базами данных



# Выводы

1. Методами генеративного ИИ было предсказано 10 632 комплекса белков ПТФ (SOX2, OCT4, KLF4) человека с другими белками ядра. Из числа предсказанных структур 14% имеют высокое качество, большая часть полученных взаимодействий ранее не была описана в литературе.
2. Были выявлены ключевые участки ПТФ в белок-белковых комплексах для которых характерен переход в упорядоченное состояние при связывании. Подтверждены специфические взаимодействия KLF4 с ядерными рецепторами гормонов. Предложен механизм действия онкологической мутации K409Q/N в KLF4: показано, что она слабо дестабилизирует белок-белковые комплексы.
3. Выявлено, что среди белков-партнеров ПТФ преимущественно представлены транскрипционные факторы, гистон-модифицирующие ферменты, белки репарации ДНК и ядерные рецепторы.

# Благодарности

Выражаю благодарность

- **Научным руководителям** Шайтану Алексею Константиновичу, Грибковой Анне Кирилловне, Армееву Григорию Алексеевичу
- **Факультету фундаментальной физико-химической инженерии МГУ имени М.В. Ломоносова**
- **моей семье**

за поддержку на всех этапах работы.