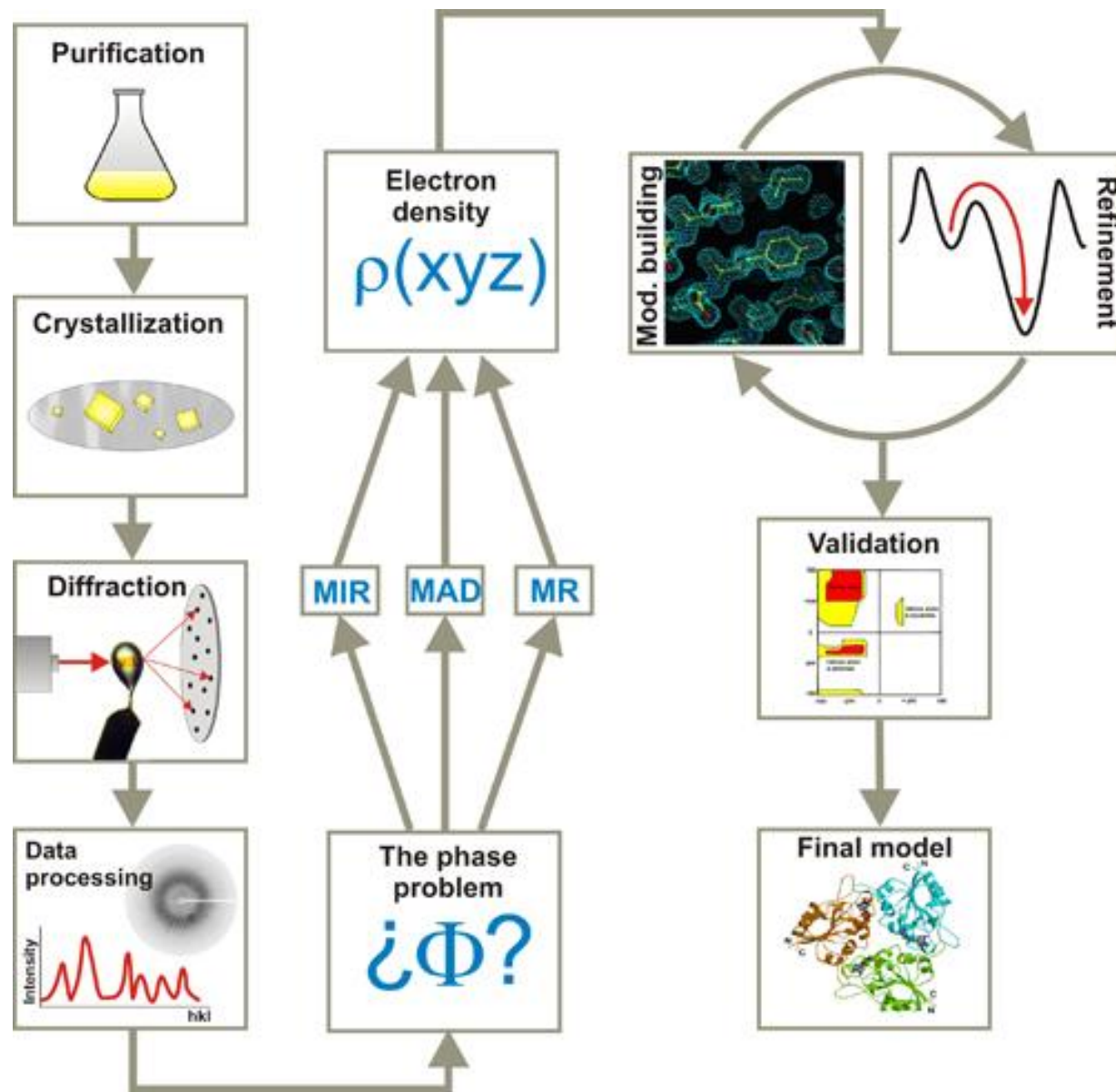

Лекция № 14
Рентгеноструктурный анализ биомакромолекул.
Представление структуры белков и НК во
внутренних координатах.
Карты Рамачандрана.

— Армеев Григорий Алексеевич —
к.ф.-м.н., м.н.с. каф. биоинженерии
armeev@intbio.org

Сайт курса <http://intbio.org/bioinf2018-2019>

Цель лекции.

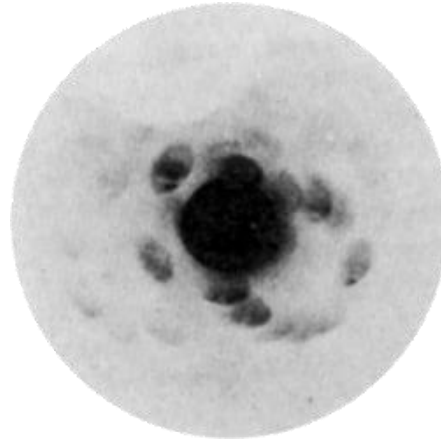
В общих чертах
описать эту
картинку:



Рентгеноструктурный анализ. Вместо введения.



Max von Laue



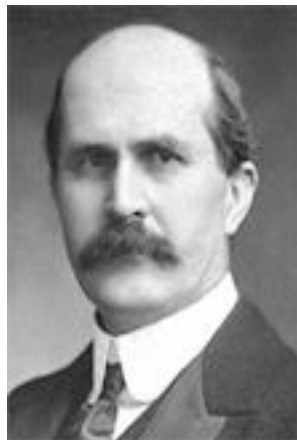
William H. Bragg

William L. Bragg

Рентгеноструктурный анализ. Вместо введения.



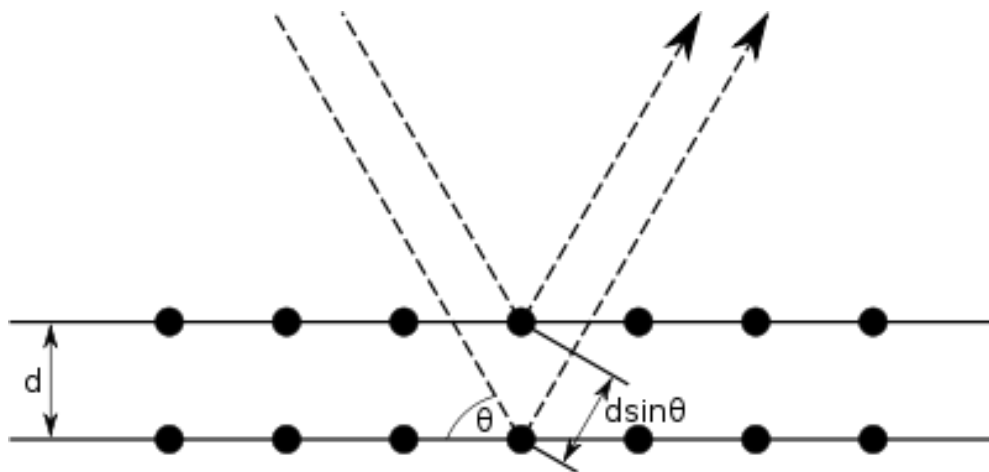
Max von Laue



William H. Bragg



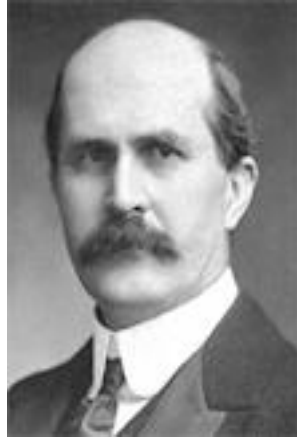
William L. Bragg



Рентгеноструктурный анализ. Вместо введения.



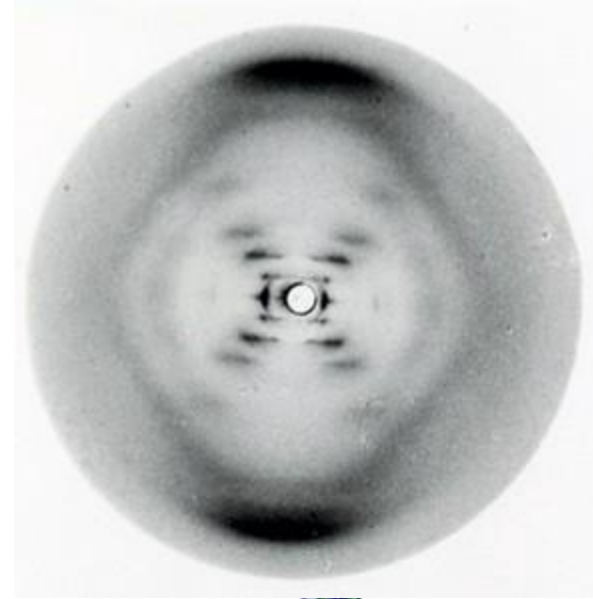
Max von Laue



William H. Bragg



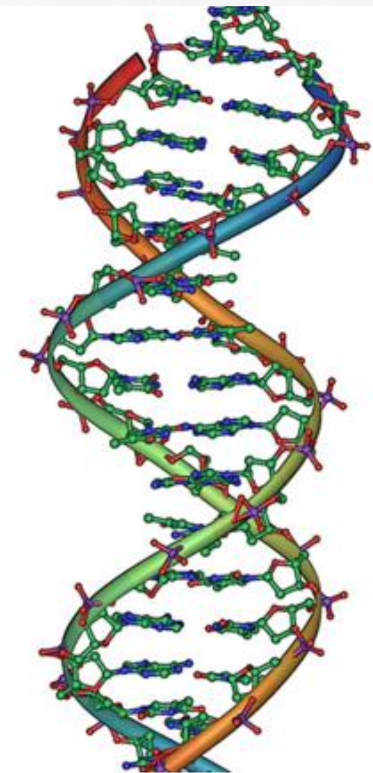
William L. Bragg



Raymond Gosling



Rosalind Elsie Franklin



Рентгеноструктурный анализ. Вместо введения.



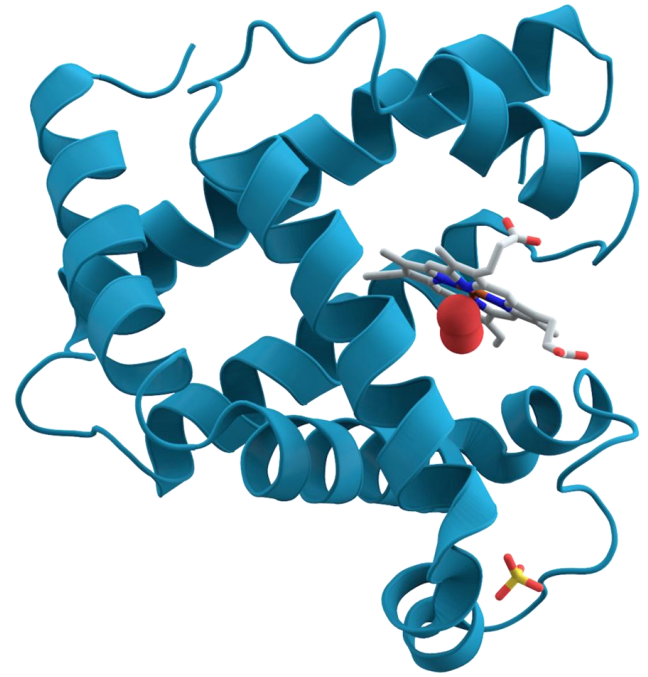
Max von Laue



William H. Bragg



William L. Bragg



Raymond Gosling

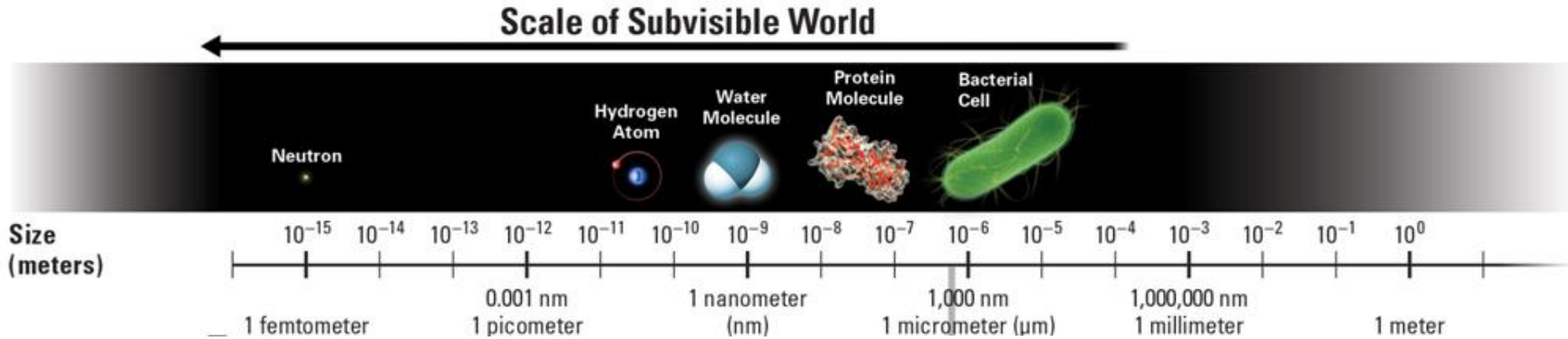


Rosalind Elsie Franklin



John Kendrew

Рентгеноструктурный анализ. Масштабы микромира.



Химическая связь 1 Å (C-C 1.5 Å)

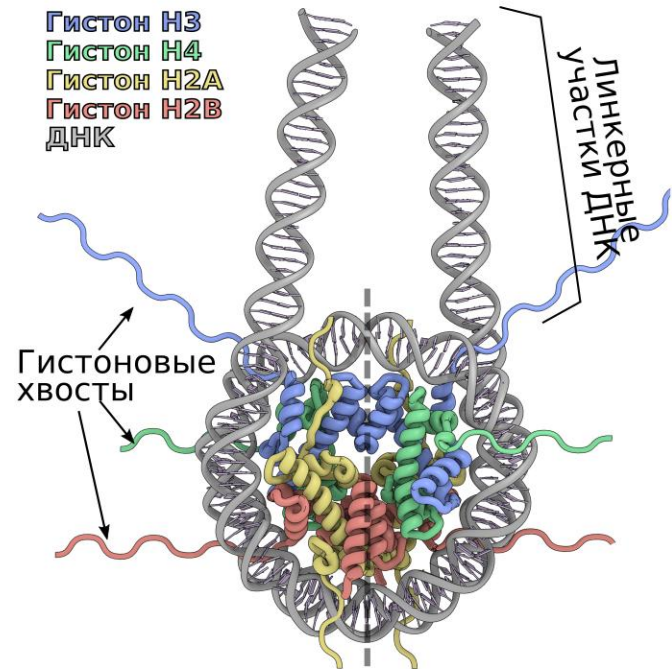
Белковый домен ~50 Å

Нуклеосома ~100 Å

Рибосома ~250 Å

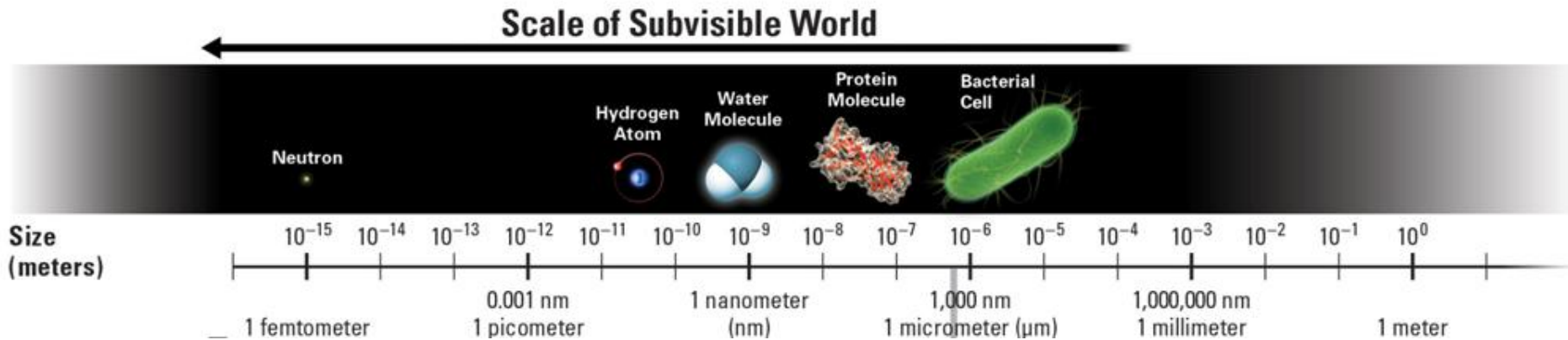
Вирус ~700 Å

*1 Å = 0,1 нм = 10⁻¹⁰м



Диадная ось симметрии

Рентгеноструктурный анализ. Масштабы микромира.



Химическая связь 1 Å (C-C 1.5 Å)

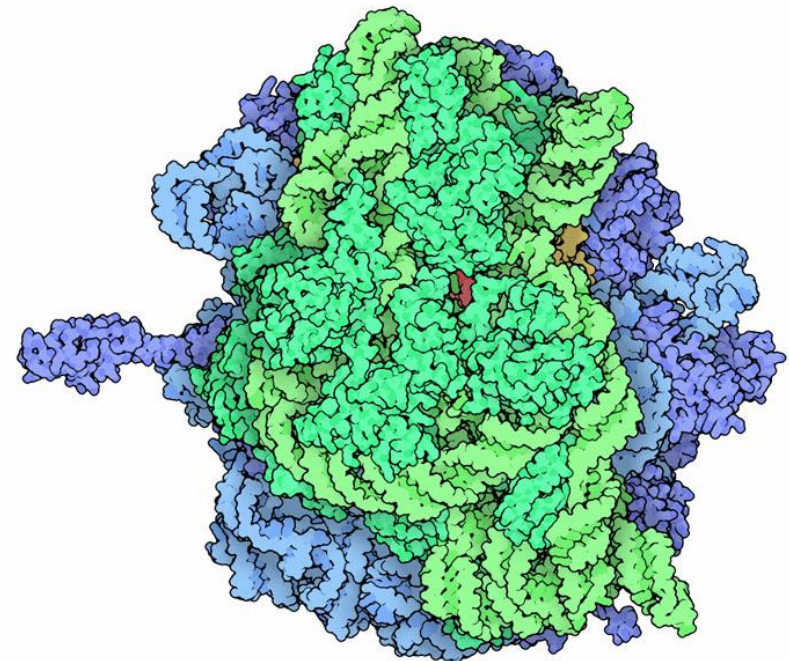
Белковый домен ~50 Å

Нуклеосома ~100 Å

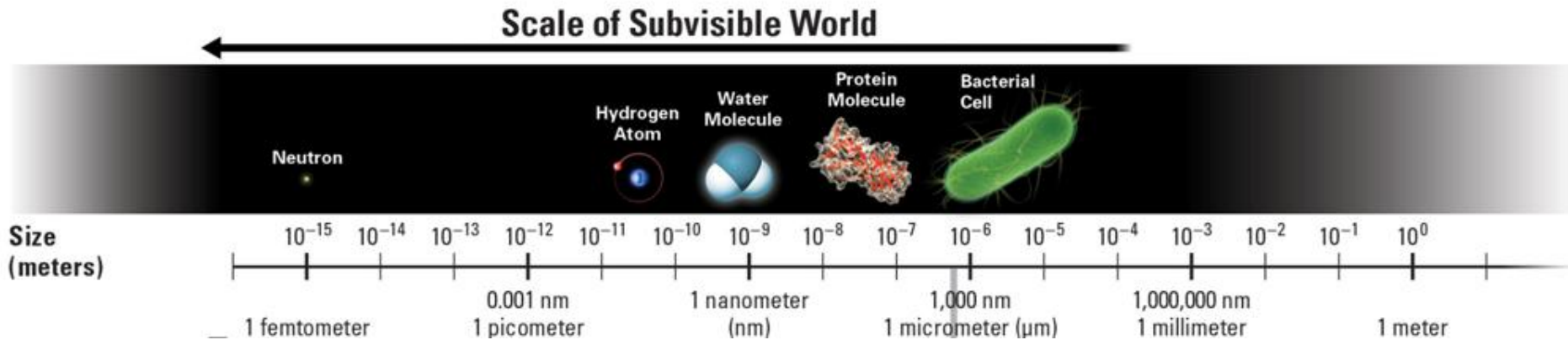
Рибосома ~250 Å

Вирус ~700 Å

*1 Å = 0,1 нм = 10⁻¹⁰м



Рентгеноструктурный анализ. Масштабы микромира.



Химическая связь 1 Å (C-C 1.5 Å)

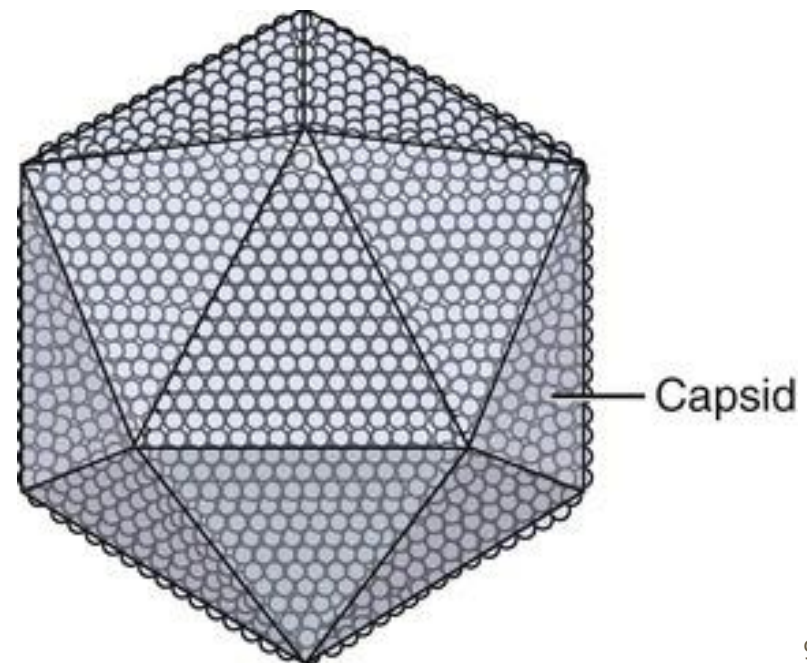
Белковый домен ~50 Å

Нуклеосома ~100 Å

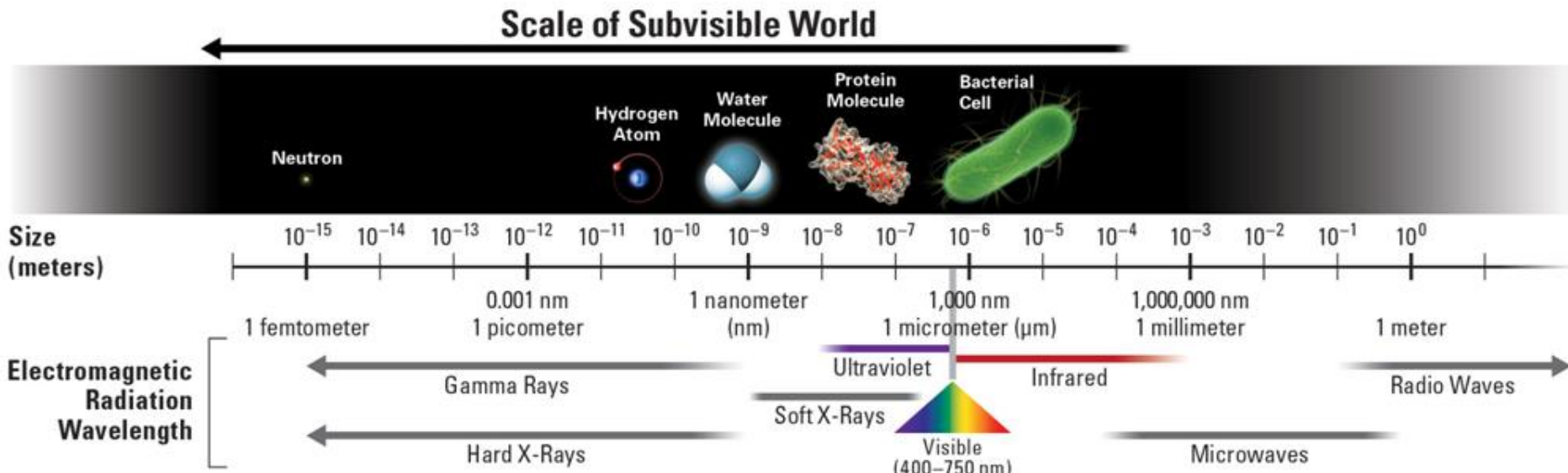
Рибосома ~250 Å

Вирус ~700 Å

*1 Å = 0,1 нм = 10^{-10} м



Рентгеноструктурный анализ. Почему рентгеновские лучи?



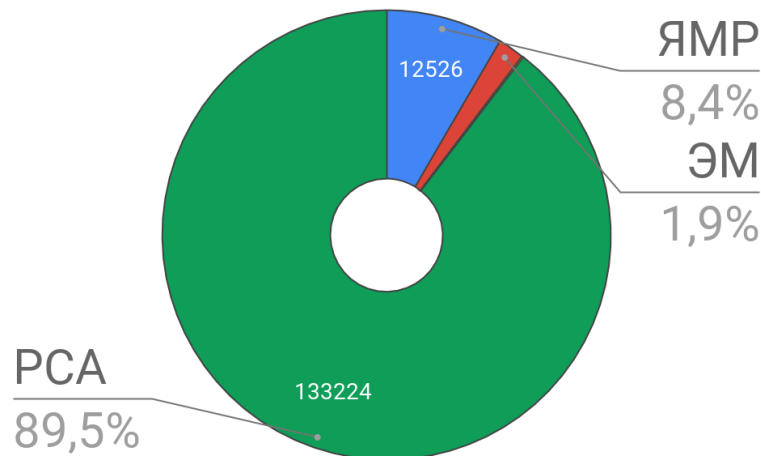
Распределение числа определенных структур по методам

Видимый свет $\lambda = 200 - 800 \text{ nm}$

Рентгеновские лучи $\lambda = 0.6 - 3 \text{ \AA}$

Нейтронное излучение $\lambda = 2 - 3 \text{ \AA}$

Пучок электронов $\lambda = 0.04 \text{ \AA}$ (50 кэВ)

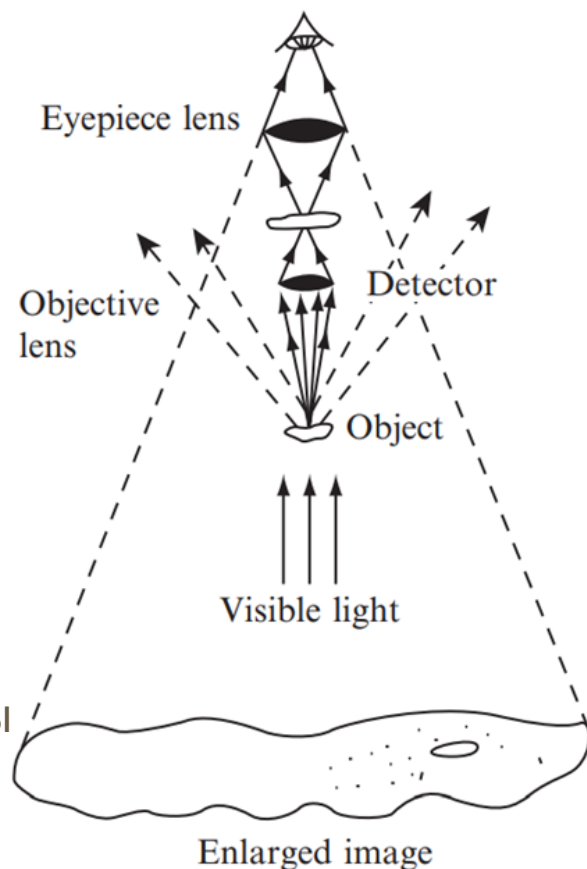


Рентгеноструктурный анализ. Нам нужен микроскоп!

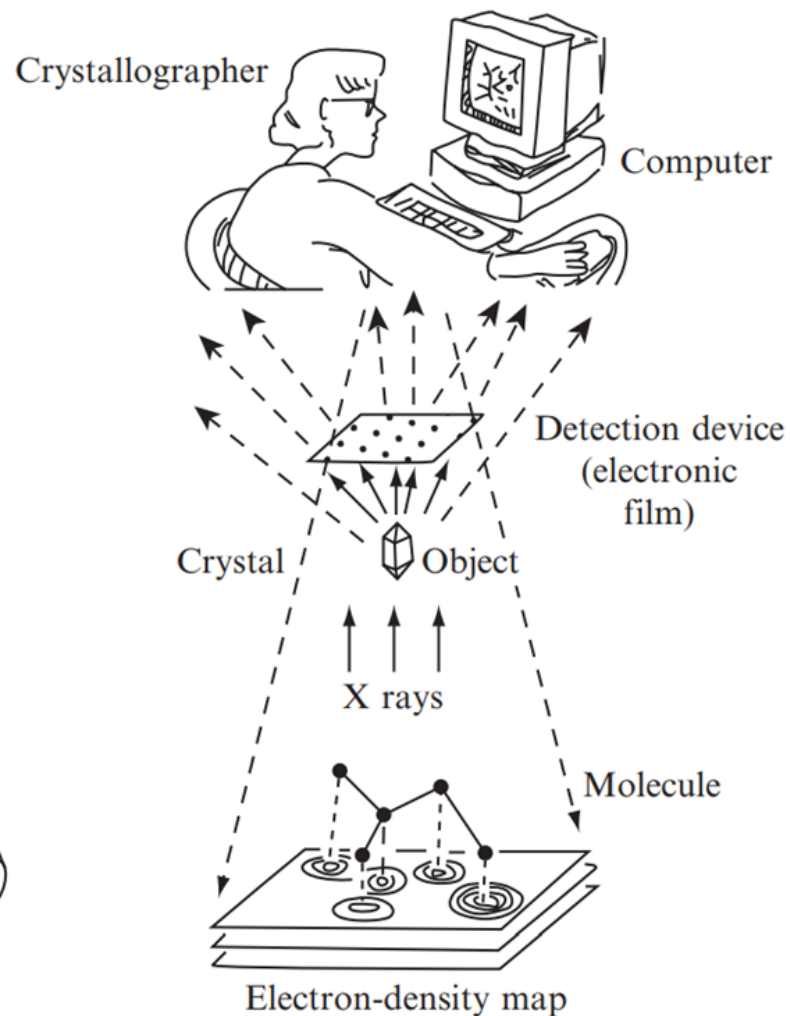
Микроскопия подразумевает наличие фокусирующих линз.

Рентгеновские лучи **сложно сфокусировать!**

В РСА в роли линзы выступает исследователь.



MICROSCOPE



X-RAY DIFFRACTION

J. P. Glusker and K. N. Trueblood. Crystal Structure Analysis: A Primer. 2010

Рентгеноструктурный анализ.

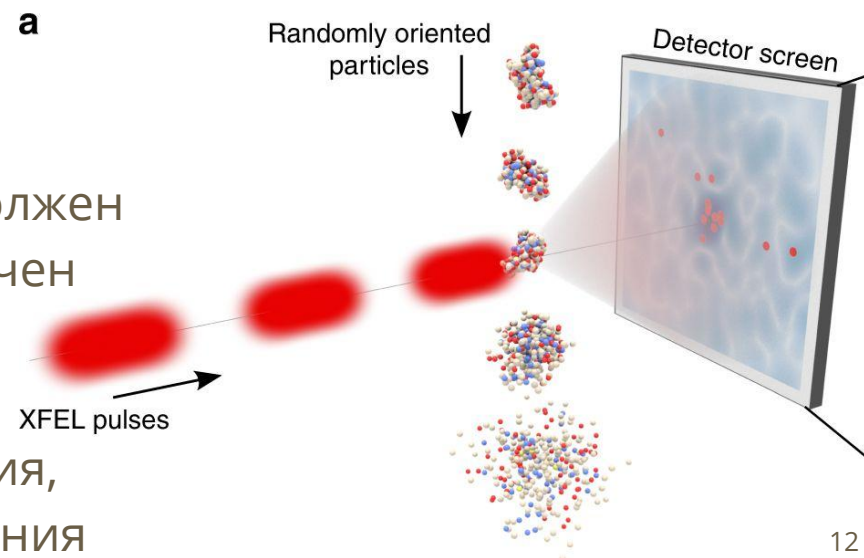
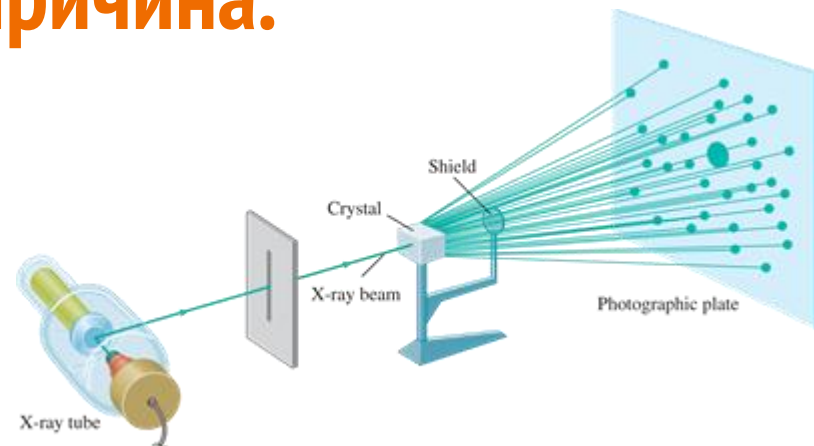
Зачем нужны кристаллы? Причина.

Исследование единичных молекул при помощи Рентгеновского излучения затруднено. **Слишком мало фотонов взаимодействует с молекулой.**

В Крио ЭМ - взаимодействие достаточно сильно, но нужен лед, да и электроны заряжены.

Что делать?

1. Нужно больше материала, но он должен быть упорядочен, причем упорядочен строго определенным образом.
(кристалл)
2. Увеличить интенсивность излучения, чтобы увеличить амплитуду рассеяния



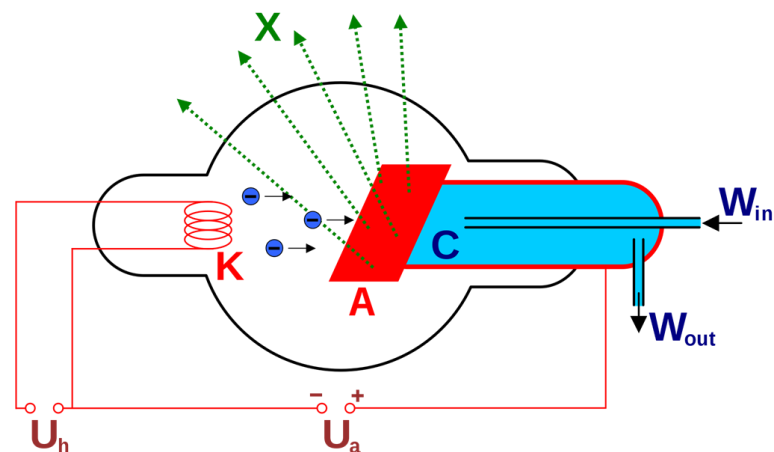
Рентгеноструктурный анализ.

Источники рентгеновского излучения.

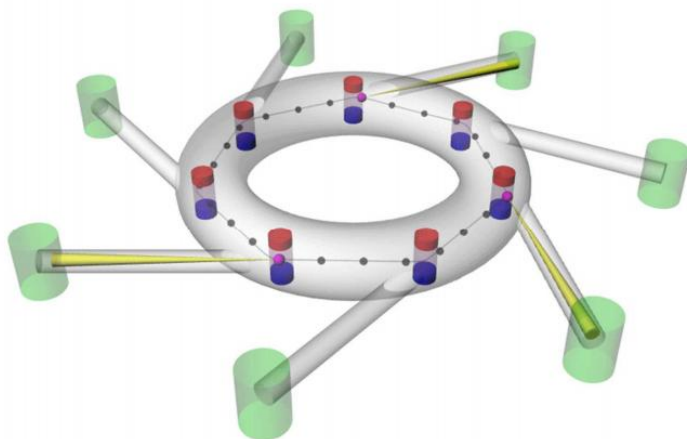
Рентгеновские лучи возникают в моменты, когда электроны ускоряются при (но не только):

1. Торможении при соударении с веществом
2. Изменении направления в магнитном поле

Рентгеновская трубка



Synchrotron radiation (dipoles)

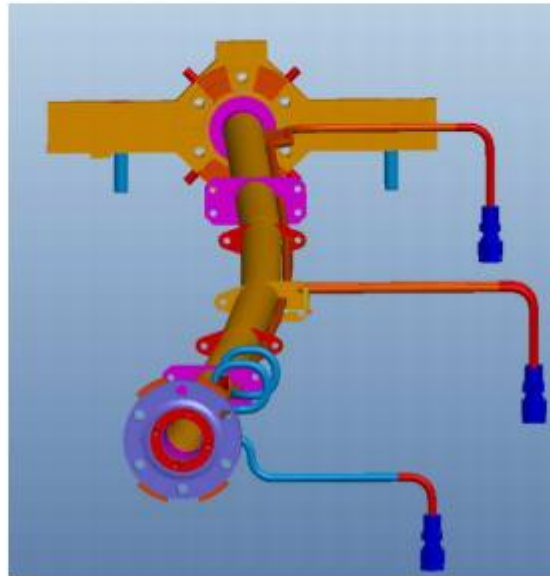
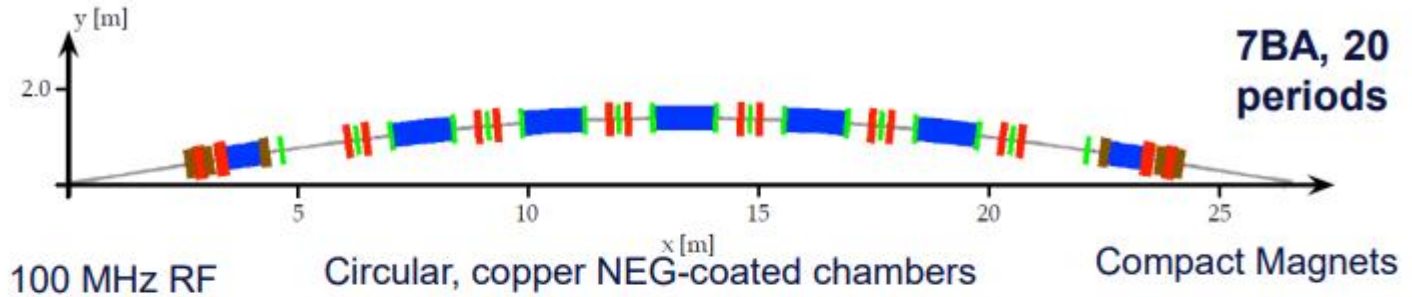


Synchrotron radiation (sources + exp. stations)



Рентгеноструктурный анализ. Источники рентгеновского излучения.

MAX IV 3 GeV ring: 528 m, 330 pm rad



Рентгеноструктурный анализ.

Взаимодействие фотонов с атомами.

Рентгеновские лучи рассеиваются на каждом атоме молекулы:

- Рассеяние происходит на электронной оболочке, **не на ядрах!**
- В основном происходит упругое рассеяние (практически без изменения энергии фотонов).
- Рассеяние происходит во **всех направлениях.**

Следствия:

- РСА не различает атомы с близким и малым числом электронов.

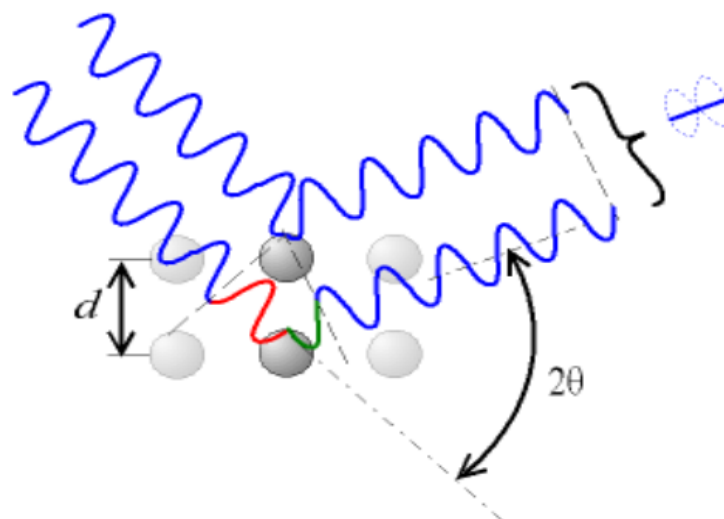
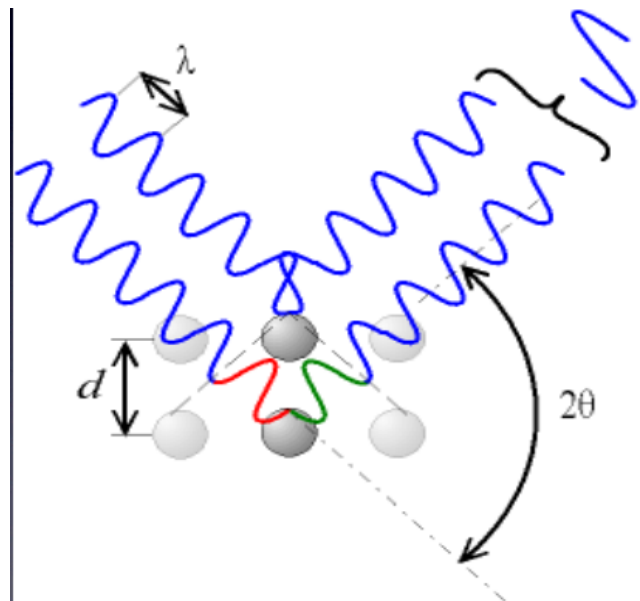
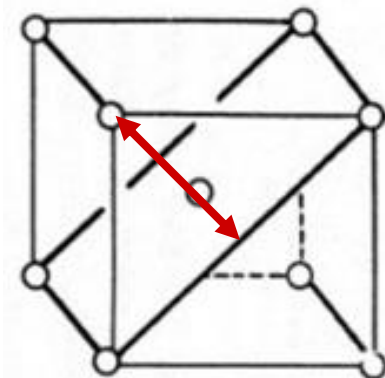
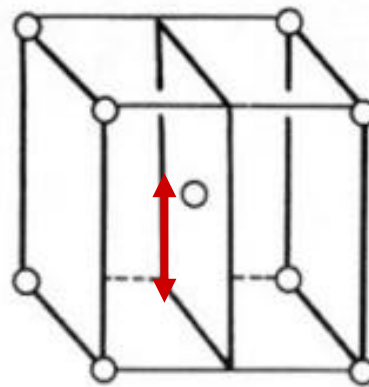
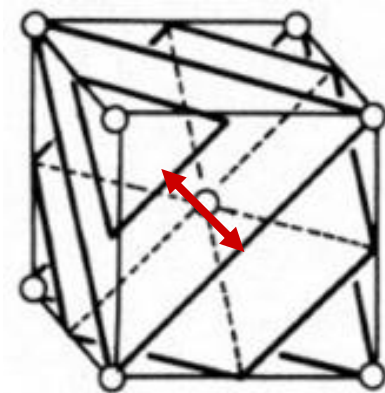


Рентгеноструктурный анализ.

Рассеяние на решетке/кристалле.

Рассеяние в кристалле происходит **на плоскостях** в решетке из атомов. Появляются картины дифракции.

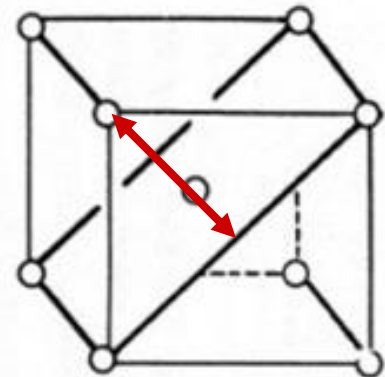
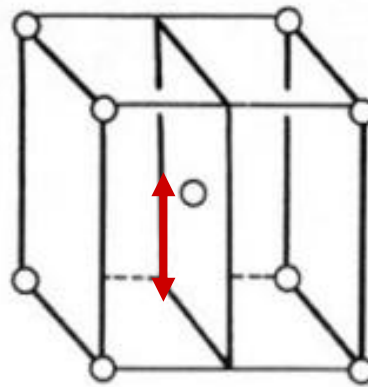
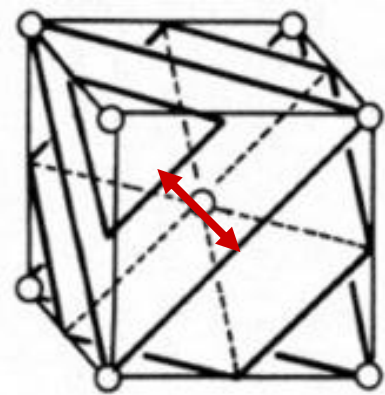
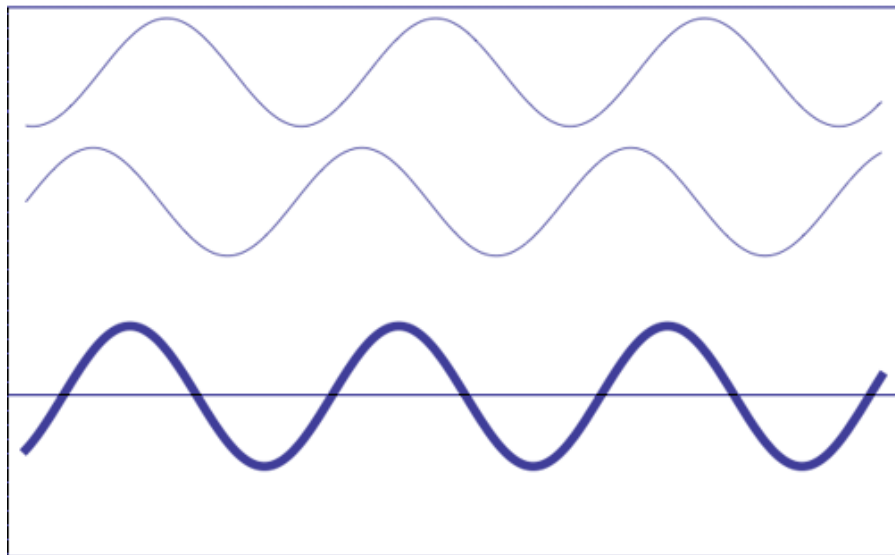
Конструктивная дифракция появляется при выполнении закона Брэгга: **$n \lambda = 2d \sin\theta$**



Рентгеноструктурный анализ. Рассеяние на решетке/кристалле.

Рассеяние в кристалле происходит **на плоскостях** в решетке из атомов. Появляются картины дифракции.

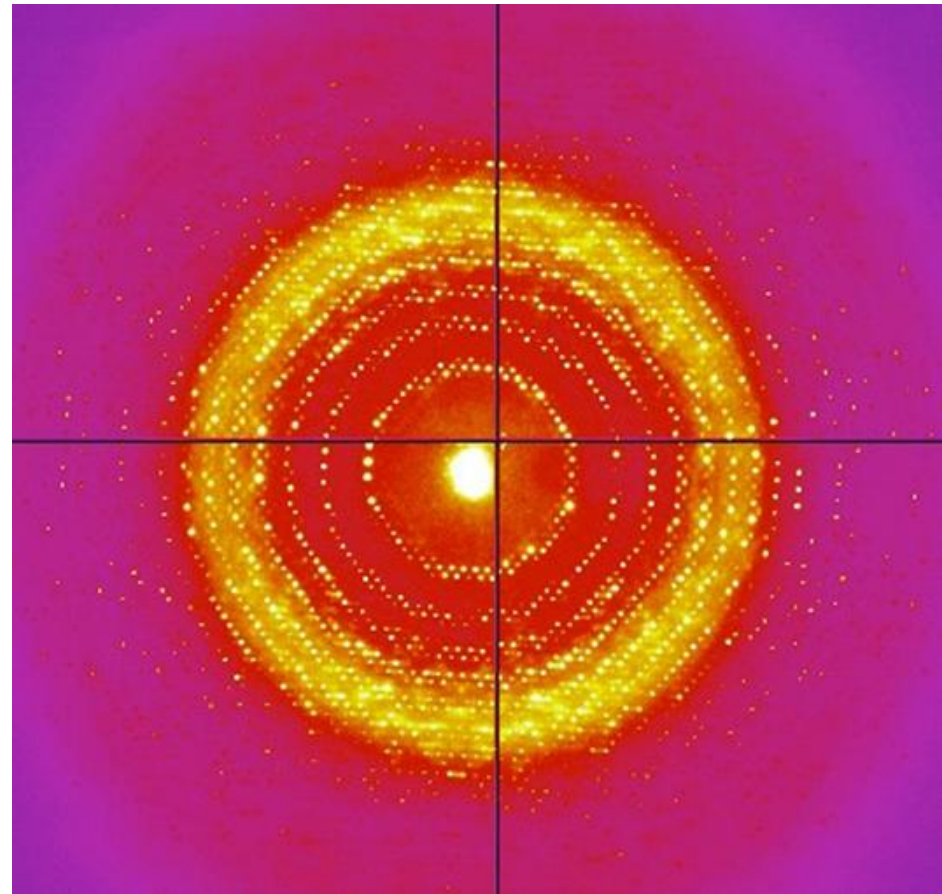
Конструктивная дифракция появляется при выполнении закона Брэгга: $n \lambda = 2d \sin\theta$



Рентгеноструктурный анализ.

Последствия рассеяния на решетке.

- Дифракционная картина представляет собой набор точек - рефлексов.
- Для каждой точки можно определить ее интенсивность и положение.
- Положение точки на дифрактограмме зависит от **элементарной ячейки** кристалла.
- Каждой точке присваивают индексы h , k и l , которые описывают положение пиков в **обратном пространстве**.

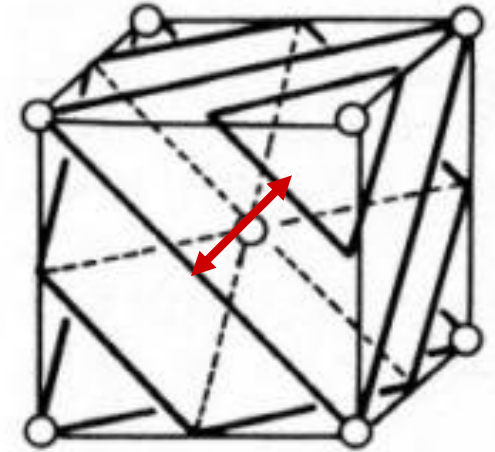
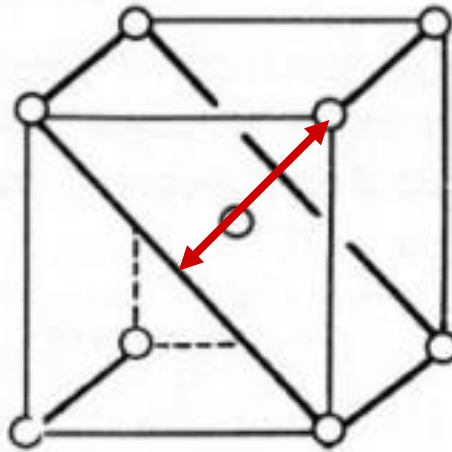
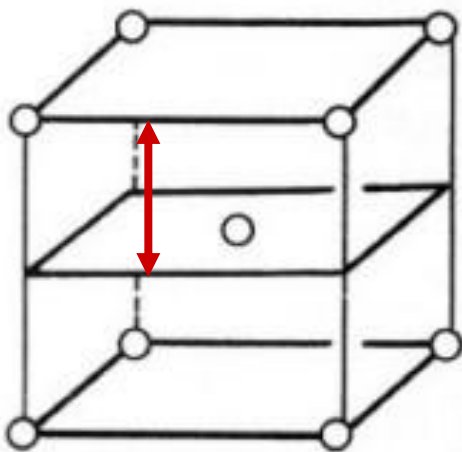
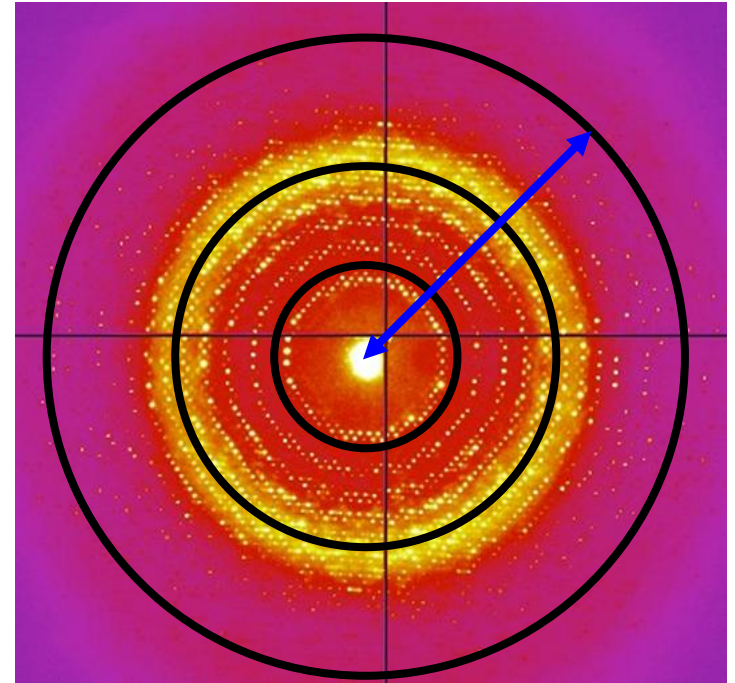


Рентгеноструктурный анализ. Разрешение РСА.

Закона Брэгга: $n \lambda = 2d \sin \theta$

Разрешение: $d = \lambda / 2 \sin \theta$

При длине волны 2 \AA и
максимальном угле
дифракции 45° $d = 1.4 \text{ \AA}$



Рентгеноструктурный анализ.

Кратко о кристаллах

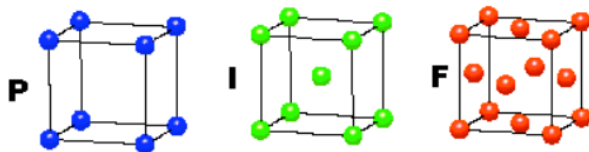
Типы кристаллических решеток

Вещество может быть в разных кристаллических формах

CUBIC

$$a = b = c$$

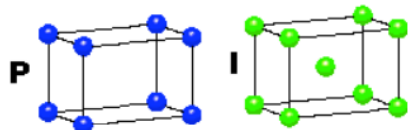
$$\alpha = \beta = \gamma = 90^\circ$$



TETRAGONAL

$$a = b \neq c$$

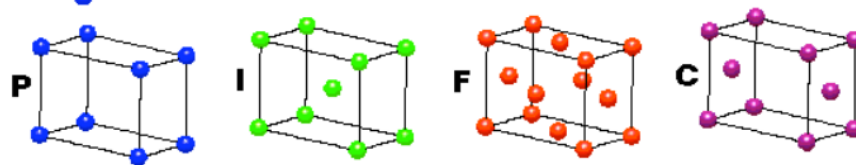
$$\alpha = \beta = \gamma = 90^\circ$$



ORTHORHOMBIC

$$a \neq b \neq c$$

$$\alpha = \beta = \gamma = 90^\circ$$

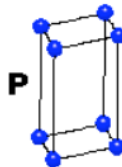


HEXAGONAL

$$a = b \neq c$$

$$\alpha = \beta = 90^\circ$$

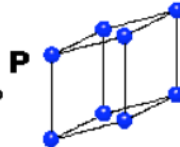
$$\gamma = 120^\circ$$



TRIGONAL

$$a = b = c$$

$$\alpha = \beta = \gamma \neq 90^\circ$$

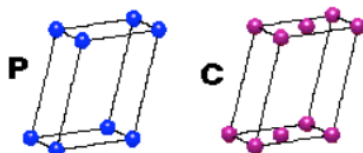


MONOCLINIC

$$a \neq b \neq c$$

$$\alpha = \gamma = 90^\circ$$

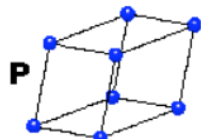
$$\beta \neq 120^\circ$$



TRICLINIC

$$a \neq b \neq c$$

$$\alpha \neq \beta \neq \gamma \neq 90^\circ$$



4 Types of Unit Cell

P = Primitive

I = Body-Centred

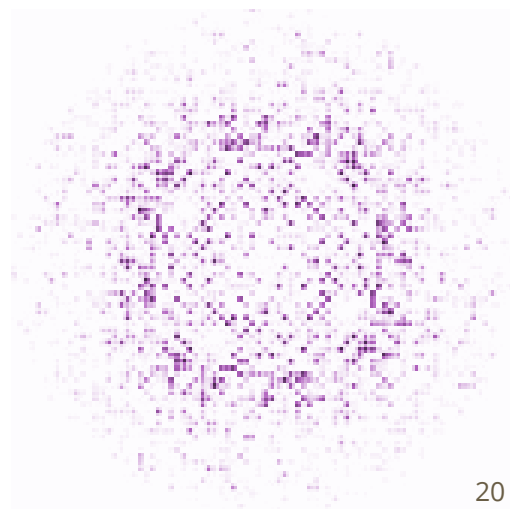
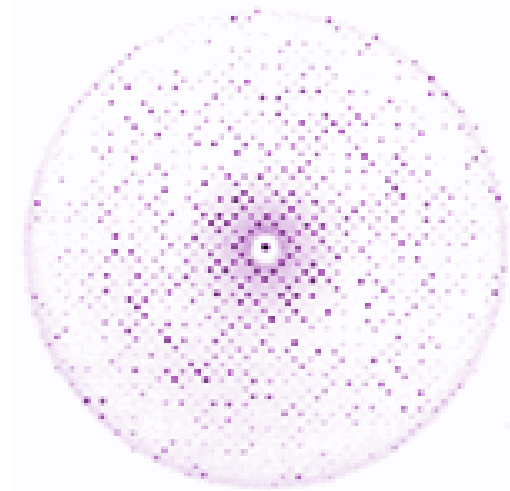
F = Face-Centred

C = Side-Centred

+

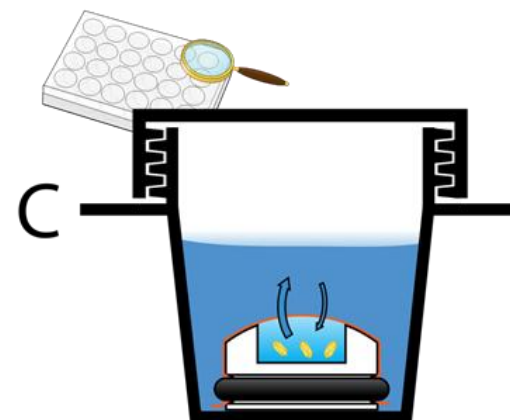
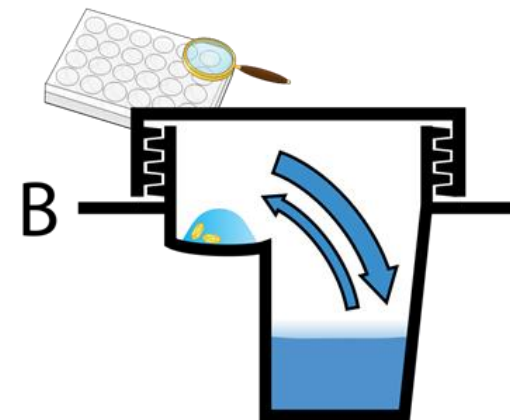
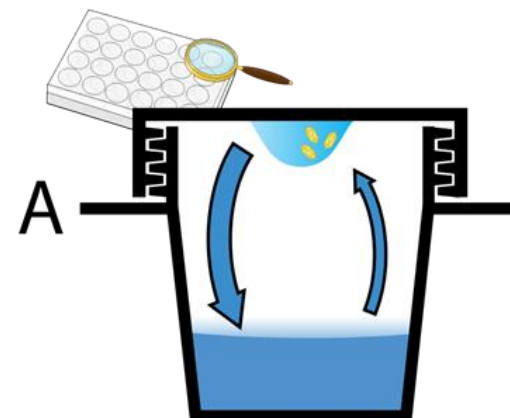
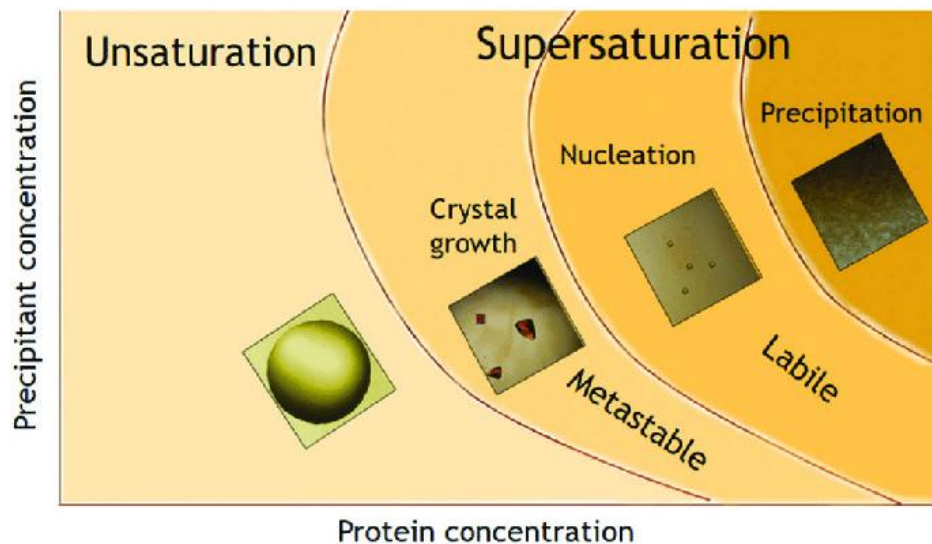
7 Crystal Classes

→ 14 Bravais Lattices



Рентгеноструктурный анализ. Где взять белковые кристаллы?

- Необходимо много **очень чистого растворимого** белка (**10-50 мг**)
- Кристаллизация происходит за счет испарения растворителя
- Кристаллы растут долго (день - неделя)
- Размер кристалла 0,1-0,5 мм (лучше больше)
- Обычно кристаллы на 50% **состоят из воды**
- Зачастую кристаллы нестабильны

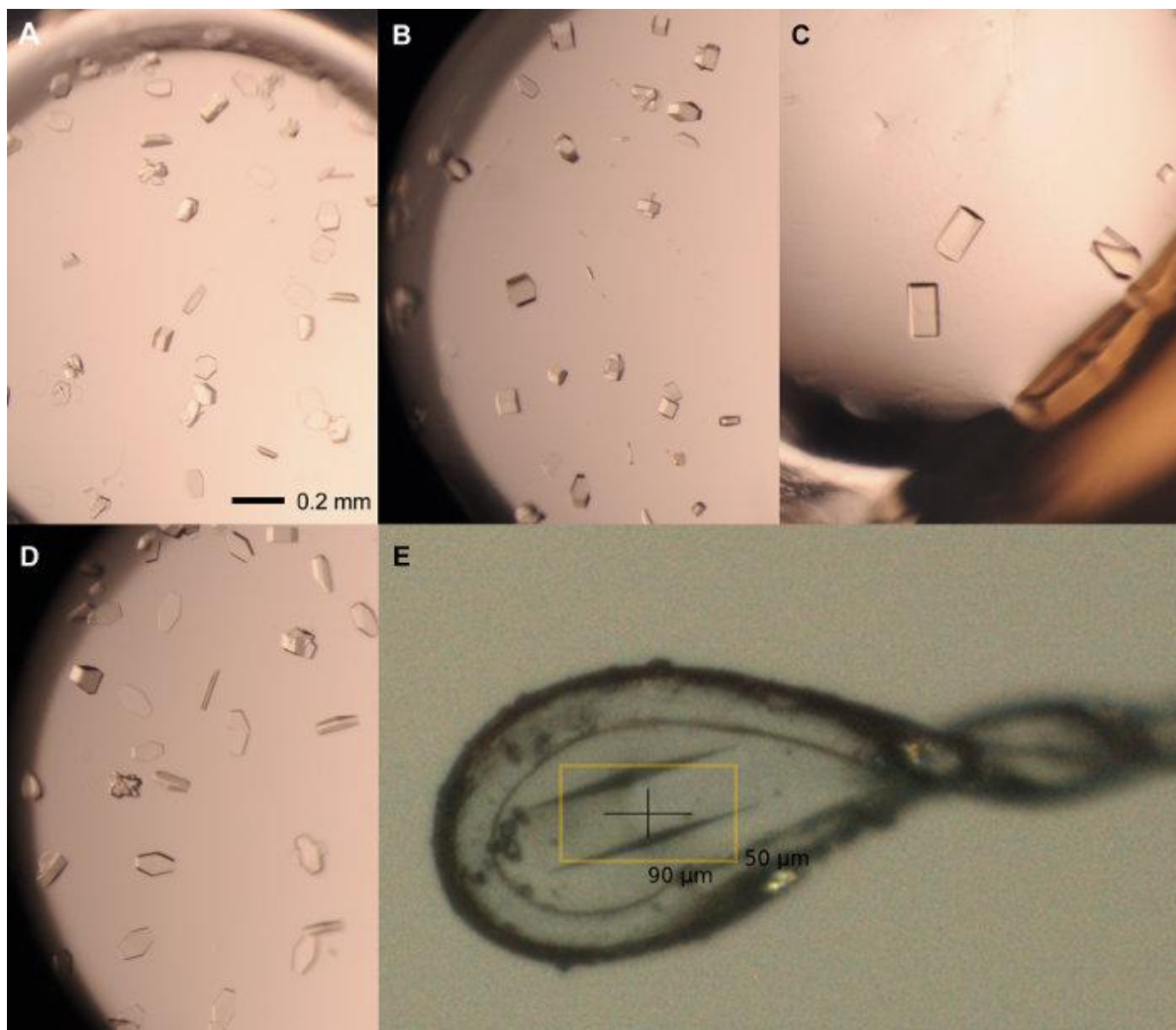


Рентгеноструктурный анализ. Где взять белковые кристаллы?

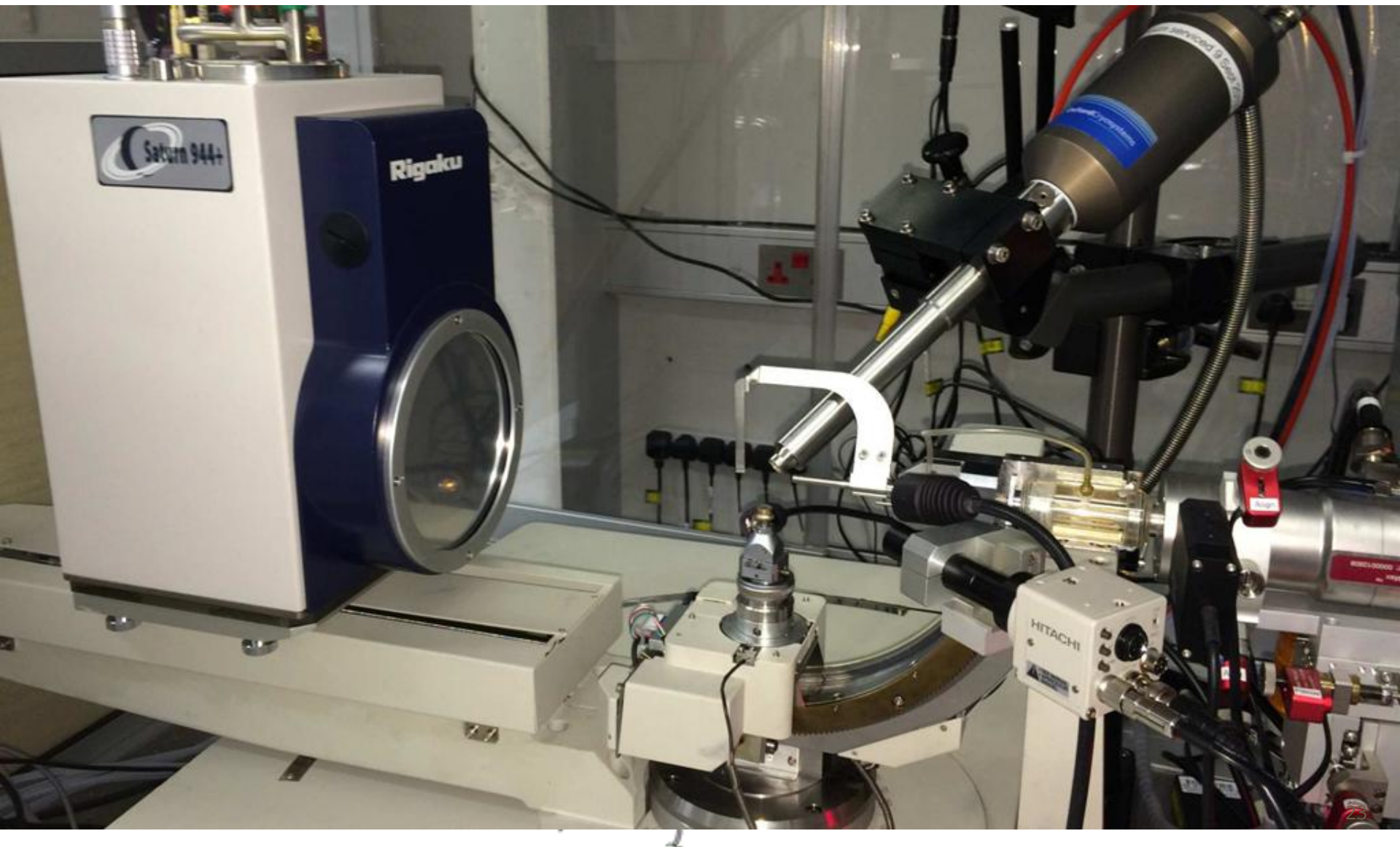
На поиск условий кристаллизации и подбор “преципитанта” (осадителя) уходит много времени и материала.

На данный момент этот процесс роботизирован.

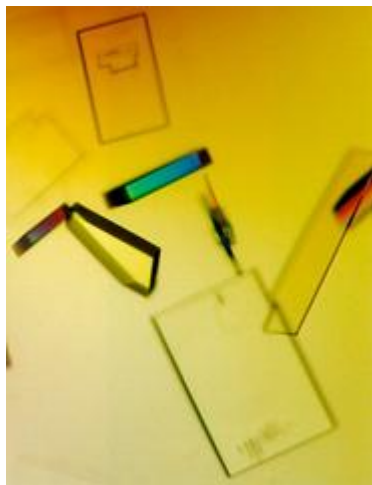
Не факт, что большой кристалл даст хорошую дифракцию.



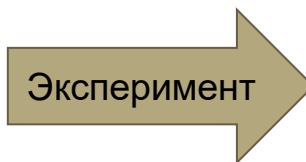
Рентгеноструктурный анализ. Схема проведения эксперимента.



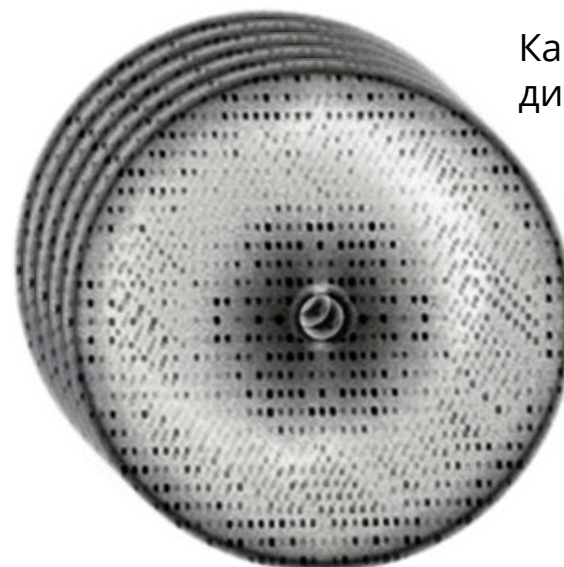
Рентгеноструктурный анализ. Анализ и интерпретация данных.



Кристаллы



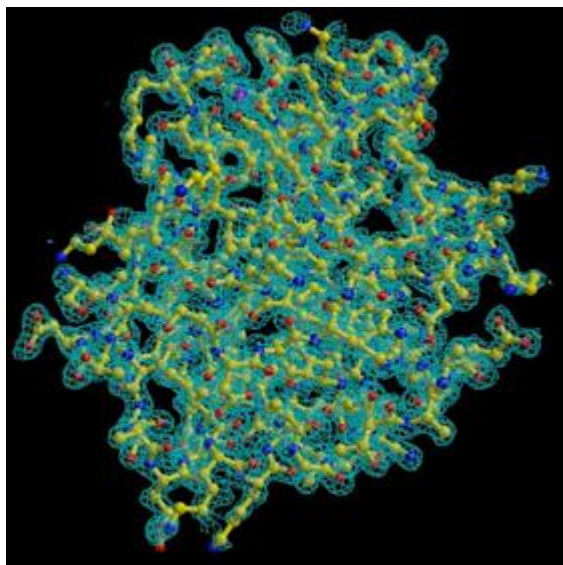
Эксперимент



Карты дифракции



Магия
Обработка
данных

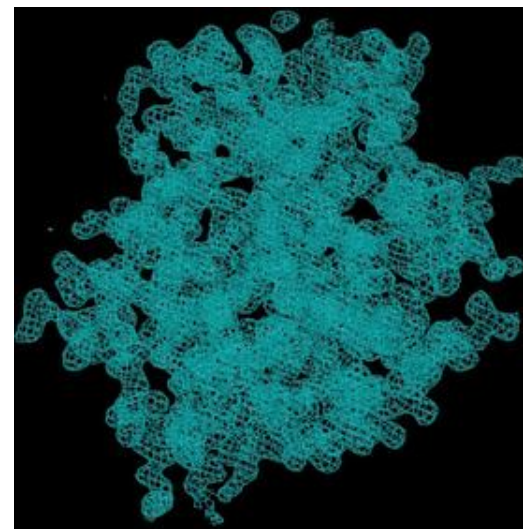


Итоговая
модель



Вписыва
ние

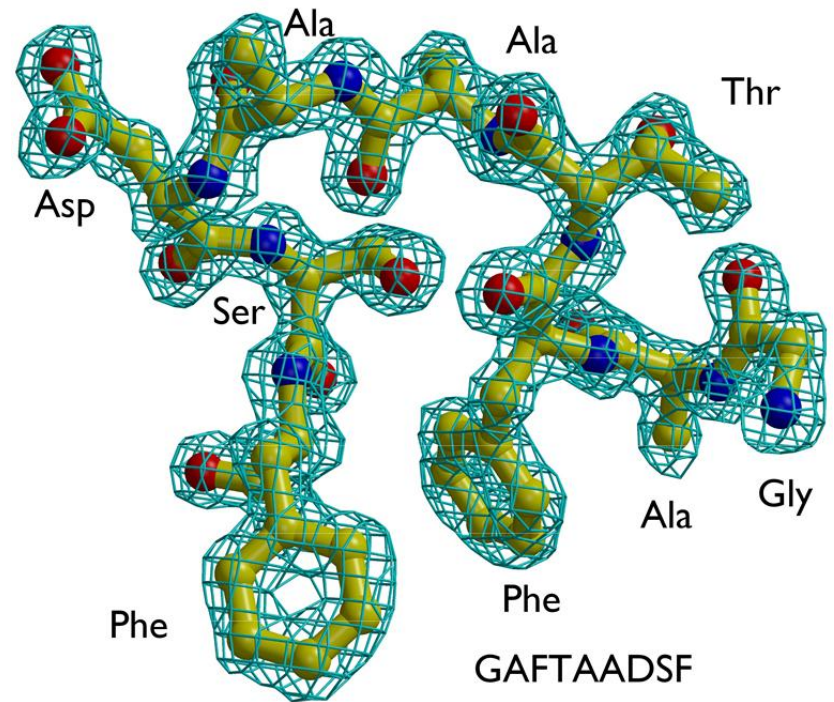
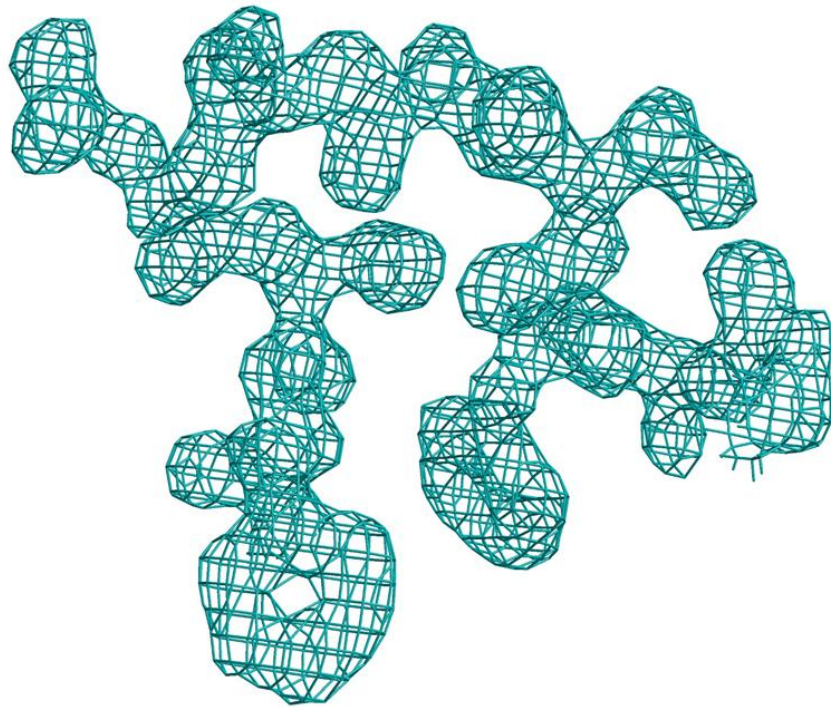
"Docking"
"Fitting"



Карта
распределения
электронной
плотности

Рентгеноструктурный анализ. Создание подходящих молекулярных моделей.

Задача докинга моделей в электронную плотность - задача поиска характерных форм. Необходимо знание первичной структуры белка.



Рентгеноструктурный анализ.

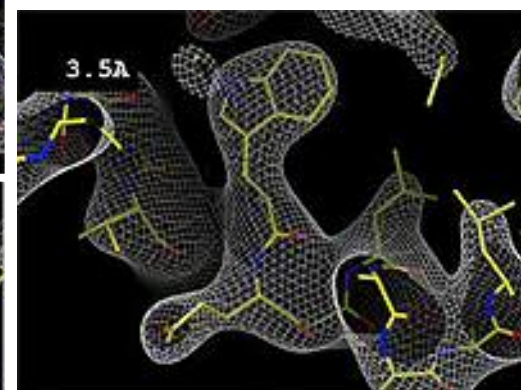
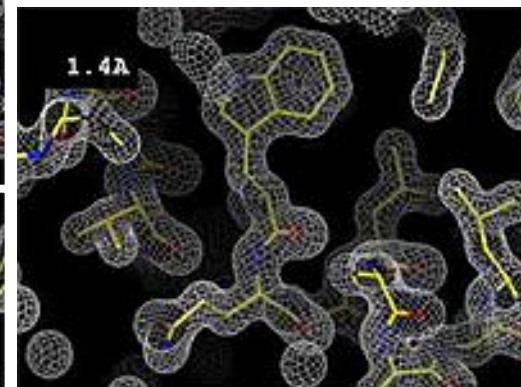
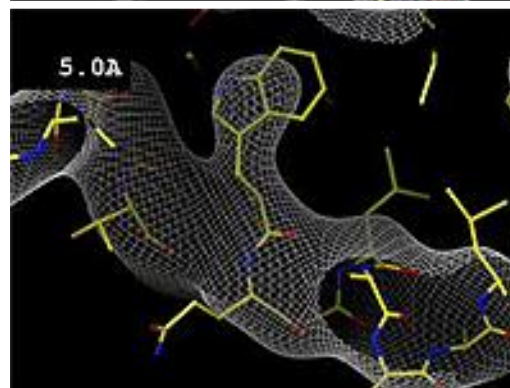
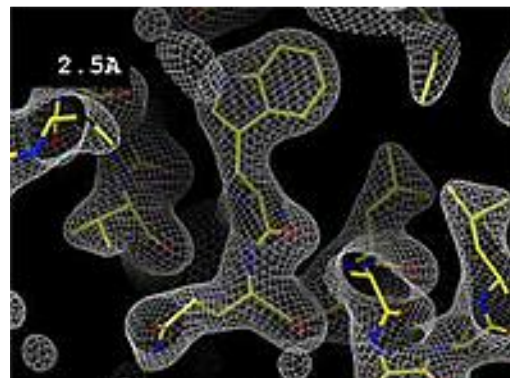
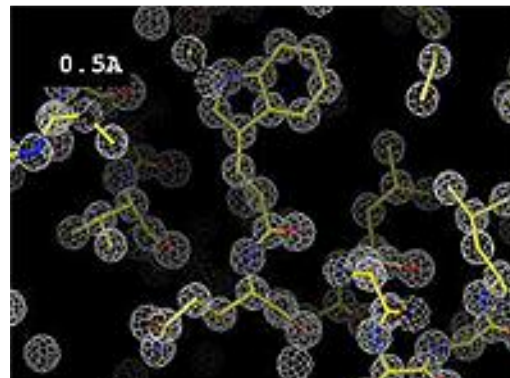
Влияние разрешения на карту электронной плотности.

< 1,2 Å Отличное разрешение, виден ход белкового остова и боковые цепи. Могут быть определены протоны.

< 2,5 Å Хорошее разрешение, виден ход белкового остова и большинство боковых цепей

< 3,5 Å Нормально.

5,0 Å Плохо. Виден лишь белковый остов.



CC James Holton

Рентгеноструктурный анализ.

Структурные файлы. *.pdb

Descriptor		Conformation index					Atom element					
Atom number		Residue name					Beta-factor					
Atom name		Chain name					Occupancy					
		Residue number										
...												
ATOM	22	CG	AGLN	A	3	-30.506	25.807	16.125	0.50	43.30		C
ATOM	23	CD	AGLN	A	3	-29.018	25.546	15.959	0.50	43.01		C
ATOM	24	OE1	AGLN	A	3	-28.187	26.374	16.332	0.50	43.76		O
ATOM	25	NE2	AGLN	A	3	-28.675	24.391	15.403	0.50	41.73		N
ATOM	26	N	BGLN	A	3	-33.005	25.819	17.852	0.50	45.16		N
ATOM	27	CA	BGLN	A	3	-32.788	24.834	16.796	0.50	44.25		C
ATOM	28	C	BGLN	A	3	-33.529	23.510	17.024	0.50	43.75		C
...												
HETATM	4140	C1	F09	C1131		-21.017	-3.092	-1.563	1.00	60.56		C
HETATM	4141	C2	F09	C1131		-21.015	-1.597	-1.357	1.00	60.40		C
...												
						X	Y	Z	coordinates in Å			

Рентгеноструктурный анализ.

В сухом остатке:

РСА позволяет:

- Исследовать кристаллизованные белки и их комплексы с другими веществами
- Точнее других методов определять взаимное расположение атомов в кристалле

Трудности:

- Не все белки можно кристаллизовать
- Нужно очень много материала
- Чаще всего необходим синхротрон
- Структура белка в кристалле может быть не природной

Дополнительная информация:

<http://www.xtal.iqfr.csic.es/Cristalografia/index-en.html>

J. P. Glusker and K. N. Trueblood. Crystal Structure Analysis: A Primer. 2010

Рентгеноструктурный анализ. Что делать?

RCSB **PDB** PROTEIN DATA BANK 148969 Biological Macromolecular Structures Enabling Breakthroughs in Research and Education

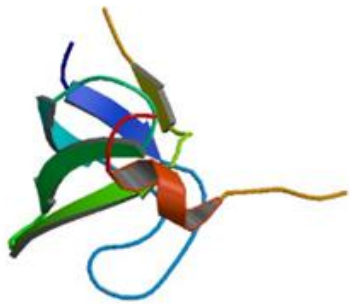
Search by PDB ID, author, macromolecule, sequence, or ligands **Go**

Advanced Search | Browse by Annotations

PDB-101 WORLDWIDE PDB PROTEIN DATA BANK EMDataResource EMBL Data Resource for 2008 Nucleic Acid Database Worldwide Protein Data Bank Foundation

Structure Summary 3D View Annotations Sequence Sequence Similarity Structure Similarity Experiment

Biological Assembly 1 ?



1UTI
Mona/Gads SH3C in complex with HPK derived peptide
DOI: 10.2210/pdb1UTI/pdb
Classification: [SIGNALING PROTEIN REGULATOR](#)
Organism(s): [Mus musculus](#)
Expression System: [Escherichia coli BL21\(DE3\)](#)
Deposited: 2003-12-09 Released: 2004-05-06
Deposition Author(s): [Lewitzky, M.](#), [Harkioliaki, M.](#), [Domart, M.C.](#), [Feller, S.M.](#)

Experimental Data Snapshot
Method: X-RAY DIFFRACTION
Resolution: 1.5 Å
R-Value Free: 0.221
R-Value Work: 0.207

wwPDB Validation **3D Report** Full Report

Metric	Percentile Ranks	Value
Rfree		0.231
Clashscore		5
Ramachandran outliers		0
Sidechain outliers		3.2%
RSRZ outliers		20.5%

Worse Better
■ Percentile relative to all X-ray structures
■ Percentile relative to X-ray structures of similar resolution

3D View: Structure | Electron Density

Standalone Viewers
[Protein Workshop](#) | [Ligand Explorer](#)

Рекомендованные программы: UCSF Chimera, PyMol, VMD

Рентгеноструктурный анализ. Что дальше?

PDB - банк данных (не база).

- Интерпретация функций белка на молекулярном уровне.
- Анализ контактов между макромолекулами (и лигандами).
- Поиск структурных гомологов макромолекул.
- Построение теоретических предсказательных моделей фолдинга белков.
- Создание “Силовых полей” для проведения экспериментов методами молекулярного моделирования.
- Создание многомасштабных моделей клеток.

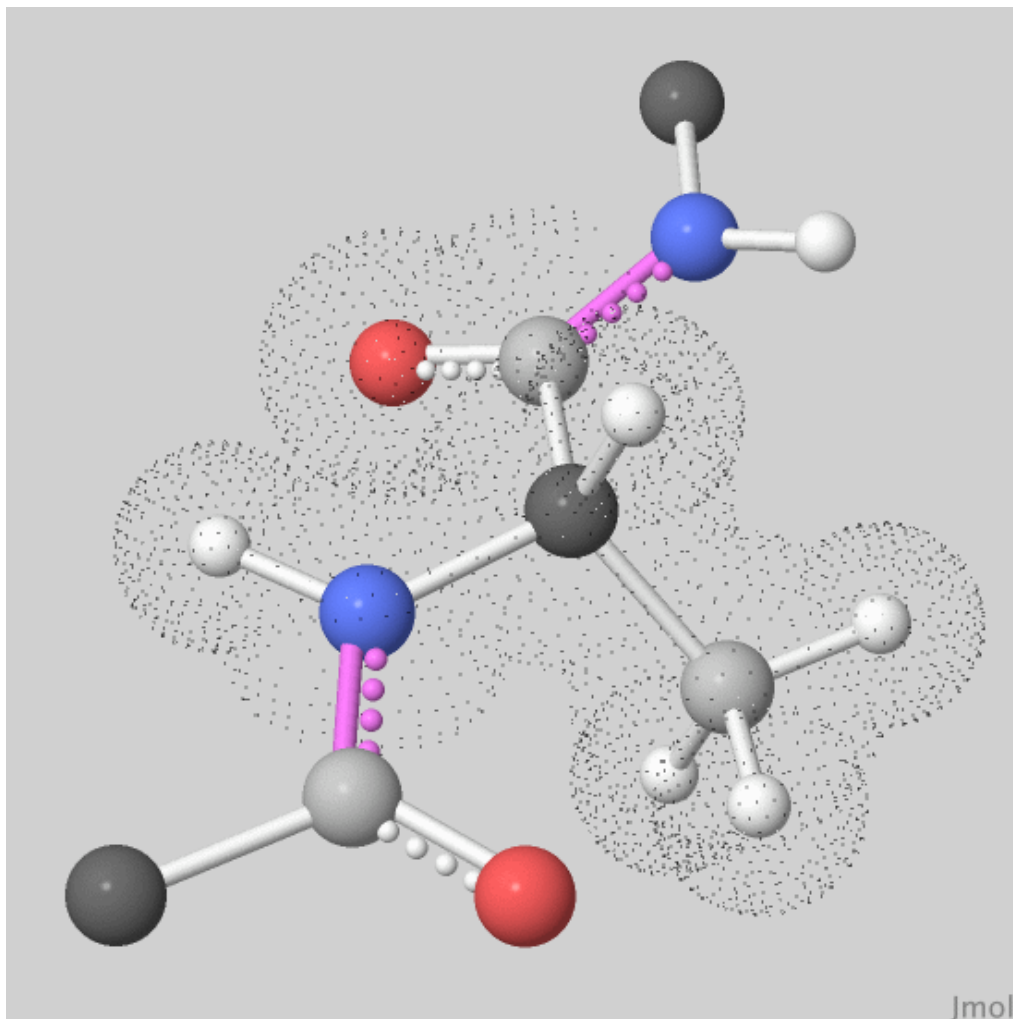


Внутренние координаты белков.

Структурные файлы. *.pdb

Descriptor		Conformation index					Atom element				
Atom number		Residue name					Beta-factor				
Atom name		Chain name					Occupancy				
		Residue number									
...											
ATOM	22	CG	AGLN	A	3	-30.506	25.807	16.125	0.50	43.30	C
ATOM	23	CD	AGLN	A	3	-29.018	25.546	15.959	0.50	43.01	C
ATOM	24	OE1	AGLN	A	3	-28.187	26.374	16.332	0.50	43.76	O
ATOM	25	NE2	AGLN	A	3	-28.675	24.391	15.403	0.50	41.73	N
ATOM	26	N	BGLN	A	3	-33.005	25.819	17.852	0.50	45.16	N
ATOM	27	CA	BGLN	A	3	-32.788	24.834	16.796	0.50	44.25	C
ATOM	28	C	BGLN	A	3	-33.529	23.510	17.024	0.50	43.75	C
...											
HETATM	4140	C1	F09	C1131		-21.017	-3.092	-1.563	1.00	60.56	C
HETATM	4141	C2	F09	C1131		-21.015	-1.597	-1.357	1.00	60.40	C
...											
						X	Y	Z	coordinates in Å		

Внутренние координаты белков.



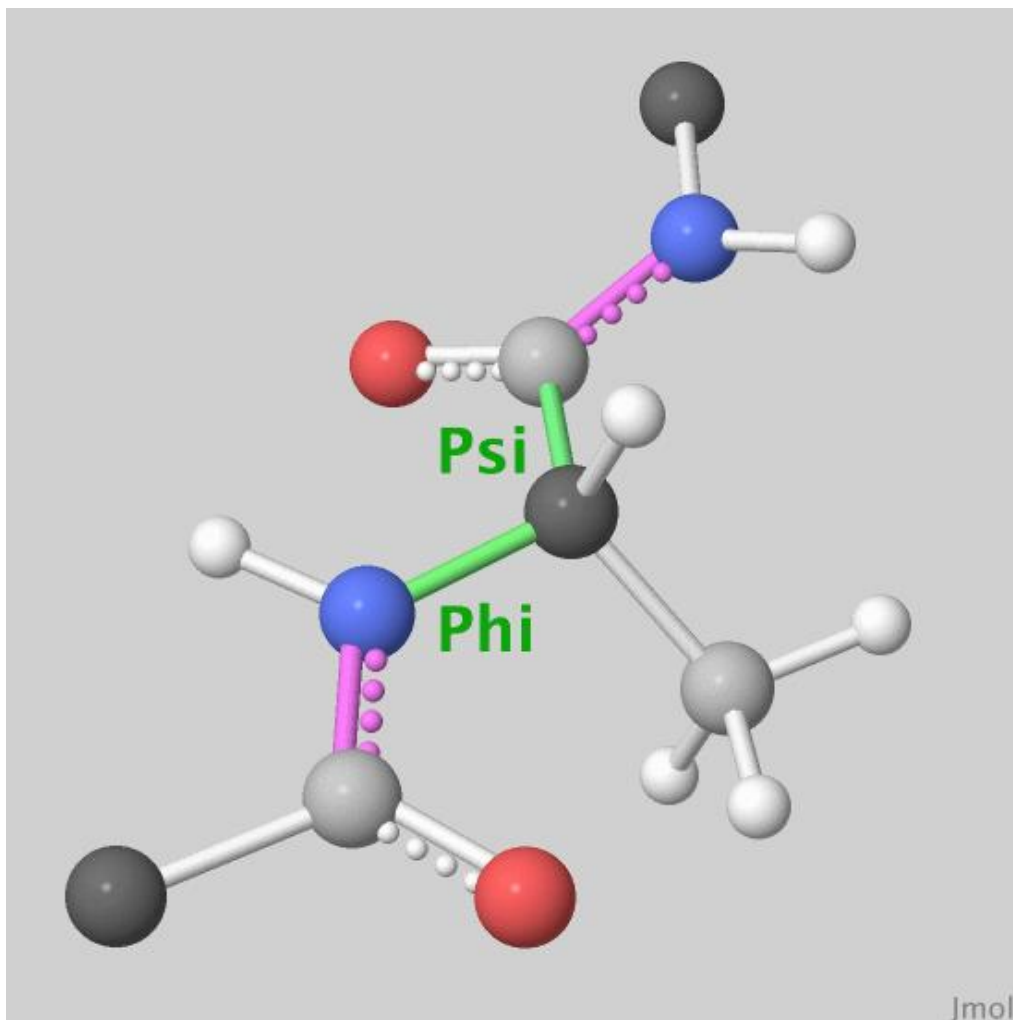
С α С Н N O

**Черными точками
выделен аланин.**

Соседи показаны лишь
частично, аланин связан
с соседями

пептидными связями.

Внутренние координаты белков.



С α С Н N O

В каждой аминокислоте
вращение может
происходить вокруг углов

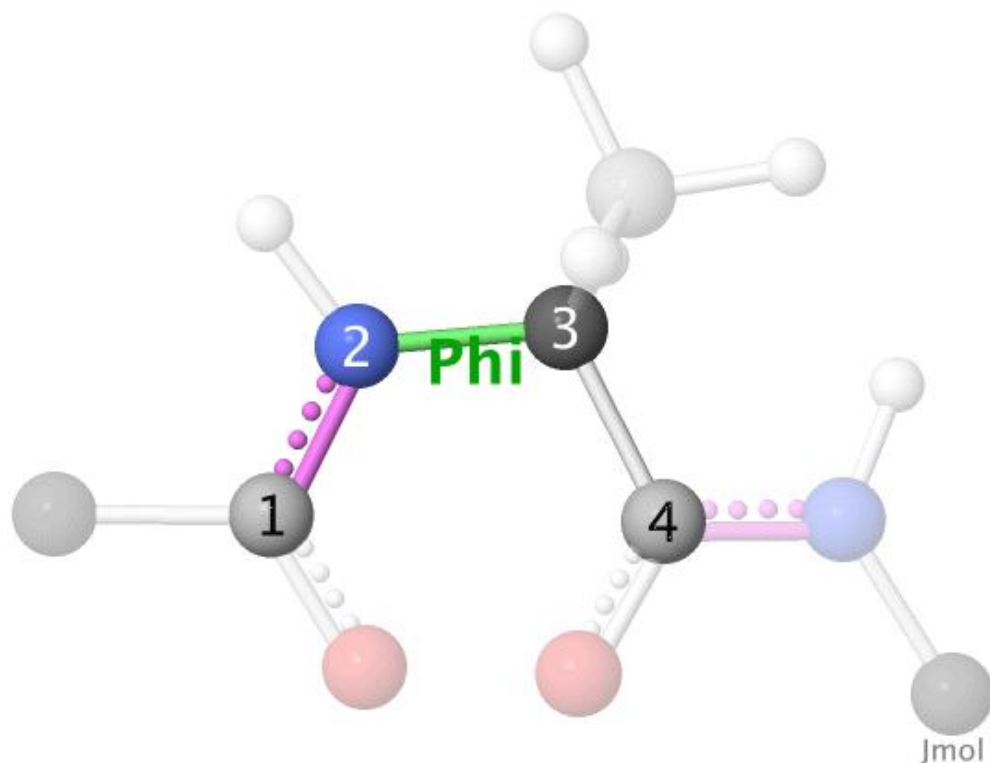
Фи (φ) и **Пси (ψ)**.

Пептидные связи не
вращаются, и обычно
находятся в транс-
ориентации.

Внутренние координаты белков.

$$\text{Phi} = 0^\circ$$

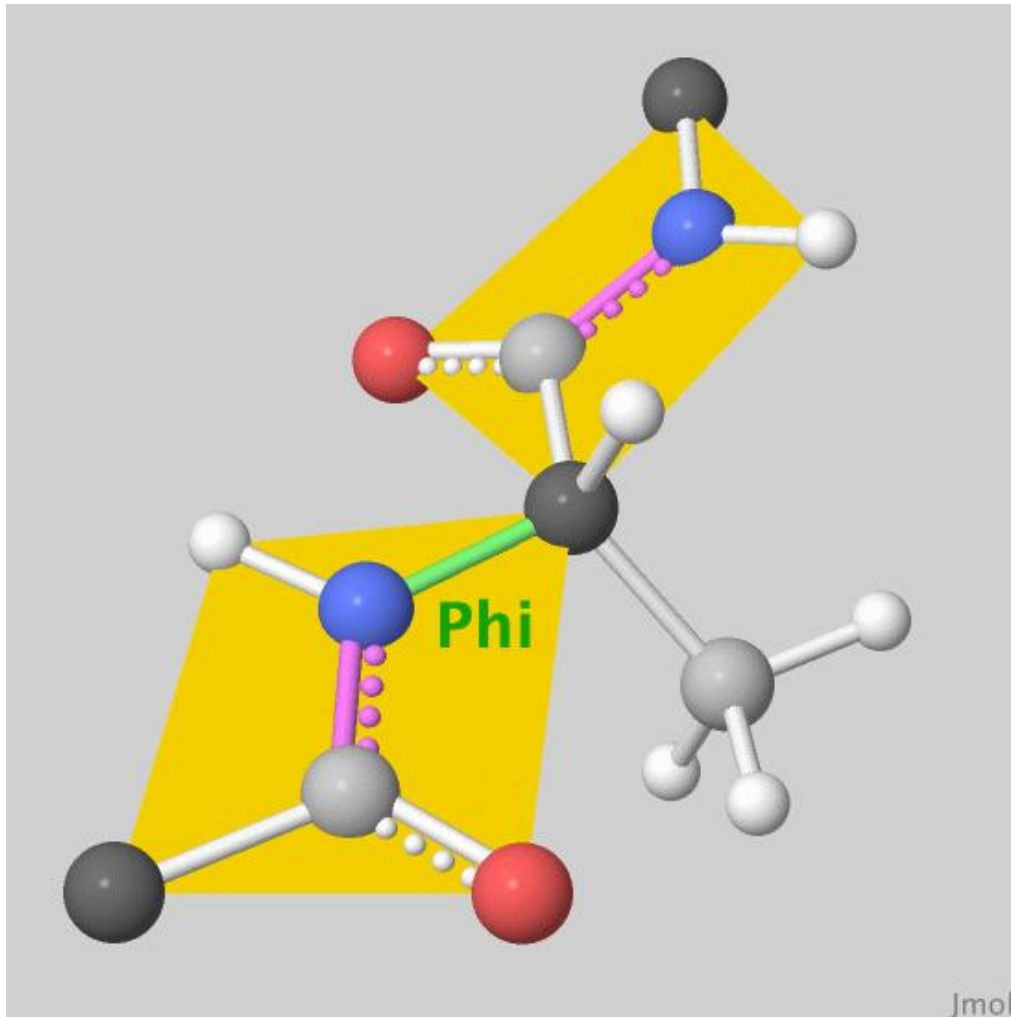
C α **C** **H** **N** **O**



Двугранные углы
могут изменяться в
пределах от -180° до
 180° .

Внутренние координаты белков.

С α С Н N O



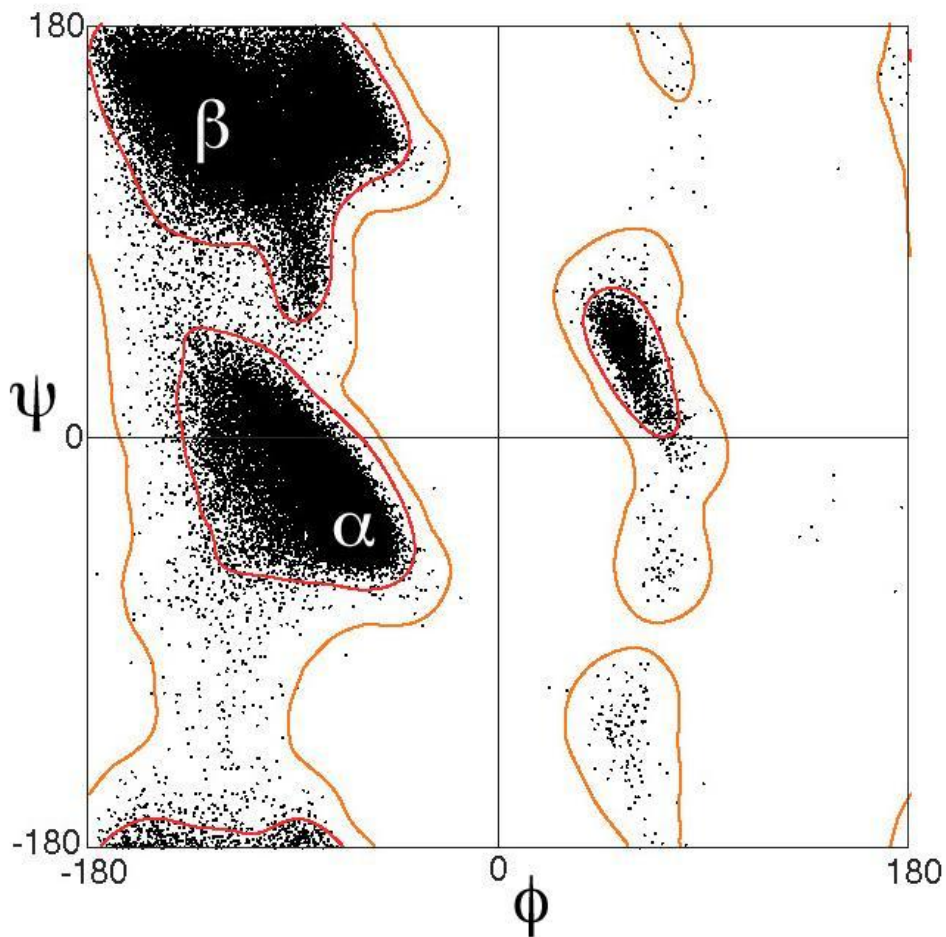
Так как **пептидные связи** не вращаются, по 6 атомов при каждой связи находятся в одной **плоскости**.

Однако, большинство значений углов ϕ и ψ не реализуемы на практике, так как приводят к серьезным перекрываниям атомов.

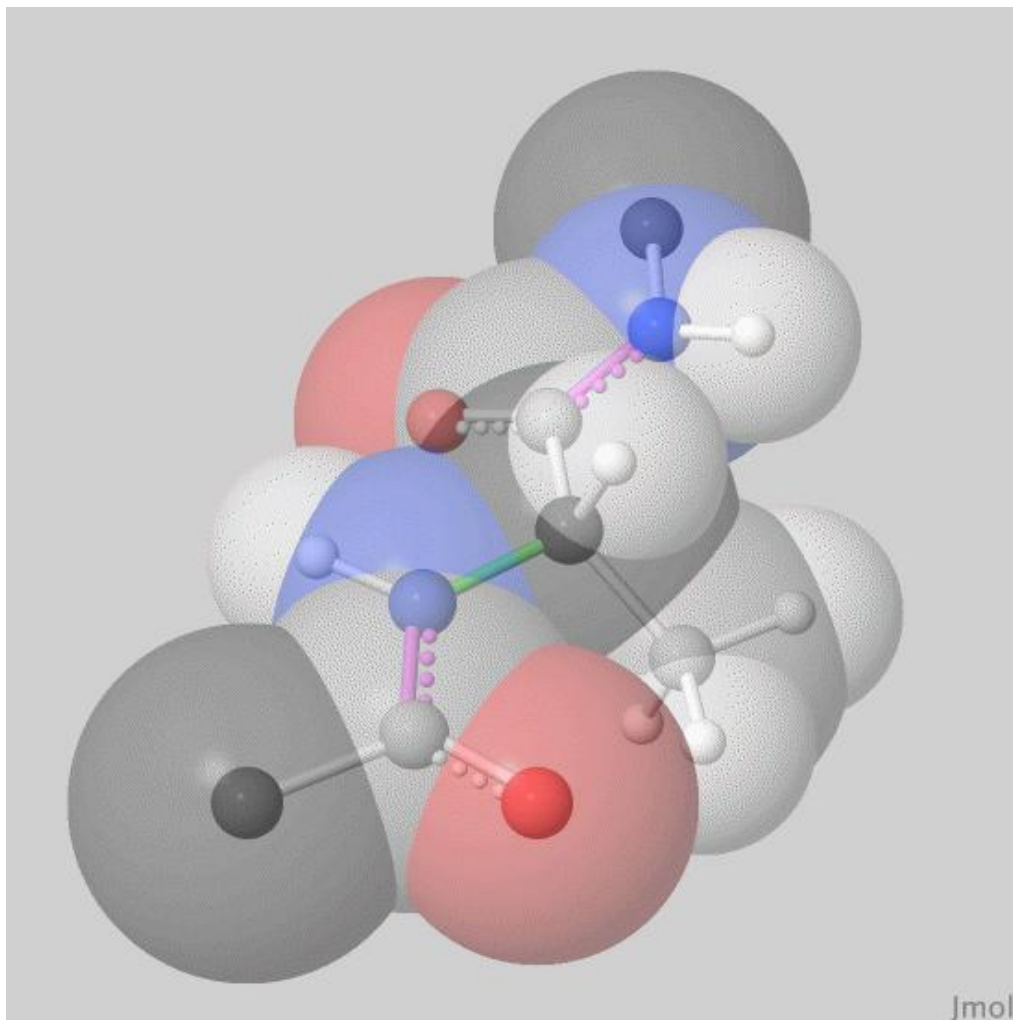
Шаростержневые модели не отражают истинного размера атомов

Внутренние координаты белков.

Взгляните на **карту Рамачандрана**. Каждая точка на нем соответствует аминокислоте с парой углов ϕ и ψ . Эти точки (аминокислоты за исключением Gly, Pro) случайным образом отобраны из PDB. Видно, что точки лежат в кластерах, причем данные комбинации углов характерны для альфа спиралей (α), бета слоев (β), и поворотов (справа).



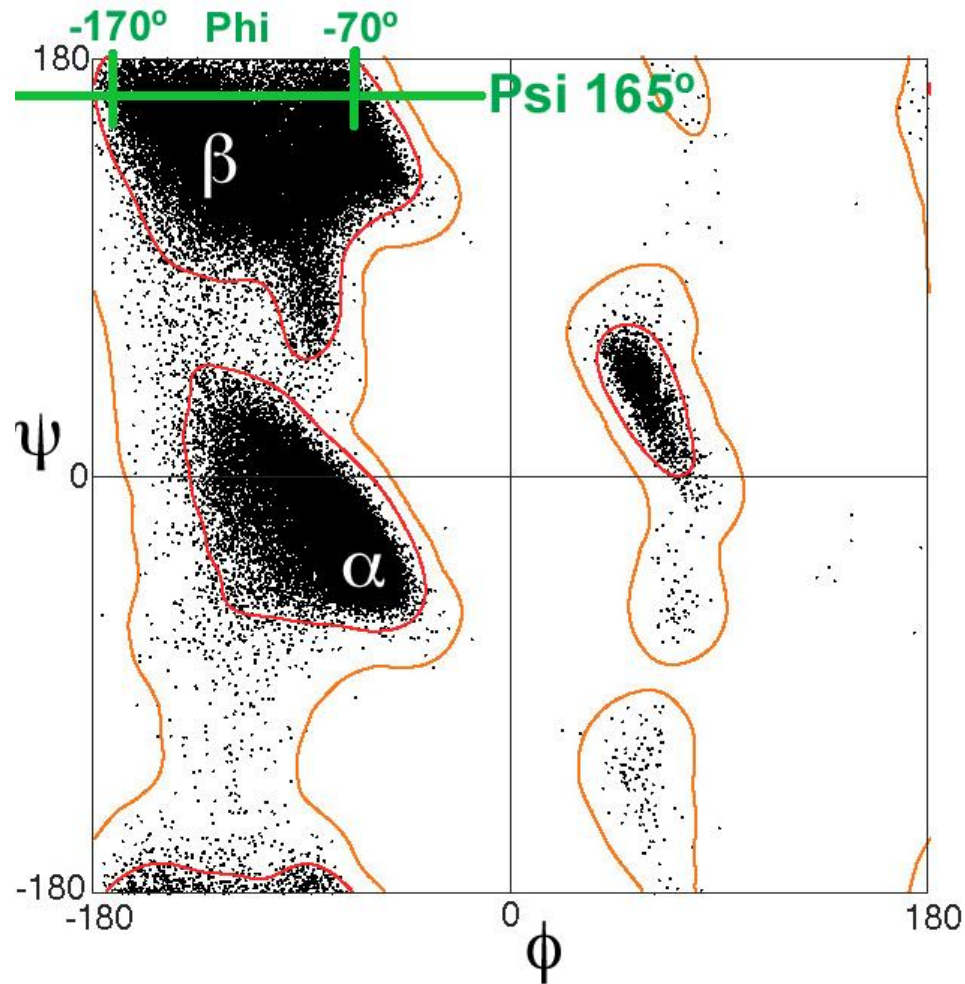
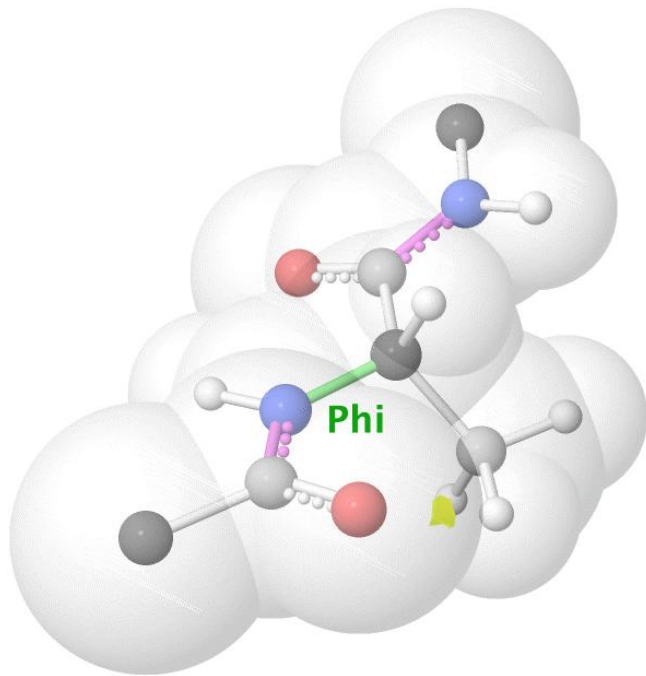
Внутренние координаты белков.



Ca C H N O

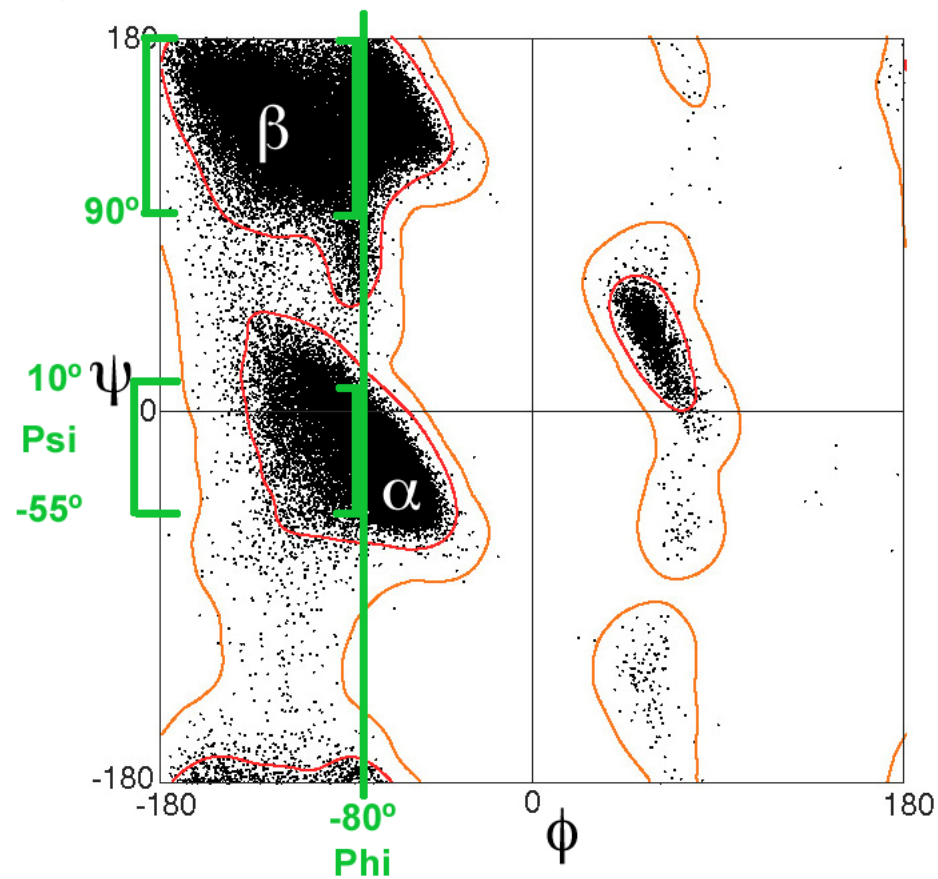
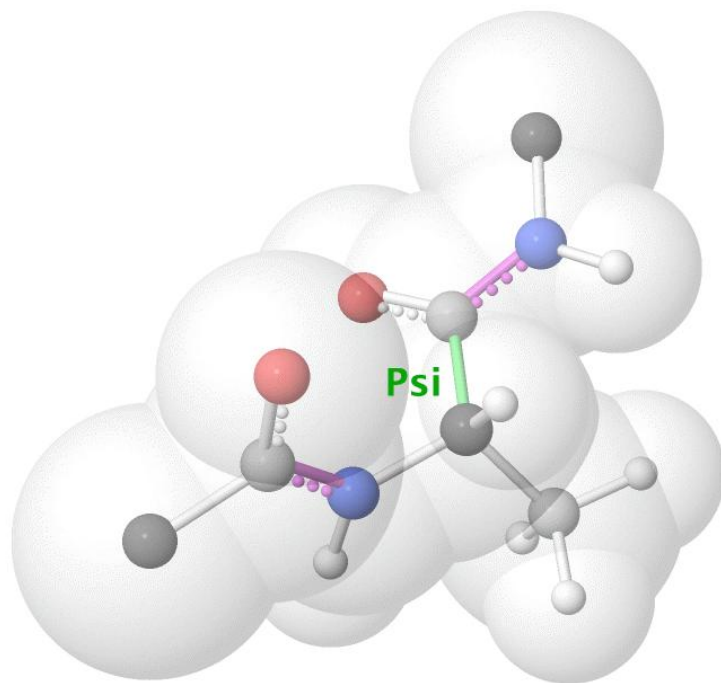
Здесь показаны
размеры равные
Ван-дер-Ваальсовым
радиусам атомов.

Внутренние координаты белков.



При условии, что ψ зафиксирован при 165° , большая часть значений угла ϕ (кроме -170 -70) приводит к возникновению **стерических перекрываний**.

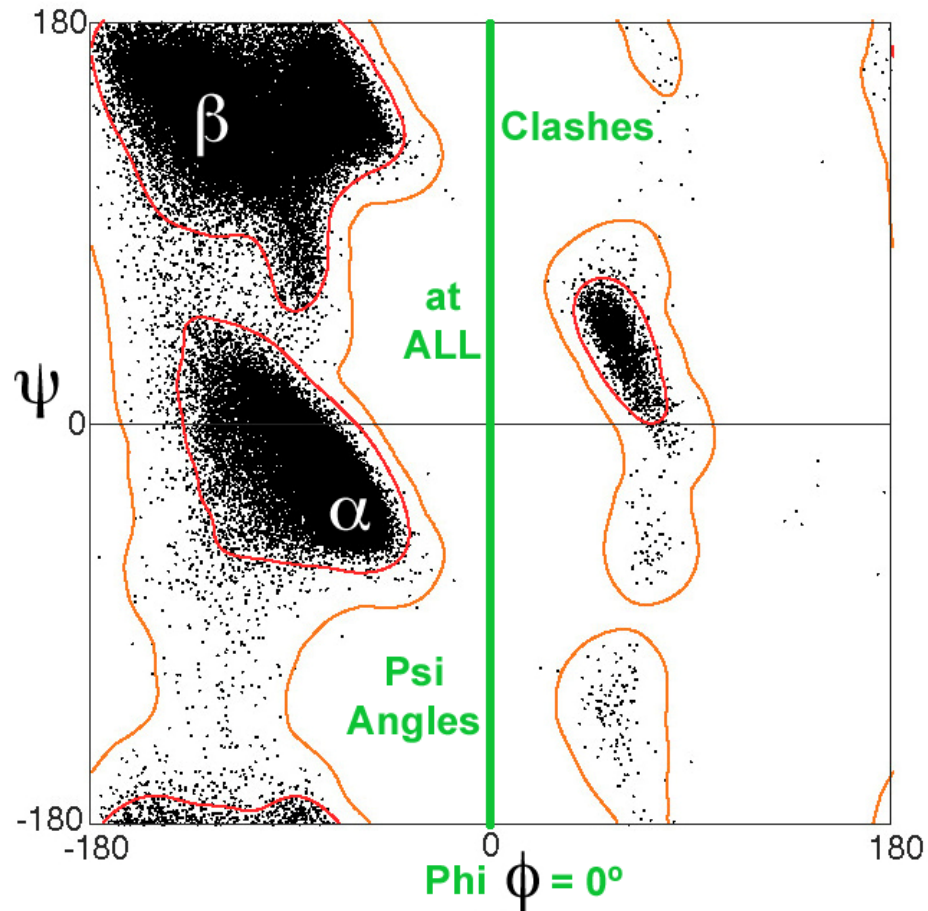
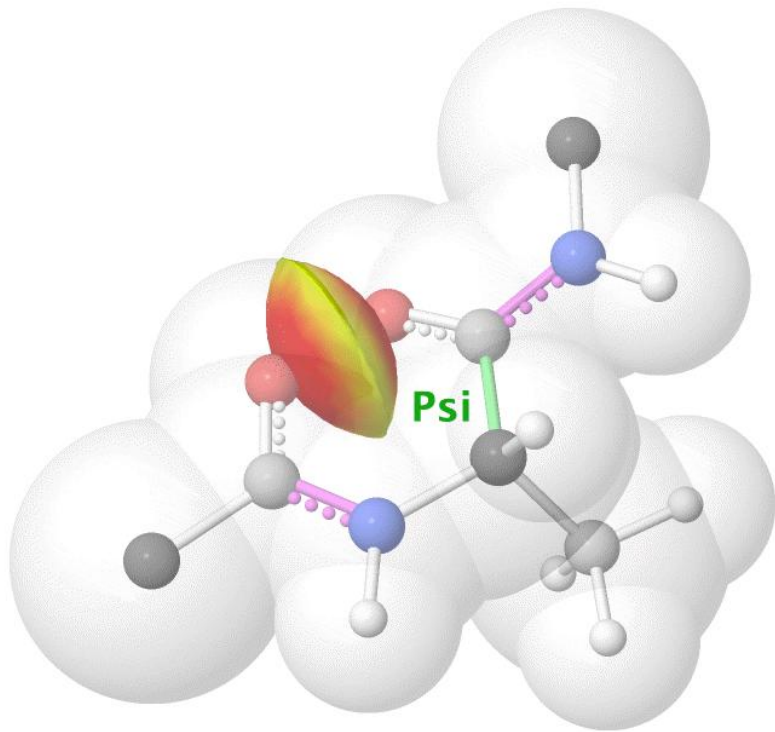
Внутренние координаты белков.



При условии, что ϕ зафиксирован при -80° , большая часть значений угла ψ приводит к возникновению **стерических перекрытий**.

- Кроме $-55 +10$ - Альфа спираль
- Кроме $+90 180$ - Бета лист

Внутренние координаты белков.

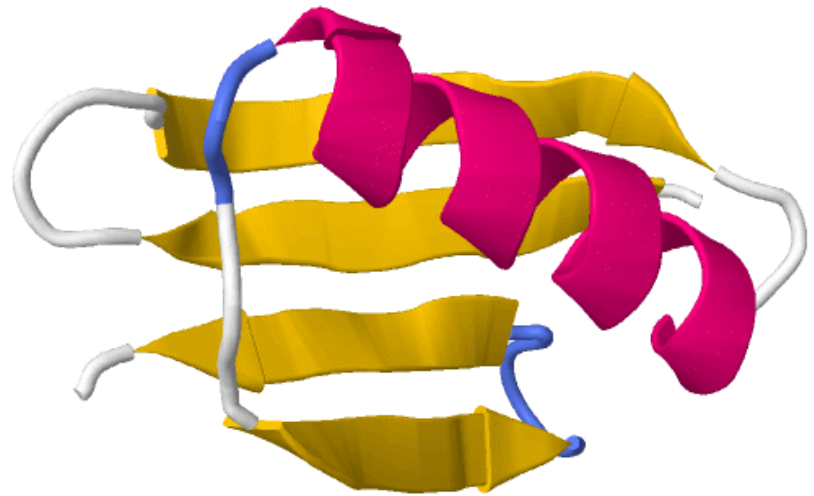


При условии, что ϕ зафиксирован при 0° , все значения угла ψ приводят к возникновению **стерических перекрытий**.

Внутренние координаты белков.

Плотно упакованный белок должен состоять из **альфа спиралей**, **бета слоев** и **поворотов**. Другие виды упаковки маловероятны в силу **внутренних перекрываний атомов**.

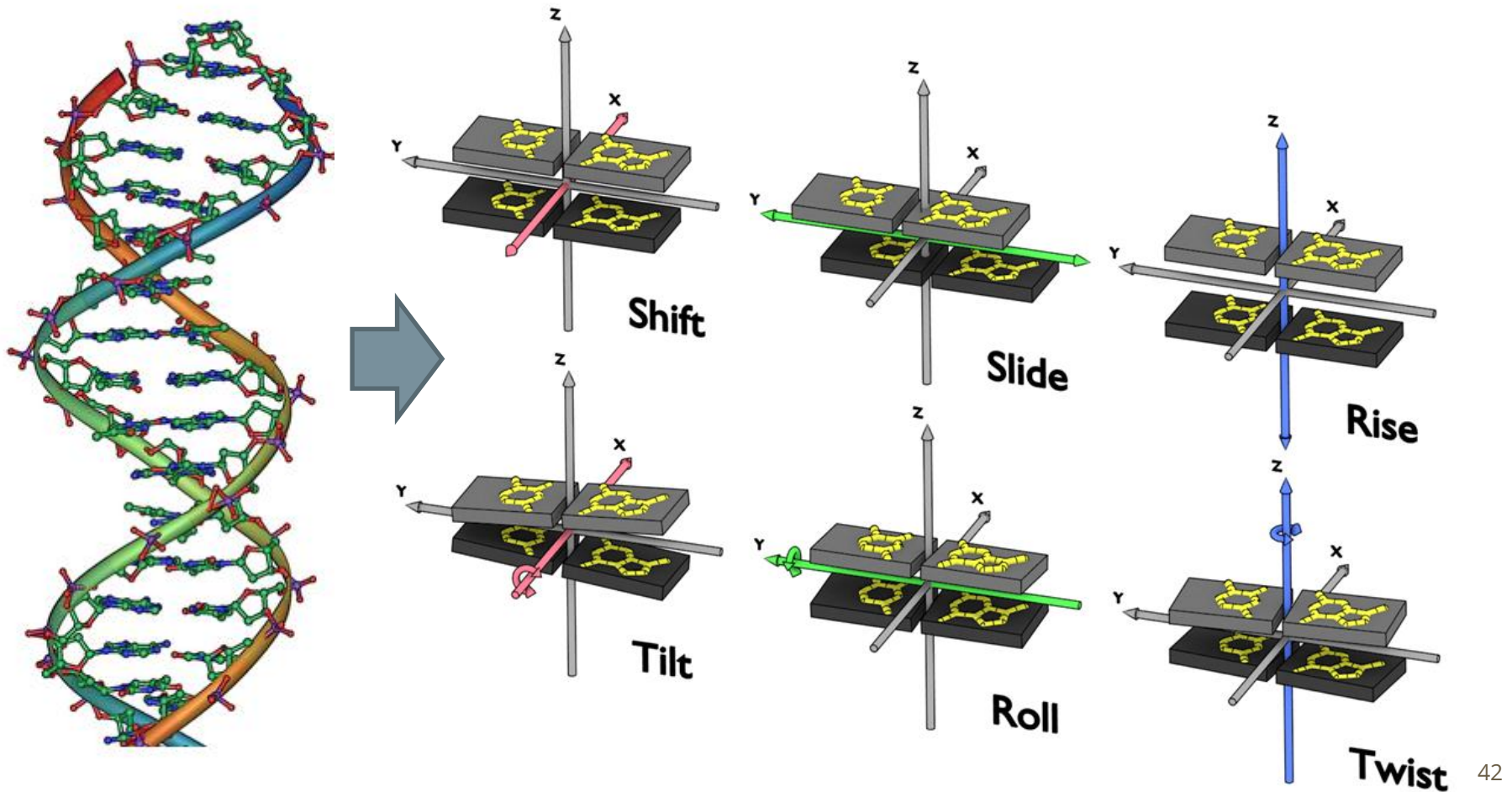
Фи(φ) и **Пси (ψ)**
являются
естественными
координатами для
определения
конфигурации белка.



Jmol

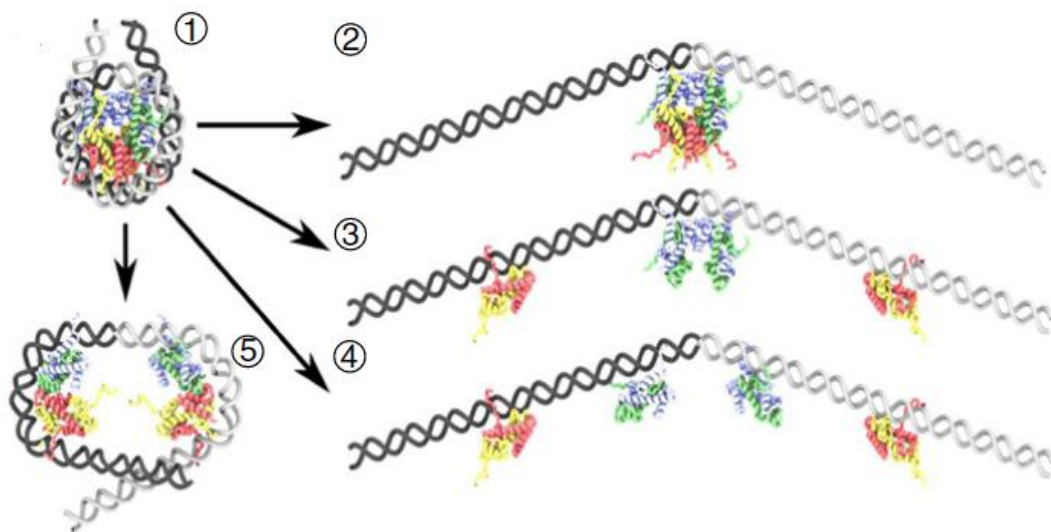
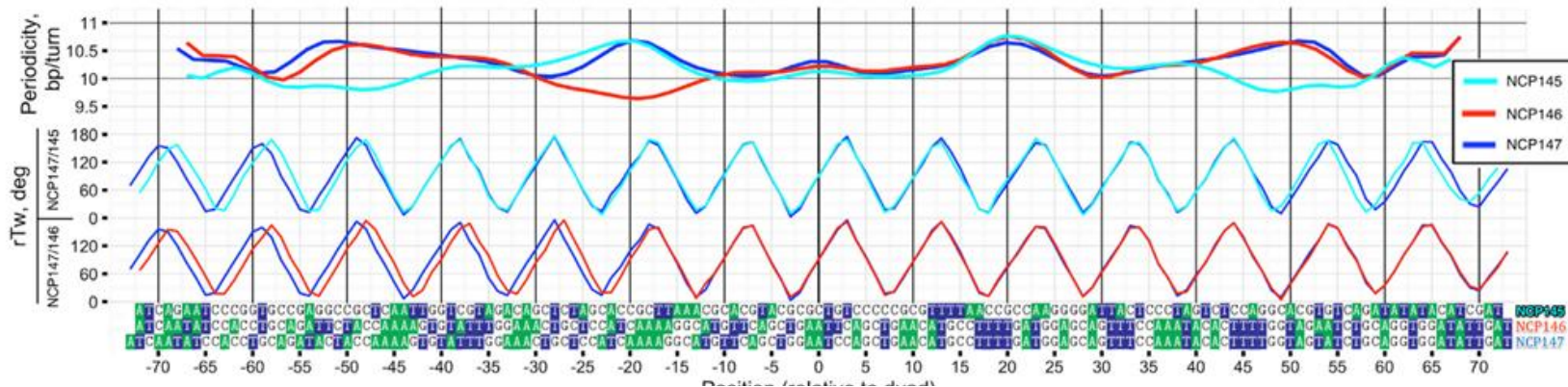
Внутренние координаты ДНК.

Внутренние координаты ДНК позволяют описывать геометрию ДНК при помощи набора из 6 переменных на каждый динуклеотидный шаг.



Внутренние координаты ДНК.

Внутренние координаты ДНК удобно использовать для углубленного анализа геометрии ДНК в комплексах с белками, а также для построения молекулярных моделей неизвестных комплексов ДНК с белками.



Продолжение следует