

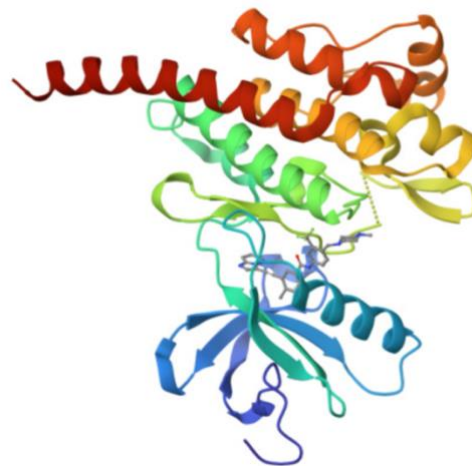
Лабораторная работа 3

1. Название белка и его UniProt ID.

Название белка: Proto-oncogene tyrosine-protein kinase ABL1 (ABL1)

UniProt ID: P00520 (ABL1_MOUSE)

Организм: Mus musculus (мышь)



 **3IK3** | **pdb_00003ik3** 

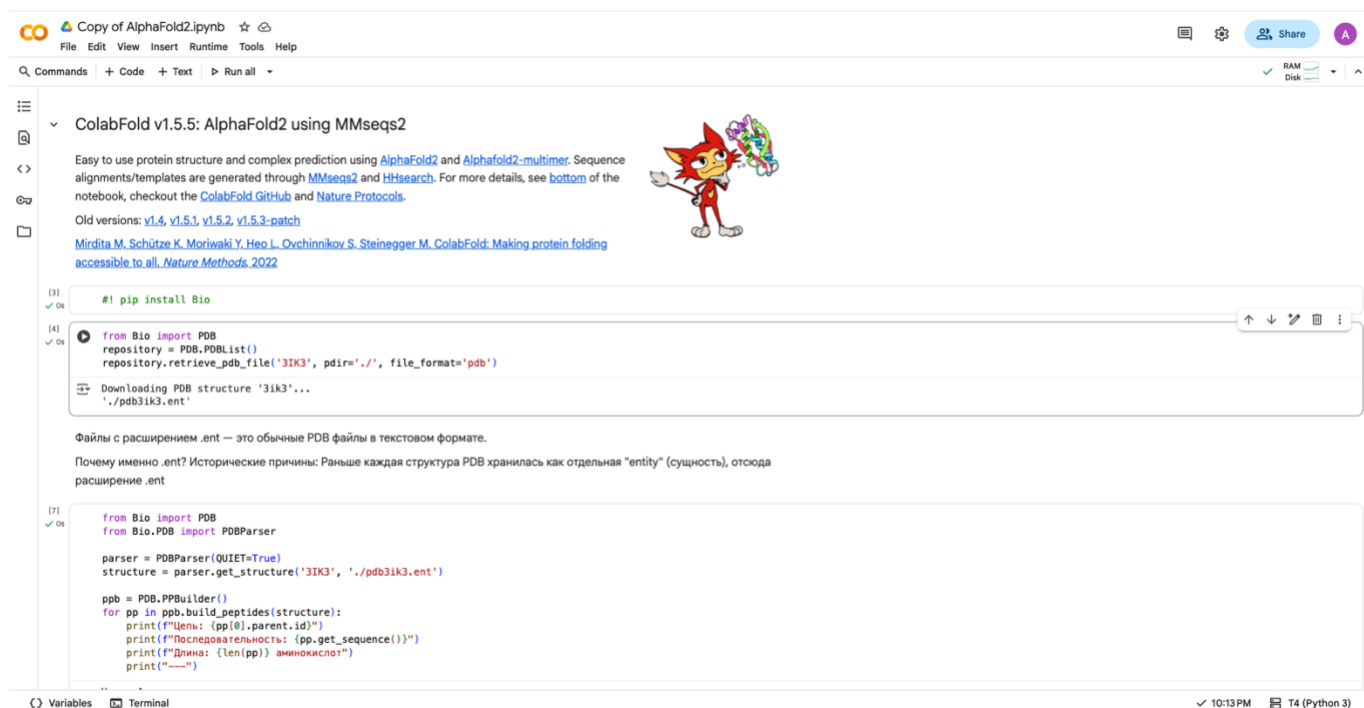
AP24534, a Pan-BCR-ABL Inhibitor for Chronic Myeloid Leukemia, Potently Inhibits the T315I Mutant and Overcomes Mutation-Based Resistance

CHAIN

B

Proto-oncogene tyrosine-protein kinase ABL1 - Mus musculus  [Explore Sequence Annotations in 3D](#)

2. Скриншоты интерфейса ColabFold и AlphaFold Server.



```
[3] ✓ Os #! pip install Bio

[4] ✓ Os from Bio import PDB
repository = PDB.PDBList()
repository.retrieve_pdb_file('3IK3', pdir='.', file_format='pdb')

Downloading PDB structure '3ik3'...
'./pdb3ik3.ent'

Файлы с расширением .ent — это обычные PDB файлы в текстовом формате.
Почему именно .ent? Исторические причины: Раньше каждая структура PDB хранилась как отдельная "entity" (сущность), отсюда
расширение .ent

[7] ✓ Os from Bio import PDB
from Bio.PDB import PDBParser

parser = PDBParser(QUiet=True)
structure = parser.get_structure('3IK3', './pdb3ik3.ent')

ppb = PDB.PPBuilder()
for pp in ppb.build_peptides(structure):
    print(f"Цепь: {pp[0].parent.id}")
    print(f"Последовательность: {pp.get_sequence()}")
    print(f"Длина: {len(pp)} аминокислот")
    print(f"-----")
```

Explore these examples of structures to see it in action – try them out without using your quota until you begin editing!

🔗 Protein-RNA-lon: PDB 8AW3 🔗 Protein-Glycan-lon: PDB 7BBV 🔗 Protein-DNA-lon: PDB 7RCE

Ok, got it

📄 Upload JSON

🗑️ Clear

Entity type

Protein

Copies

1

10

20

30

40

50

60

GSPNYDKWEM

ERTDITMKHK

LGGGQYGEVY

EGVKKYSLT

VAVKTLKEDT

MEVEEFLKEA

70

80

90

100

110

120

AVMKEIKHPN

LVQLLGVCTR

EPPFYIIIEF

MTYGNLLDYL

RECNRQEVSA

VVLLYMATQI

130

140

150

158

SSAMEYLEKK

NFIHRDLAAR

NCLVGENHLV

KVADFGLS

...

↗

+ Add entity

📄 Save job

Continue and preview job

🔍 Search History

✓ Completed ✓ Saved draft ✓ In progress ✓ Examples ✓ Failed

No jobs found with filters specified

Items per page: 10 0 of 0 |< < > >|

Job name*

3IK3

Seed

42

Seed ☒

Specify a seed

Type	Copies	Sequence
Protein	1	GSPNYDKWEMERTDITMKHKLGGGQYGE VY... (length 158)

Remaining jobs: 30

Go back and edit this job

Confirm and submit job

🔍 Search History

✓ Completed ✓ Saved draft ✓ In progress ✓ Examples ✓ Failed

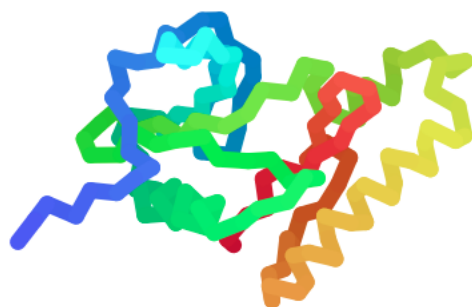
<input type="checkbox"/>	Name	Modified	
<input type="checkbox"/>	3IK3	2025-10-24 21:55	⋮

Items per page: 10 1 – 1 of 1 |< < > >|

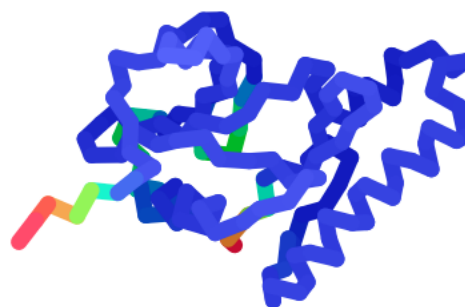
Search History		
<input checked="" type="checkbox"/>	Completed	
<input checked="" type="checkbox"/>	Saved draft	
<input checked="" type="checkbox"/>	In progress	
<input checked="" type="checkbox"/>	Examples	
<input checked="" type="checkbox"/>	Failed	
<input type="checkbox"/>	Name	Modified
<input type="checkbox"/> <input checked="" type="checkbox"/>	3IK3	2025-10-24 21:57
Items per page: 10		1 - 1 of 1

3. Полученные значения pLDDT и confidence score.

colored by N→C



colored by pLDDT



```

2025-10-24 19:13:23,727 reranking models by 'plddt' metric
2025-10-24 19:13:23,727 rank_001_alphafold2_ptm_model_5_seed_000 pLDDT=91.8 pTM=0.867
2025-10-24 19:13:23,728 rank_002_alphafold2_ptm_model_3_seed_000 pLDDT=91 pTM=0.857
2025-10-24 19:13:23,728 rank_003_alphafold2_ptm_model_2_seed_000 pLDDT=91 pTM=0.856
2025-10-24 19:13:23,728 rank_004_alphafold2_ptm_model_1_seed_000 pLDDT=90.8 pTM=0.849
2025-10-24 19:13:23,728 rank_005_alphafold2_ptm_model_4_seed_000 pLDDT=90.7 pTM=0.858
2025-10-24 19:13:25,012 Done

```

pLDDT = 91.8

pTM = 0.867

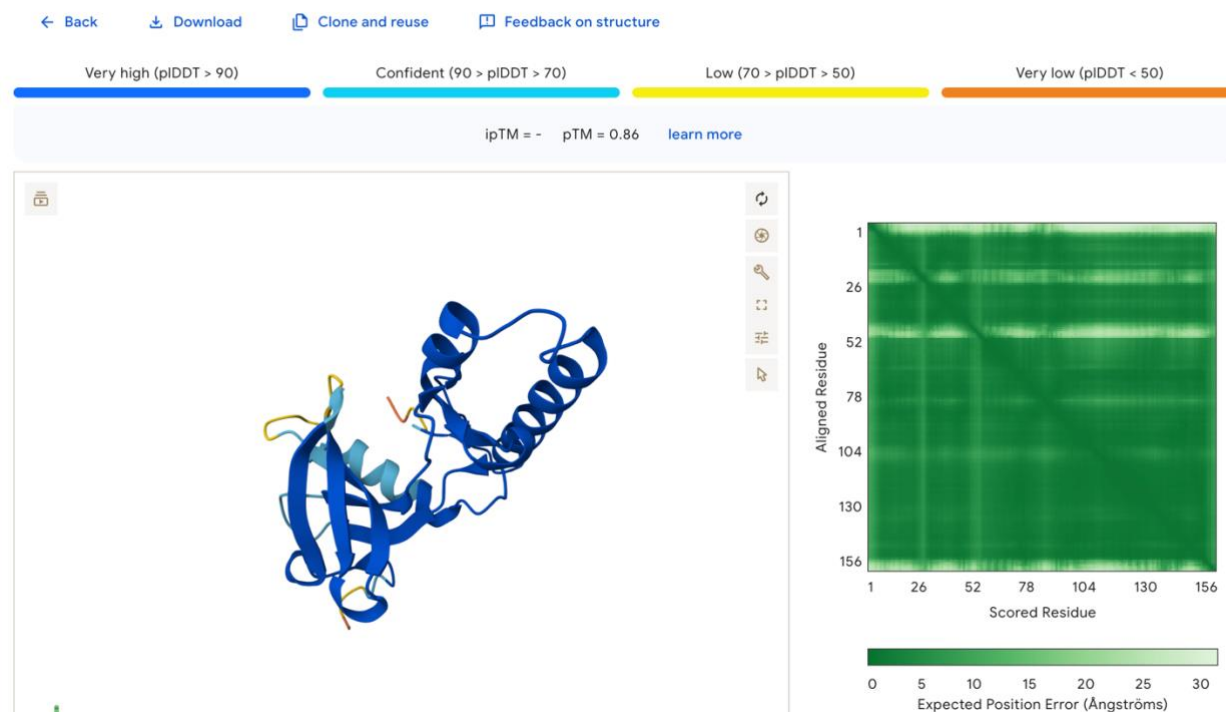
pTM (predicted Template Modeling score) - Confidence Score

0.867 - означает, что модель с 86.7% уверенностью правильно предсказала общую укладку белка

pLDDT (predicted Local Distance Difference Test)

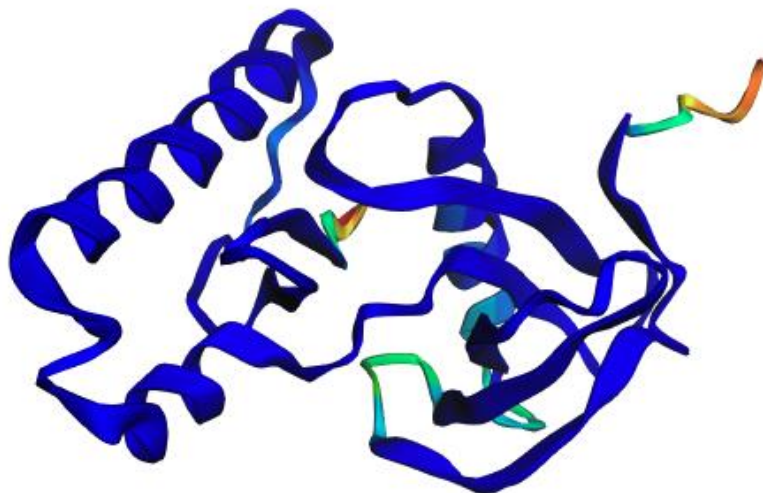
90–100 (Очень высокая) - Надёжная структура

3IK3



pTM = 0.86

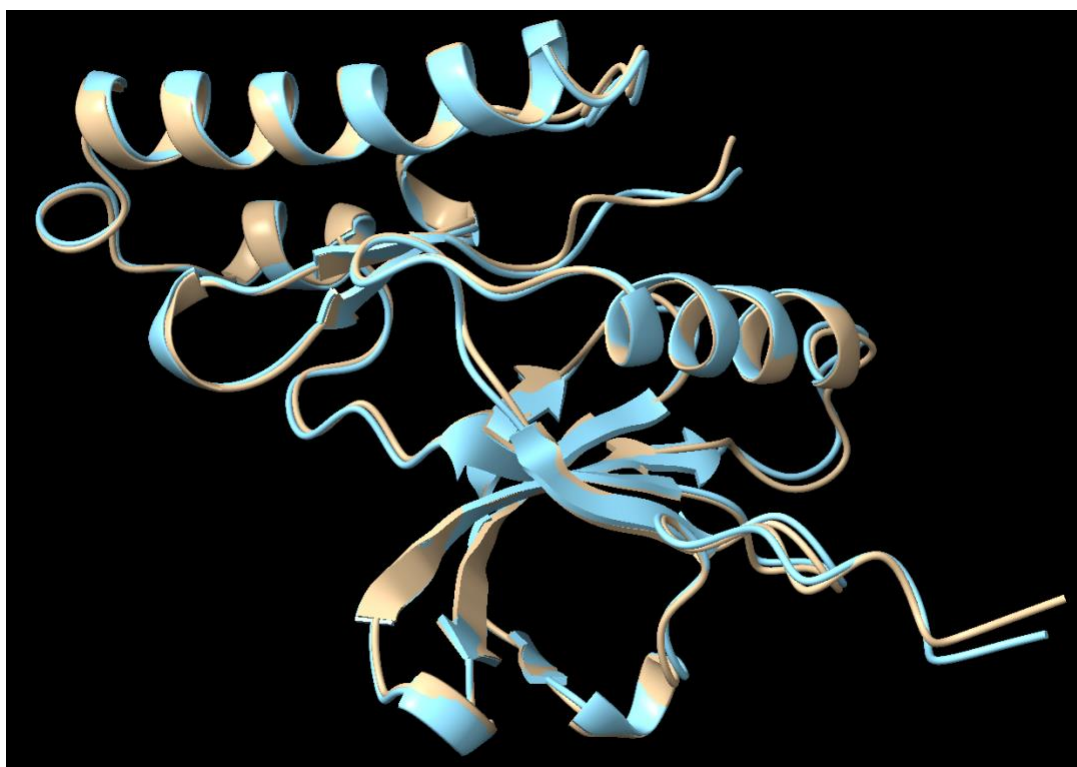
4. Скриншоты 3D-структур.
AlphaFold Colab:



AlphaFold Server:



5. Значение RMSD между моделями.



[matchmaker](#) #1 to #2

Parameters																	
Chain pairing	bb																
Alignment algorithm	Needleman-Wunsch																
Similarity matrix	BLOSUM-62																
SS fraction	0.3																
Gap open (HH/SS/other)	18/18/6																
Gap extend	1																
SS matrix	<table><tr><td></td><td>H</td><td>S</td><td>O</td></tr><tr><td>H</td><td>6</td><td>-9</td><td>-6</td></tr><tr><td>S</td><td></td><td>6</td><td>-6</td></tr><tr><td>O</td><td></td><td></td><td>4</td></tr></table>		H	S	O	H	6	-9	-6	S		6	-6	O			4
	H	S	O														
H	6	-9	-6														
S		6	-6														
O			4														
Iteration cutoff	2																

Matchmaker fold_3ik3_model_0.cif, chain A (#2) with
3IK3_1561c_unrelaxed_rank_001_alphafold2_ptm_model_5_seed_000.pdb, chain A
(#1), sequence alignment score = 815.9
RMSD between 154 pruned atom pairs is 0.471 angstroms; (across all 158 pairs:
0.671)

Models					
Name	ID				
3IK3_1561c_unrelaxed_rank_001_alphafold2_ptm_model_5...	1		<input checked="" type="checkbox"/>	<input type="checkbox"/>	Close
fold_3ik3_model_0.cif	2		<input checked="" type="checkbox"/>	<input type="checkbox"/>	Hide
					Show
					View
					Info

6. Краткие выводы о различиях методов и ответы на контрольные вопросы

Результаты структурного выравнивания в UCSF Chimera:

- RMSD между предсказанной и экспериментальной структурой: 0.471 Å
- Sequence alignment score: 815.9

Выводы:

1. Предсказанная модель демонстрирует очень высокое структурное сходство с экспериментальной структурой (RMSD ((Root Mean Square Deviation) среднеквадратичное отклонение) < 0.5 Å)
2. Качество предсказания AlphaFold подтверждено сравнением с эталонной структурой из PDB
3. Незначительные расхождения (0.671 Å по всем атомам) могут

быть связаны с гибкими участками белка

Контрольные вопросы

1. Чем отличается ColabFold от оригинального AlphaFold Colab?

ColabFold — это оптимизированная версия AlphaFold2, адаптированная для работы в облаке Google Colab. Она использует ускоренные библиотеки и серверные ресурсы для быстрого моделирования без установки программ.

2. Что нового реализовано в AlphaFold3 по сравнению с AlphaFold2?

Общий охват молекул — AF3 может предсказывать не только отдельные белки, но и **комплексы**, включающие белки + нуклеиновые кислоты (DNA/RNA), лиганды, ионы и модифицированные остатки.

Архитектура с диффузионной моделью — в AF3 используется так называемый «diffusion network» («диффузионный» подход), где модель начинает с «облака» атомов и постепенно «схлопывает» его до финальной структуры.

Повышенная точность в задачах взаимодействия — AF3 показывает намного лучшую работу в задачах: белок-лиганды, белок-нуклеиновая кислота, антитела-антиген и др.

Упрощенная/унифицированная архитектура — несмотря на расширенную функциональность, архитектура модели по некоторым источникам стала проще и более унифицированной, чем у AF2 + дополнений.

3. Как интерпретировать pLDDT и confidence score?

pLDDT (Predicted Local Distance Difference Test) — это локальная метрика уверенности AlphaFold в точности предсказанной структуры каждой аминокислоты.

Значения варьируются от 0 до 100:

- **90–100** — очень высокая уверенность, структура достоверна;
- **70–90** — хорошая уверенность, форма белка предсказана корректно;
- **50–70** — средняя уверенность, возможна гибкость участка;
- **<50** — низкая уверенность, участок вероятно неструктурирован.

Для удобства визуализации участки модели окрашиваются по шкале:

синий — высокая уверенность, голубой — хорошая, жёлтый — средняя, красный — низкая.

Confidence score (или глобальный показатель уверенности) представляет собой среднее значение pLDDT по всем остаткам белка и отражает общую надёжность предсказанной структуры.

Если средний pLDDT превышает 90, модель считается высокоточной; значения ниже 70 указывают на наличие подвижных или неопределённых областей.

4. Какие преимущества и недостатки имеют ColabFold и AlphaFold Server?

ColabFold представляет собой оптимизированную и открытую версию AlphaFold2, которая обеспечивает значительно более быструю работу за счёт использования сервиса MMseqs2 для поиска гомологов. Этот инструмент удобен для обучения и исследовательских задач, так как позволяет запускать предсказания локально или через Google Colab и гибко настраивать параметры.

В то же время **AlphaFold Server**, основанный на AlphaFold3, обеспечивает более высокую точность и способен моделировать не только белки, но и их взаимодействия с ДНК, РНК и лигандами. Однако он работает медленнее, имеет ограниченный доступ и не предоставляет возможности тонкой настройки параметров пользователем.

5. Почему RMSD может отличаться между моделями одной и той же последовательности?

RMSD (Root Mean Square Deviation) между моделями одной и той же аминокислотной последовательности может отличаться из-за различий в способах предсказания пространственной структуры. Даже при одинаковой последовательности разные модели (или даже разные запуски одной модели) могут по-разному оценивать взаимодействия между аминокислотами, вторичные структуры и конформационные состояния.

Кроме того, отличие RMSD может быть связано с различиями в исходных данных для выравнивания (MSA), случайной инициализацией модели, качеством шаблонов, параметрами релаксации или вариациями в методах оптимизации. Таким образом, RMSD отражает не только точность, но и неопределённость модели в предсказании конкретной пространственной конфигурации белка.