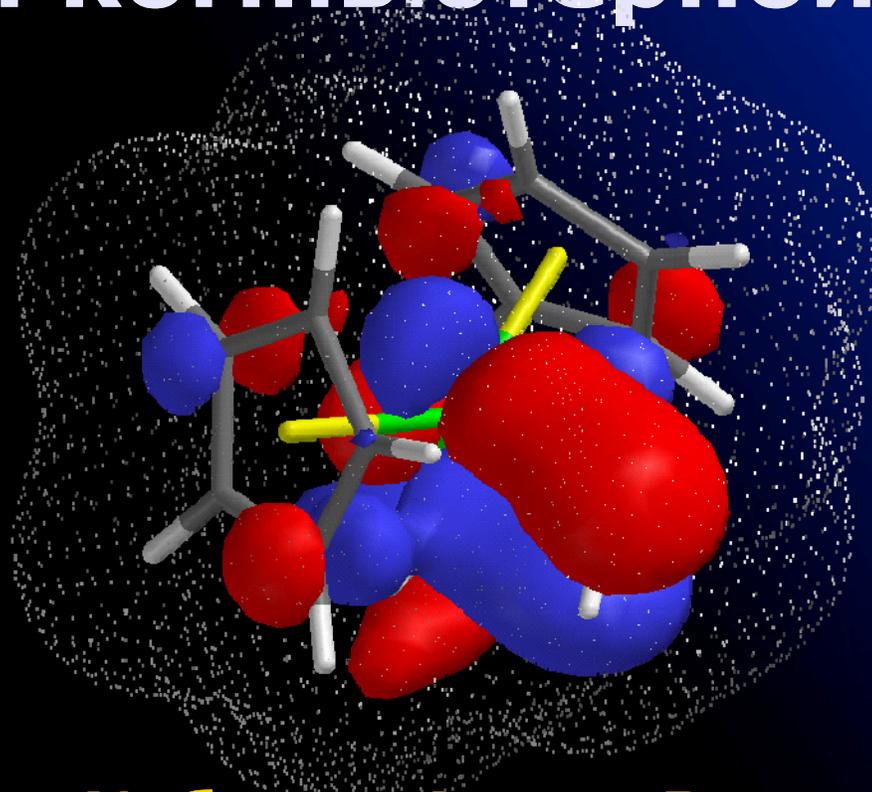


**Мастер-класс**  
**«ОСНОВЫ КОМПЬЮТЕРНОЙ ХИМИИ»**



**к.ф-м.н. Кабанов Артем Валерьевич**  
**к.ф-м.н. Кондратьев Максим Сергеевич**

**(Лаборатория структуры и динамики биомолекулярных систем ИБК РАН)**

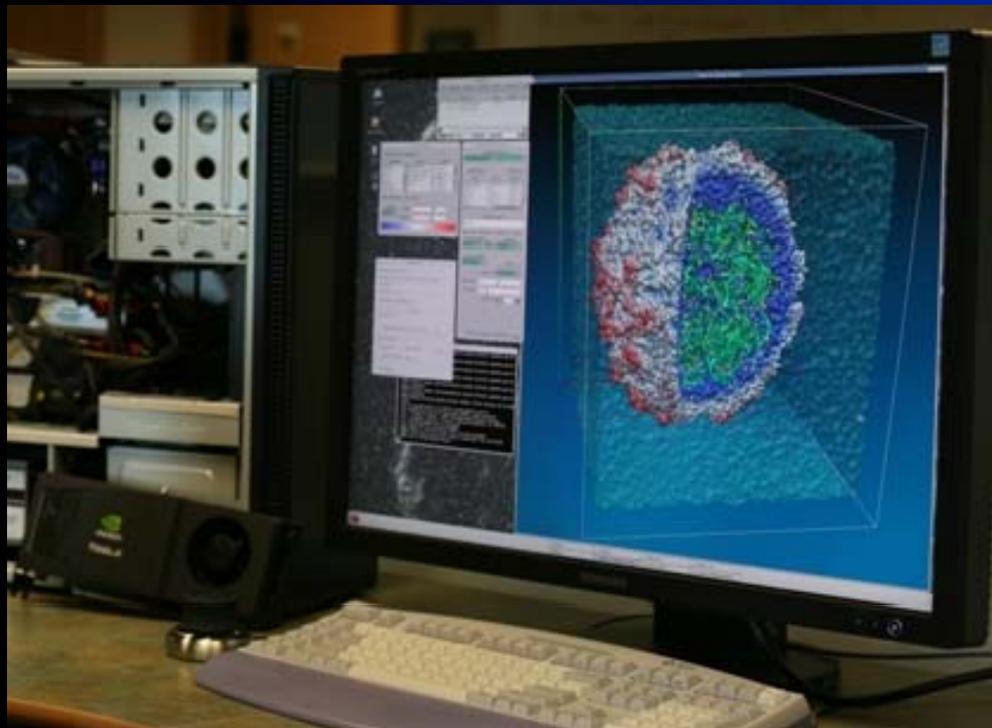
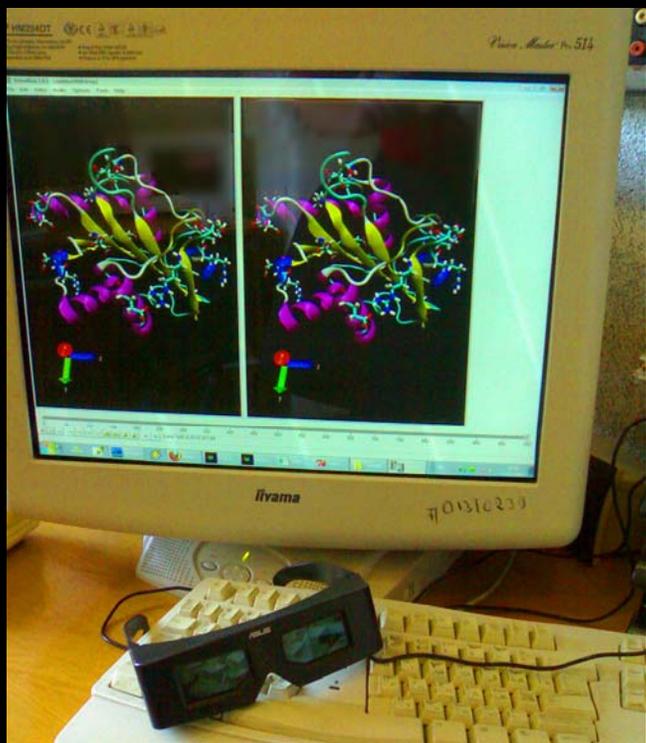
# План:

- 1) Общая информация о возможностях использования компьютеров для расчетов структуры веществ, динамики их молекул, химических реакций, зарядовых и спектральных свойств.**
- 2) Кто придумал компьютерную химию? Экскурс в историю разработки методов. Современные разработчики.**
- 3) Насколько точны теоретические методы? Как теория предвосхитила эксперимент.**
- 4) Практические занятия. Моделирование структуры простых соединений и расчет их свойств.**

# 1. Использование компьютеров для расчетов структуры и свойств биомолекул.

Используются кластеры и профессиональные рабочие станции, а также **настольные суперкомпьютеры** (Linux и Windows).

Основная часть расчетов на них ведется не на ЦП (2, 4, 8 ядер), а на графическом процессоре (до **480** ядер – у GeForce и до **1792** – у Tesla)



## 2. История разработки методов компьютерной химии.

По реализуемым теоретическим подходам, существующие компьютерные методики можно условно разделить на три группы, и расположить их в порядке возрастания строгости приближений и увеличения затратных требований к вычислительным ресурсам:

- Методы эмпирического силового поля (молекулярная механика),
- Полуэмпирические квантово-химические методы,
- Неэмпирические (*ab initio*) квантово-химические методы.

- Квантовая механика, созданная сначала как способ интерпретации атомных спектров, в настоящее время является основой для предсказания и объяснения многих физических свойств молекул, в том числе биомакромолекул.
- В принципе, любую наблюдаемую физическую величину можно вычислить квантовомеханическим путем, если только удастся решить уравнение Шредингера, найти волновую функцию и соответствующую ей энергию.
- Однако, практически невозможно найти точные решения этого уравнения, поэтому приходится искать приближенные методы решения уравнения Шредингера.



**Джон Попл**

**(Sir John Anthony Pople)**

**(31 октября 1925 — 15 марта 2004)**

**Химик-теоретик.**

**Лауреат Нобелевской премии (1998)  
и медали Копли (2002)**

**«за разработку вычислительных  
методов квантовой химии».**



**Майкл Дьюар**

**(Michael Dewar)**

**(24 сентября 1918 - 10 октября 1997)**

**Химик-теоретик.**

**Разработчик принципов минимизации энергии молекулы.**

**Создатель методов MNDO и AM1.**

**Являлся членом Международной Академии квантовой молекулярной науки**



**Джеймс Стюарт**  
**(James Stewart)**

**Химик-теоретик. Программист.  
Разработчик метода РМЗ и РМБ.  
Автор пакета МОРАС**

**<http://openmorac.net>**

# MOPAC

## Molecular Orbital PACkage

- **MOPAC timelines:**

1977: MNDO - M. Dewar & W. Thiel

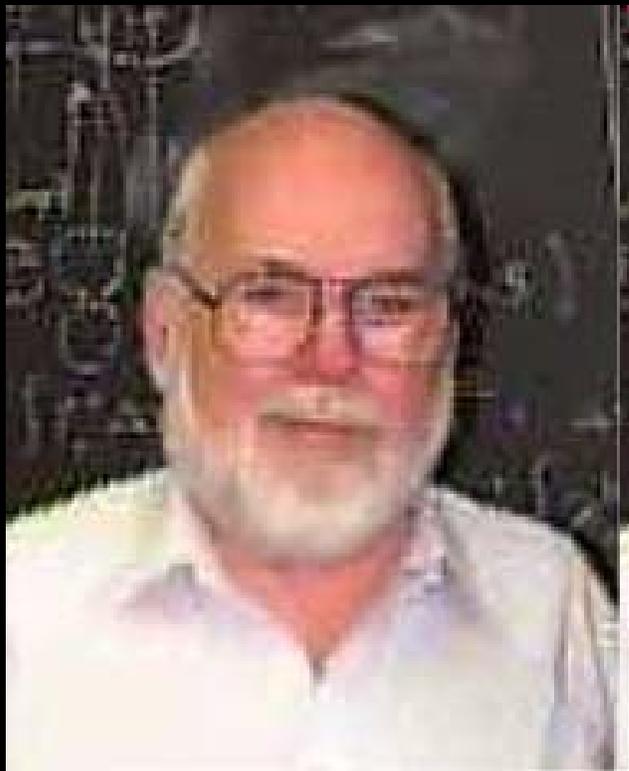
1981: MOPAC project started at Univ. of Texas, Austin by J.J.P. Stewart

1985: AM1 (Austin Method 1) - M. Dewar et al

1989: PM3 (Parametric Method 3) - J.J.P. Stewart

1996: MNDO-d (T-metals added) - W. Thiel & A.A. Voityuk

2007: PM6 (all elements) - J.J.P. Stewart



## **Роберт Парр**

**Химик-теоретик.**

**Разработчик квантово-химических методов *ab initio*, в том числе DFT.**

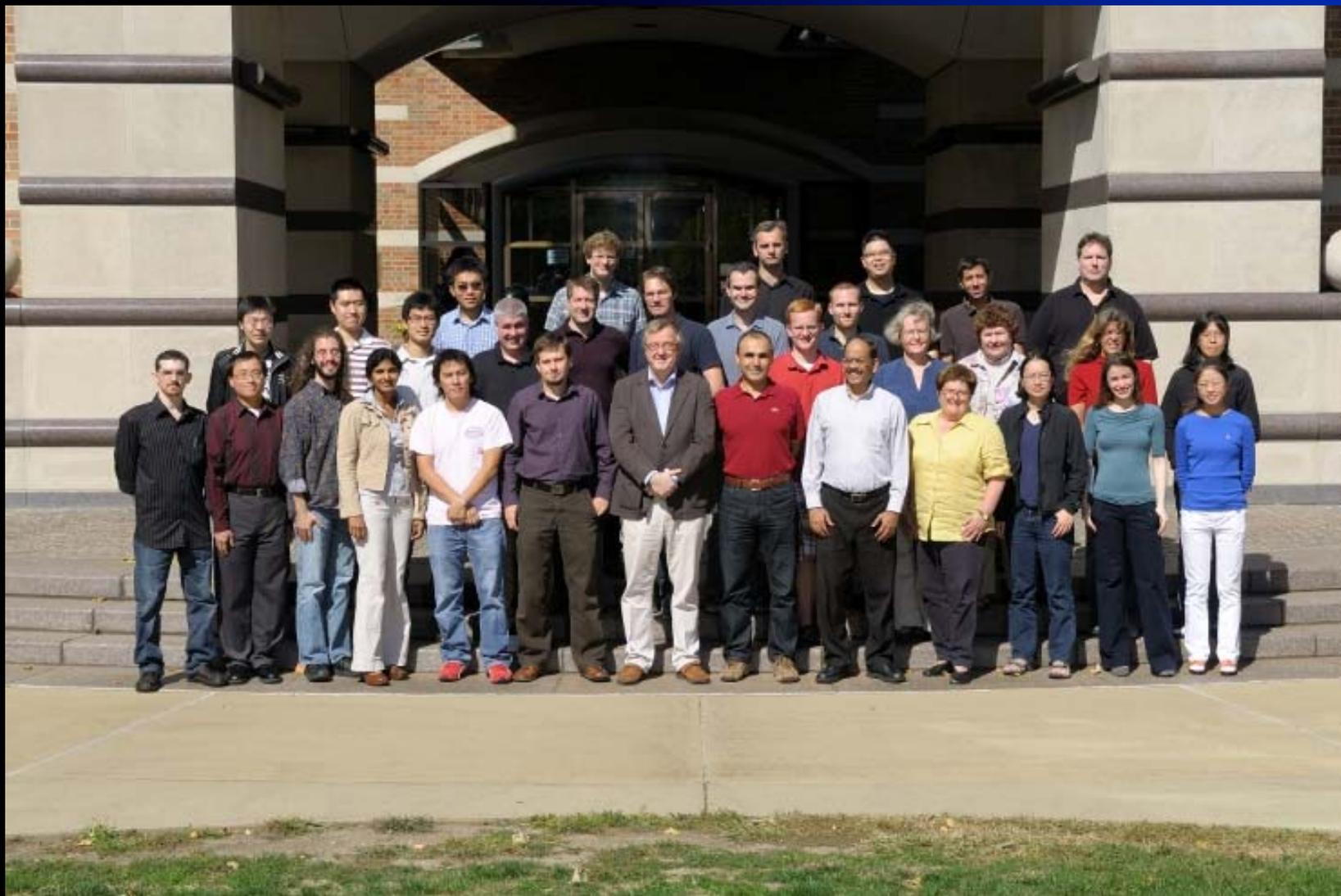
# Метод молекулярной динамики

**Метод молекулярной динамики (MD) — метод, в котором временная эволюция системы взаимодействующих атомов или частиц отслеживается интегрированием их уравнений движения.**

**Комбинацию потенциальных функций, лежащих в основе метода молекулярной динамики и описывающих растяжение связей, деформацию валентных углов, торсионное взаимодействие, называют валентным силовым полем.**

**Это поле соответствует свойствам, приписываемым химическим связям в молекуле. Для выполнения более точных количественных расчетов валентное силовое поле должно включать в себя также функции, отвечающие стерическим взаимодействиям.**

# Разработчики NAMD



США, штат Иллинойс (Университет Урбана-Шампейн)

# Разработчики Gaussian



США, штат Коннектикут

# **Российские разработчики методов компьютерной химии**

**МГУ им.Ломоносова**

**(К.В.Шайтан, А.В.Немухин)**

**Институт математических проблем биологии РАН**

**(В.Д.Лахно, Н.К.Балабаев, А.В.Теплухин)**

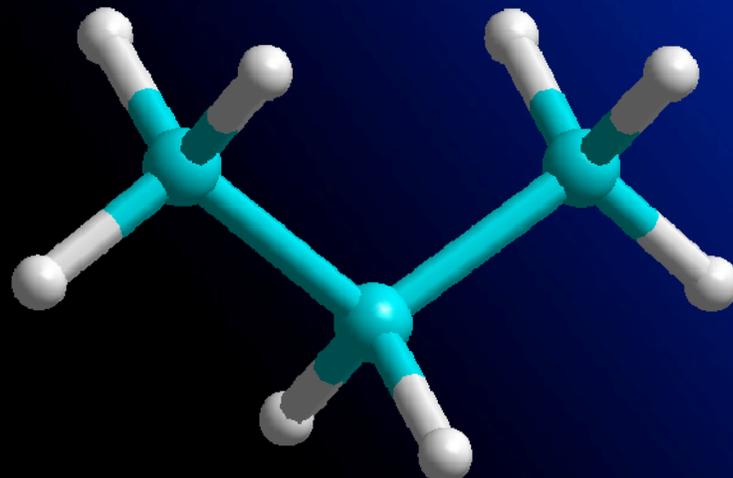
**Институт биологии Карельского научного центра РАН**

**(В.А.Иванов, А.Л.Рабинович)**

**НИИ Биомедицинской химии им. В.Н. Ореховича**

**(В.В.Поройков и соавт.)**

### 3. Точность методов. Связь с экспериментами.



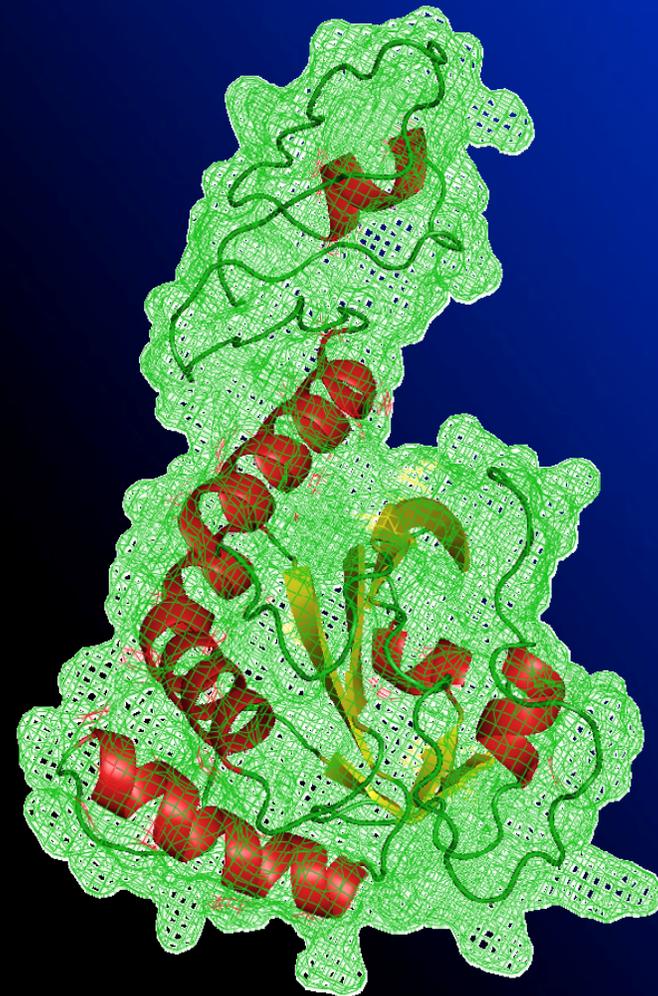
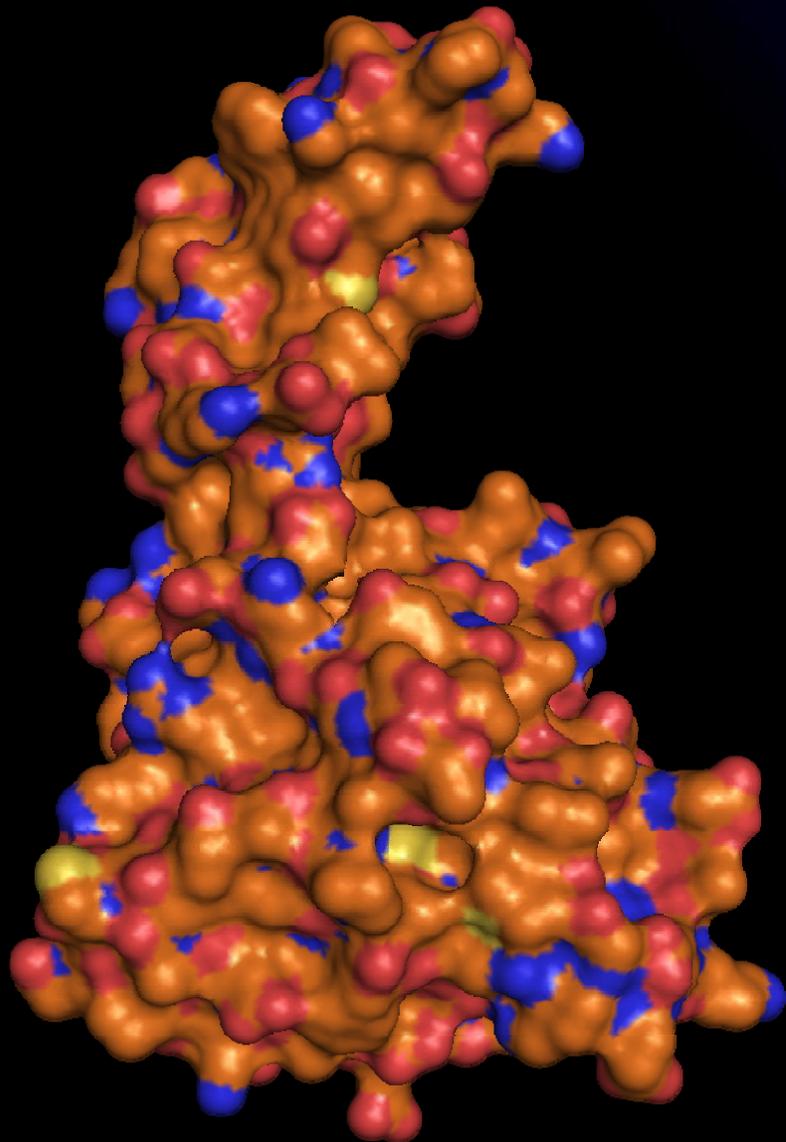
	MM2	MINDO/3	MNDO	3-21G	6-31G(d)	Эксперимент
Время расчета, с	0,83	9,75	10,32	550	4702	----
$R_{CC}$ , Å	1,534	1,495	1,530	1,541	1,528	1,526
$\angle CCC$ , °	111,7	121,5	115,4	111,6	112,7	112,4
$\Delta H_f^0$ , ккал/моль	-24,8	-26,5	-24,9	--	--	-25,0

**Затрачиваемое машинное время и результаты расчетов характеристик молекулы пропана, проведенных различными методами.**

# Программные пакеты для компьютерной химии

- **Hyperchem**
- **Avogadro**
  
- **Zephyr (GROMACS)**
- **NAMD+VMD**
- **AMBER**
  
- **MOPAC**
- **Gaussian**
  
- **HEX**
- **Coot**

**Опыт моделирования  
структурно-динамических особенностей  
малых белков с помощью настольных  
суперкомпьютеров**



# Методы

## Квантово-химические:

полуэмпирические (PM3, AM1)  
неэмпирический (MP2/6-311++G\*\*)

## Молекулярная динамика

в различных вариантах силовых полей  
(AMBER99, CHARMM27, OPLS-AA)

# Часть I. Расчеты в рамках фундаментальных исследований

1. Моделирование ранних стадий фолдинга белков  
(исследование процесса инициирования альфа-спирализации)
2. Определение предпочтительных конформационных состояний  
междоменного линкера альфа-субъединицы  
РНК-полимеразы *E.coli*

-Arg1-Asp2-Val3-Arg4-Gln5-Pro6-Glu7-Val8-Lys9-Glu10-Glu11-Lys12-Pro13-Glu14-Phe15-

# Предпочтительная локализация кислых и основных аминокислотных остатков в белках: в кэп(«сар»)-зонах $\alpha$ -спиралей

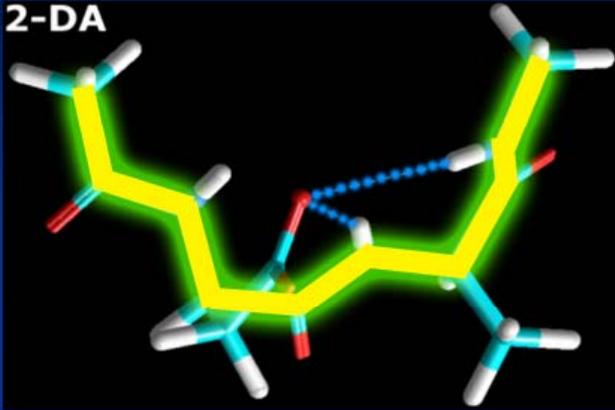
А.к. ост.	Наличие		число $\gamma$	Диполь/заряд	рК	Яркая тенденция быть в:							
	NH	C $\beta$				до		спираль		за		клубок	ядро
						$\alpha_N$	$ \alpha_N $	$\alpha$	$ \alpha_C $	$\alpha_C$	$\beta$		
Gly	✓							-		-	+		
Ala	✓	✓						+			-		
Pro		✓	1				+	-	-	-	-	+	
Glu	✓	✓	1	COOH $\Rightarrow$ CO $_2^-$	4.3	+	+			-	-	-	
Asp	✓	✓	1	COOH $\Rightarrow$ CO $_2^-$	3.9	+	+	-	-	-	-	+	-
Gln	✓	✓	1	OCNH $_2$									-
Asn	✓	✓	1	OCNH $_2$		+		-		+	-	+	-
Ser	✓	✓	1	OH		+						+	
His	✓	✓	1	NH, и N $\Rightarrow$ NH $^+$	6.5		-		+	+			
Lys	✓	✓	1	NH $_2 \Rightarrow$ NH $_3^+$	10.5	-	-		+	+	-		-
Arg	✓	✓	1	HNC(NH $_2$ ) $_2^+$	12.5	-	-		+	+	-	+	-
Thr	✓	✓	2	OH		+					+		
Ile	✓	✓	2								+	-	+
Val	✓	✓	2								+	-	+
Leu	✓	✓	1						+		+	-	+
Met	✓	✓	1						+		+	-	+
Phe	✓	✓	1								+	-	+
Tyr	✓	✓	1	OH $\Rightarrow$ O $^-$	10.1						+		+
Cys	✓	✓	1	SH $\Rightarrow$ S $^-$	9.2						+		+
Trp	✓	✓	1	NH							+		+

# Объем выполненных расчетов

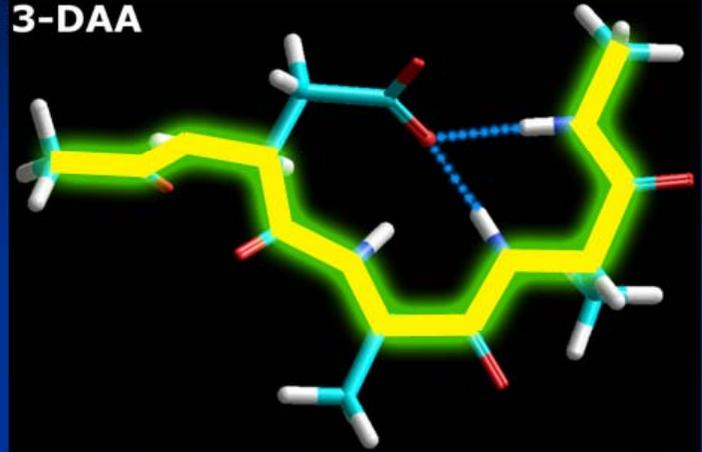
Название объекта	Методы	Число изученных структур
Изолированные аминокислоты	PM3, AM1, MP2/6-311++G**	160 конформеров. 20 а.к. - по 8 на каждую.
Цвиттер-ионы аминокислот	PM3 MP2/6-311++G**	40 конформеров. 20 а.к. - по 2 на каждую.
Цвиттер-ионы + H <sub>2</sub> O	PM3	20 молекул, добавляя от 1 до 14 молекул H <sub>2</sub> O.
Монопептиды (метиламиды N-ацетил- $\alpha$ -L-аминокислот)	PM3, AM1, MM3Pro, MP2/6-311++G**	20 молекул. От 14 до 5800 конформеров.
Дипептиды, трипептиды, тетрапептиды, пентапептиды	PM3	10 молекул. От 3 до 20 конформеров.
8-членные гомо- и гетерогенные цепочки а.к. (октапептиды)	OPLS-AA PM3	20 молекул. Для 10 построены PM3-профили спирализации.
12-членные гомо- и гетерогенные цепочки а.к.	OPLS-AA	20 молекул.
16-членный полиаланин с аспаратом на N-конце	AMBER, CHARMM, OPLS-AA	5 конформеров.
Октапептиды в водной (TIP3P) среде	AMBER OPLS-AA	Три последовательности: DAAAAAAAA, EAAAAAAAA и AAAAAAAAAH.

# Механизм образования первого витка альфа-спирали. Образование бифуркационных водородных связей

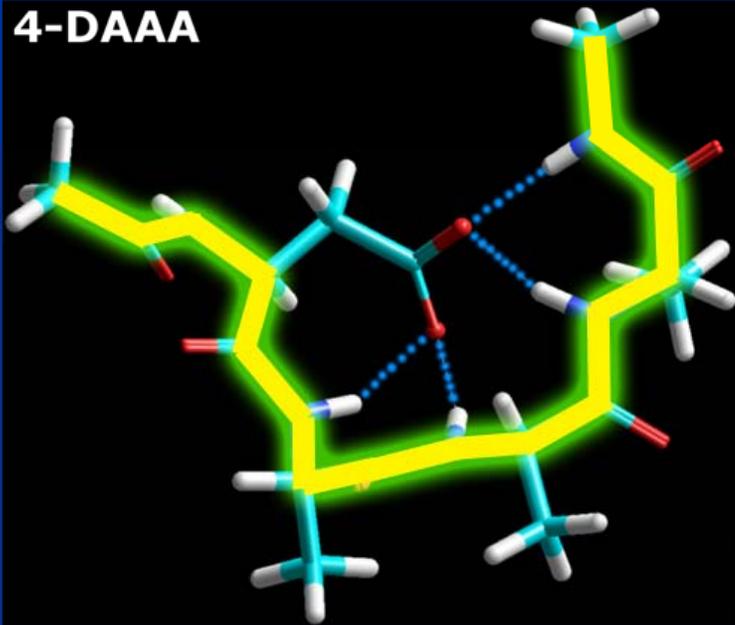
2-DA



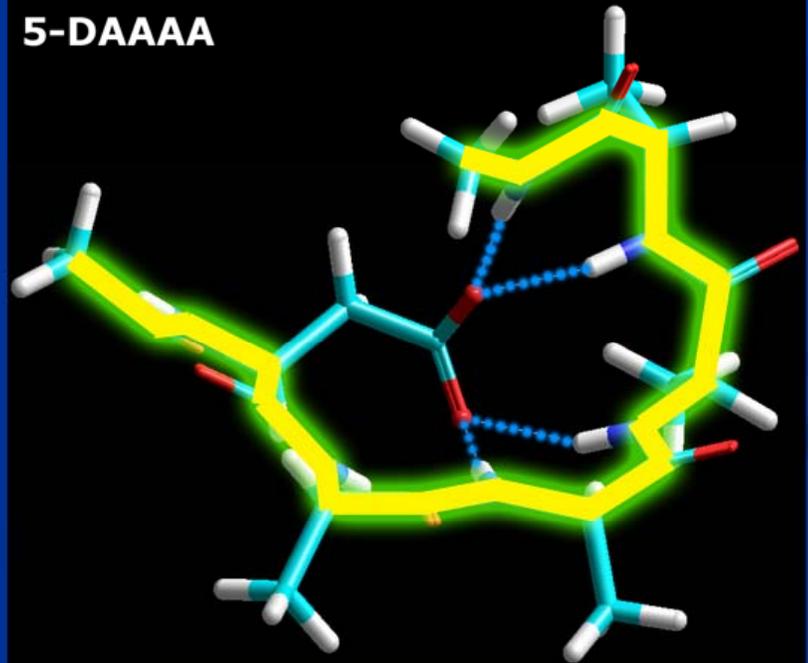
3-DAA



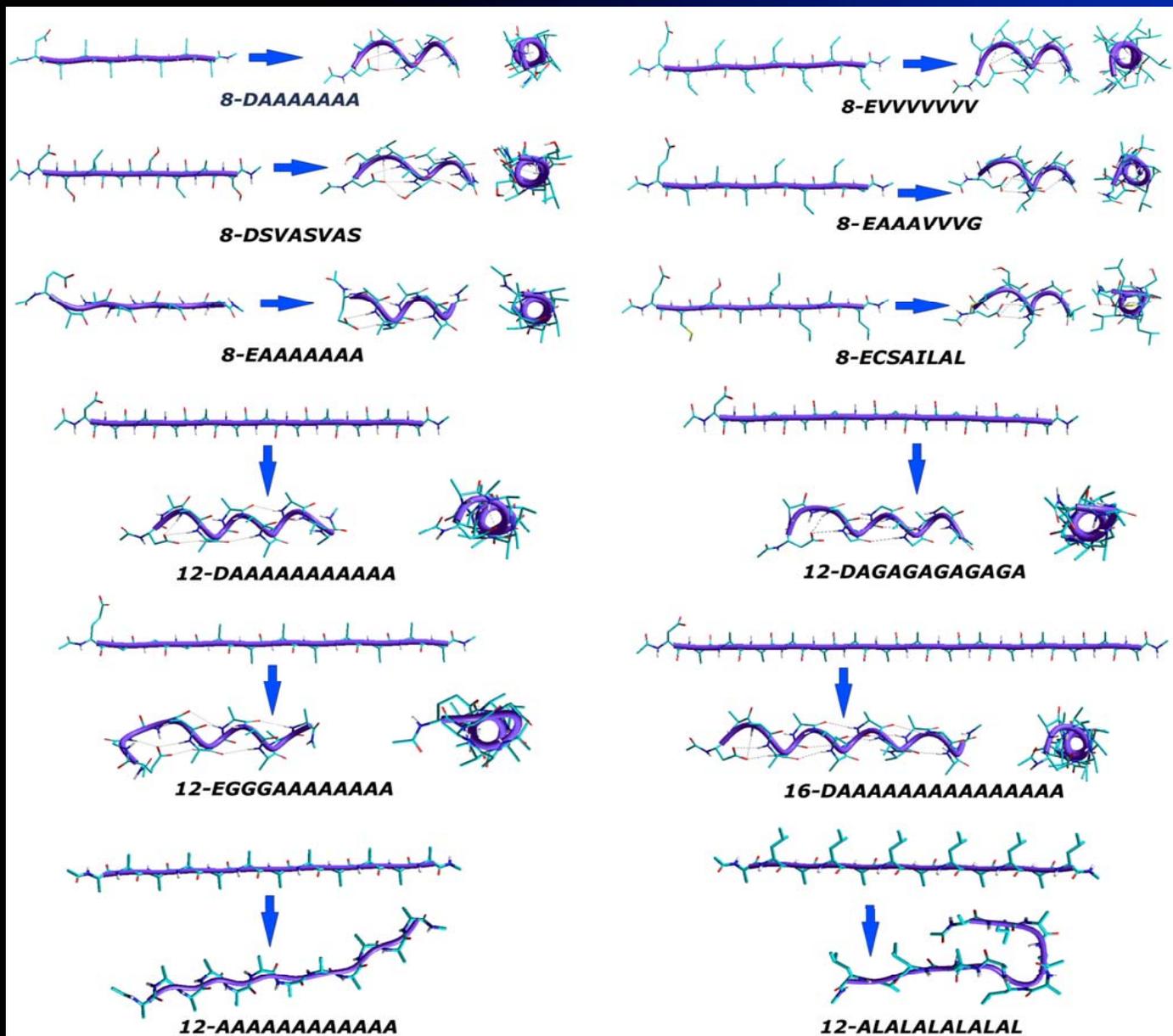
4-DAAA



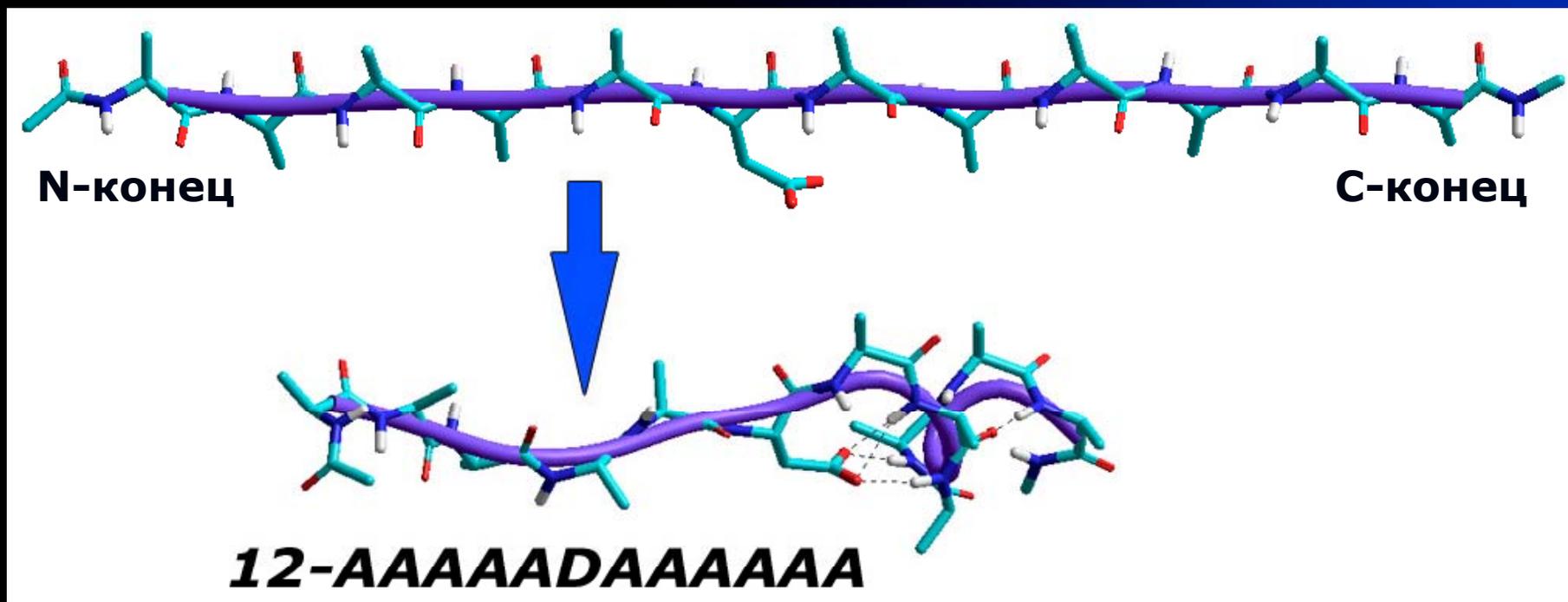
5-DAAAA



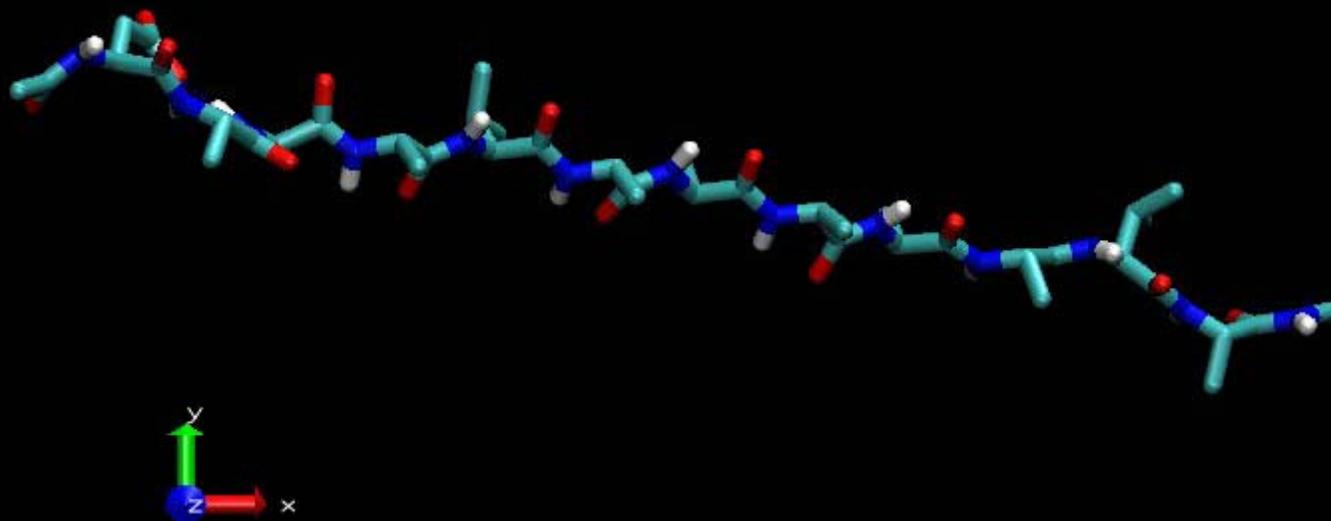
# MD-результаты спирализации олигопептидов при наличии или отсутствии остатков аспартата (D) или глутамата (E)



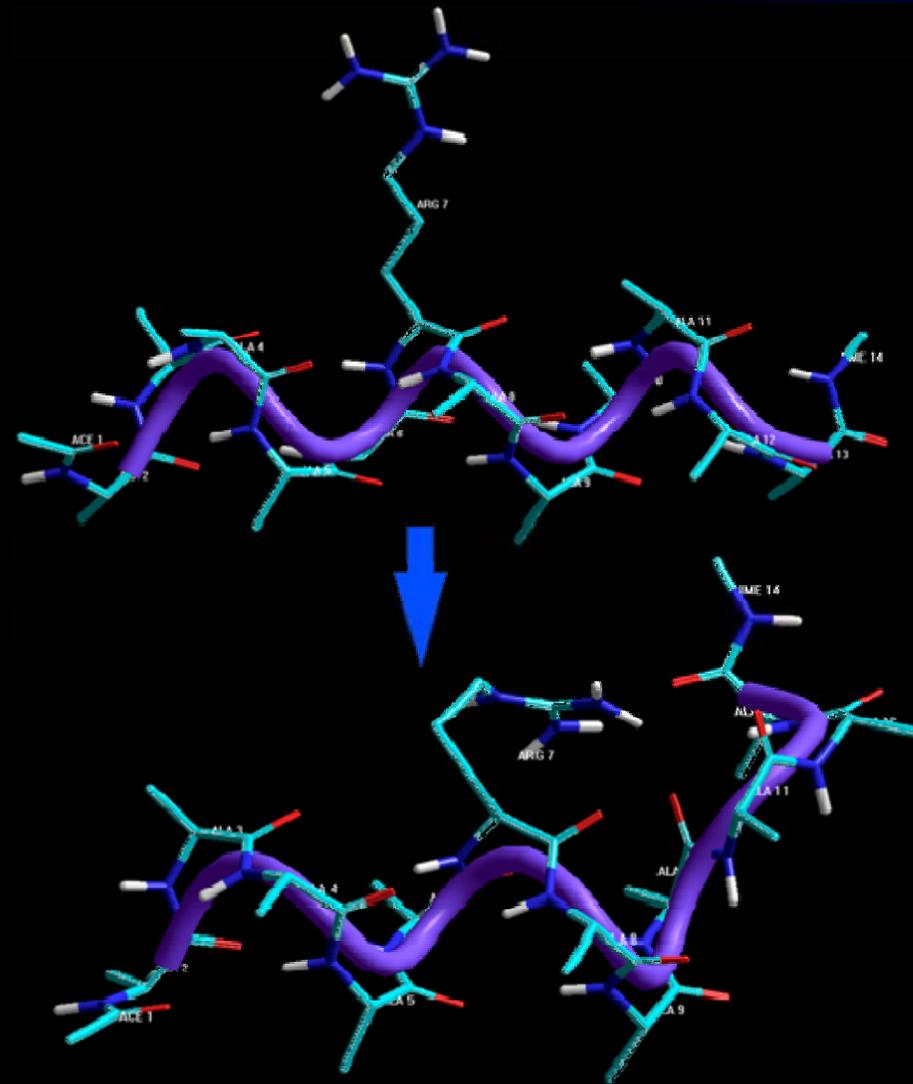
**Локализация спирального участка олигопептида зависит от местоположения иницирующего аминокислотного остатка**



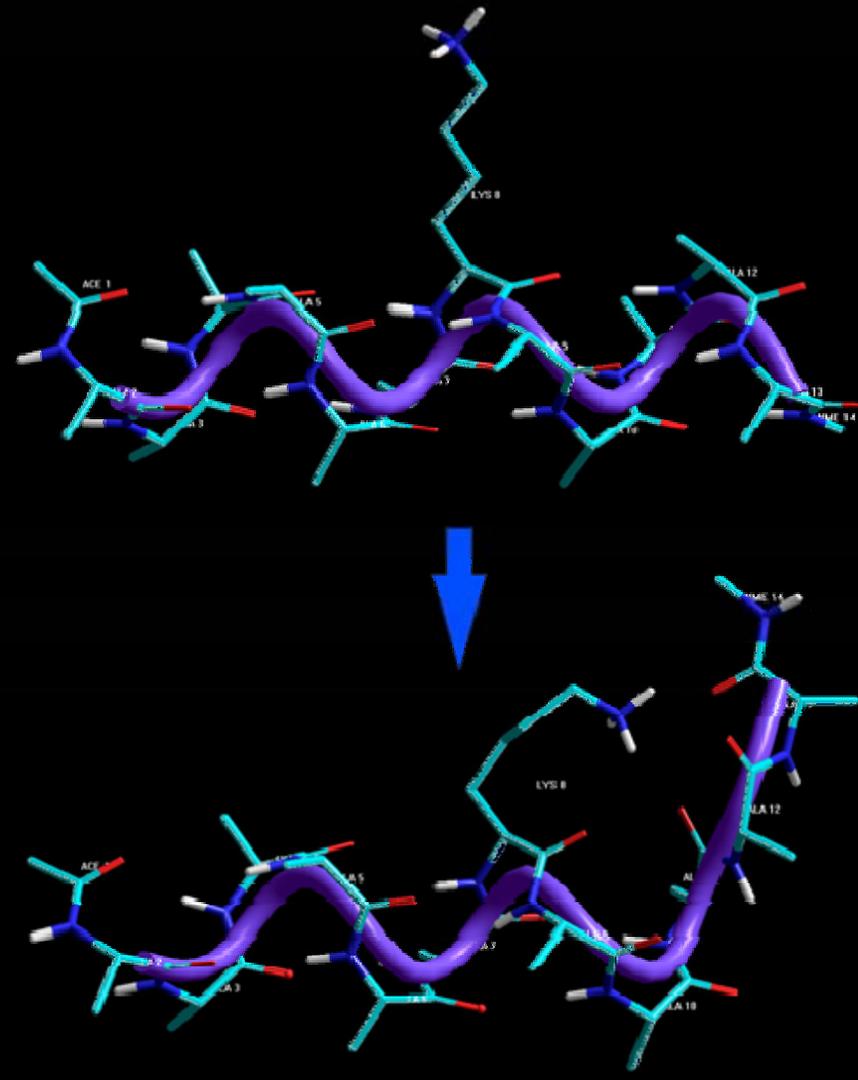
## Спирализация пептида, запускаемая на N-конце последовательности



**Терминирование спирализации последовательностей может осуществляться заряженными основными остатками (на примере аргинина и лизина)**



**12-A-A-A-A-A-R-A-A-A-A**



**12-A-A-A-A-A-K-A-A-A-A**

# Два типа спирализации?

**Прямая:**

направление синтеза белка

направление спирализации

**N-конец**

**C-конец**

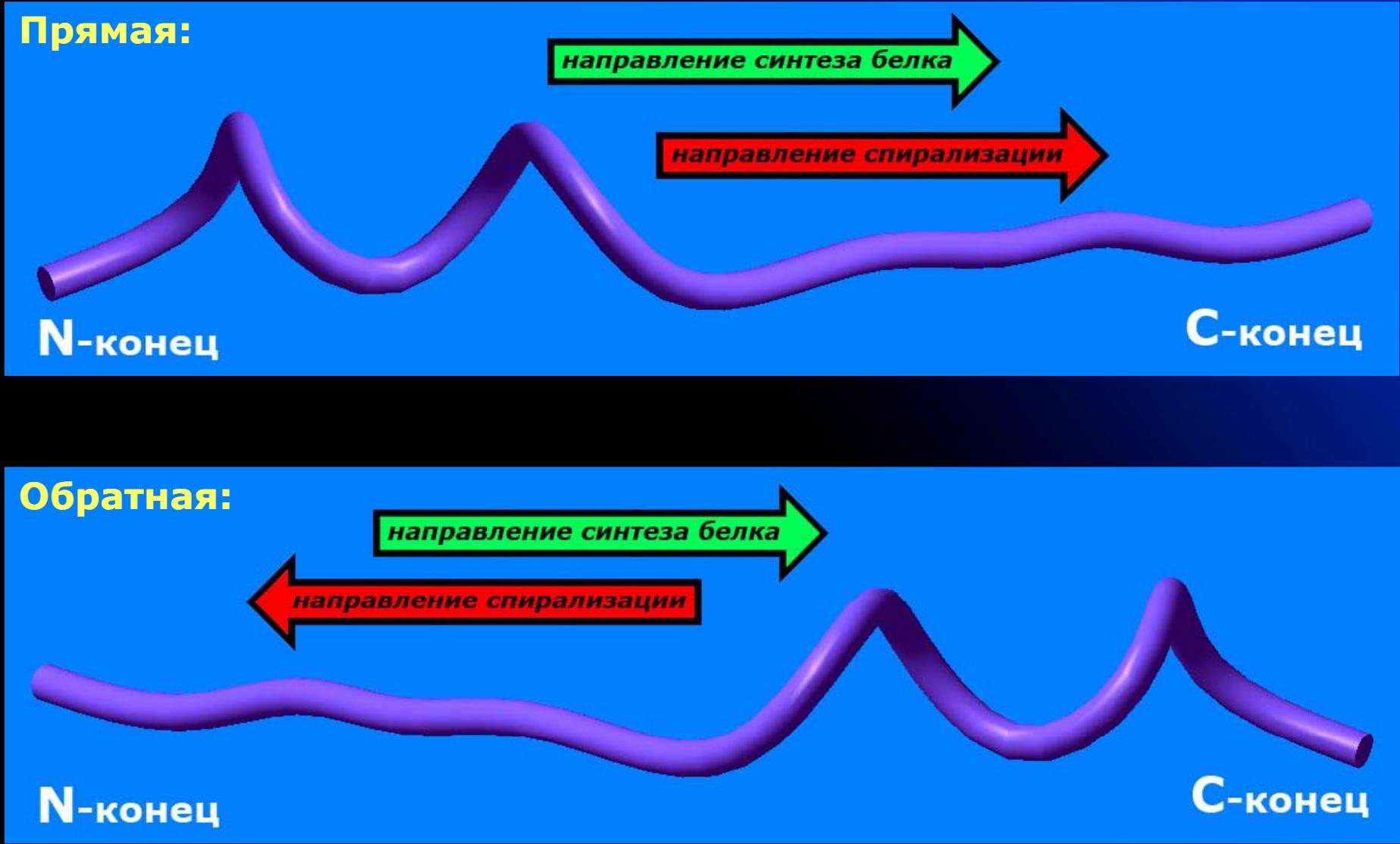
**Обратная:**

направление синтеза белка

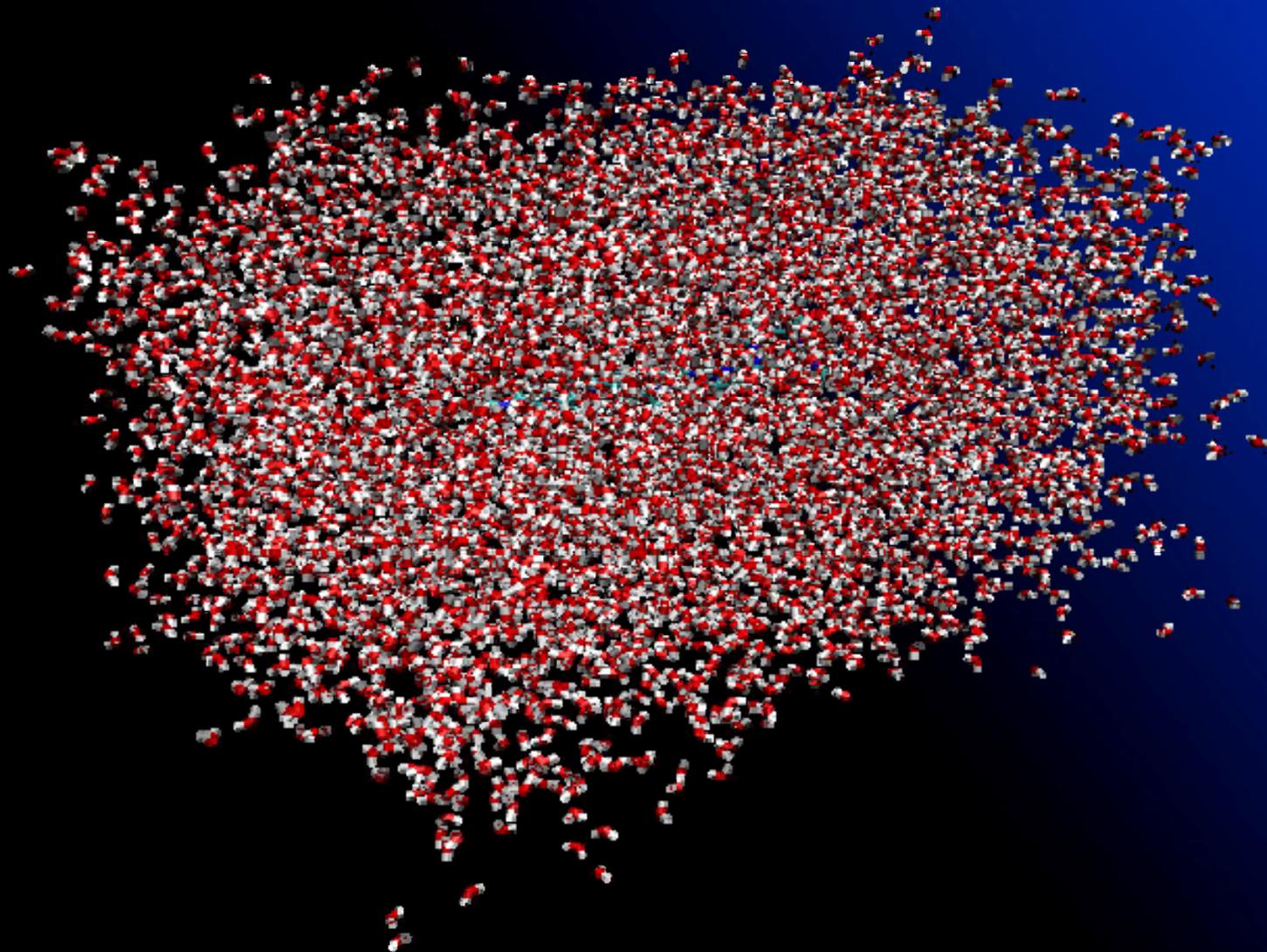
направление спирализации

**N-конец**

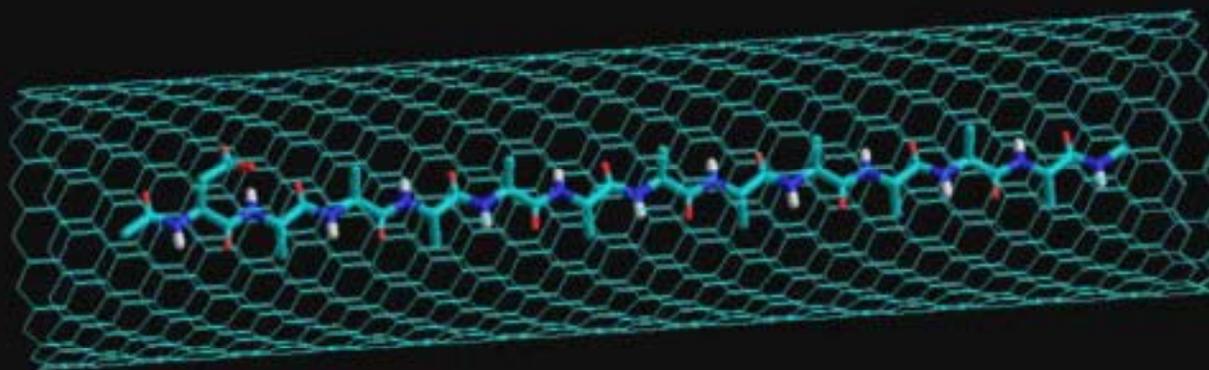
**C-конец**



**Оптимизированная структура комплекса  
«DAAAAAAAA (полиаланин с аспартатом на N-конце) + 6237 H<sub>2</sub>O»**



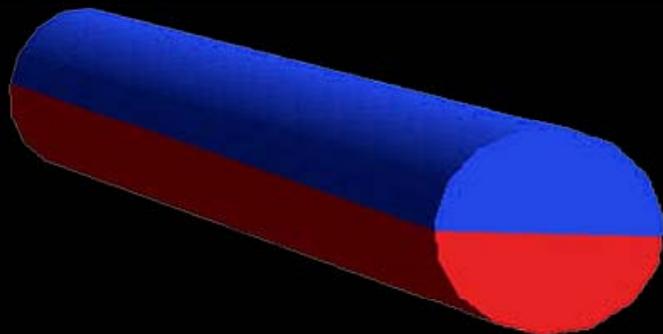
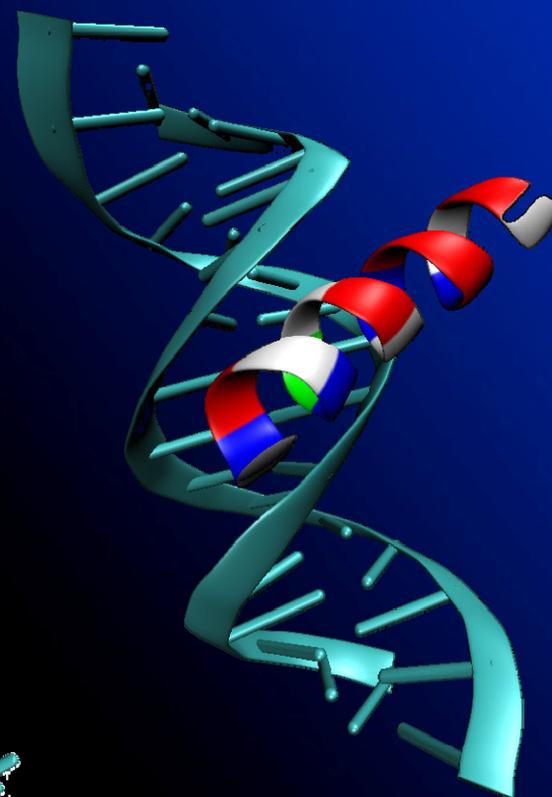
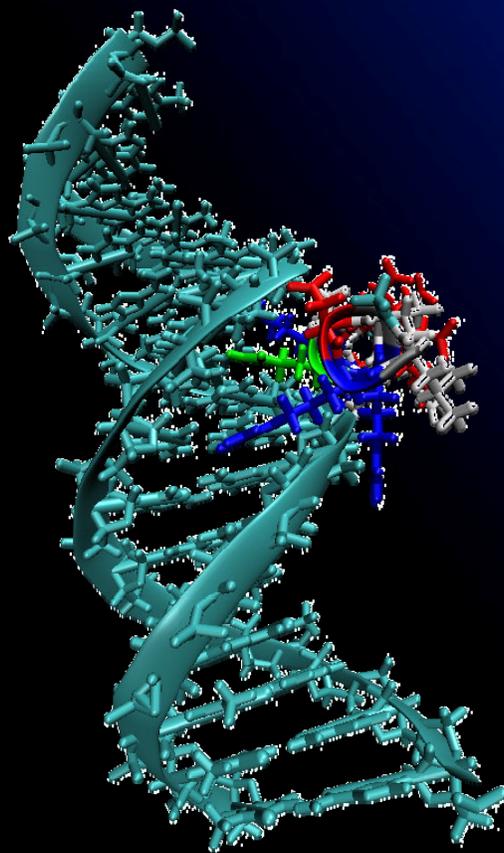
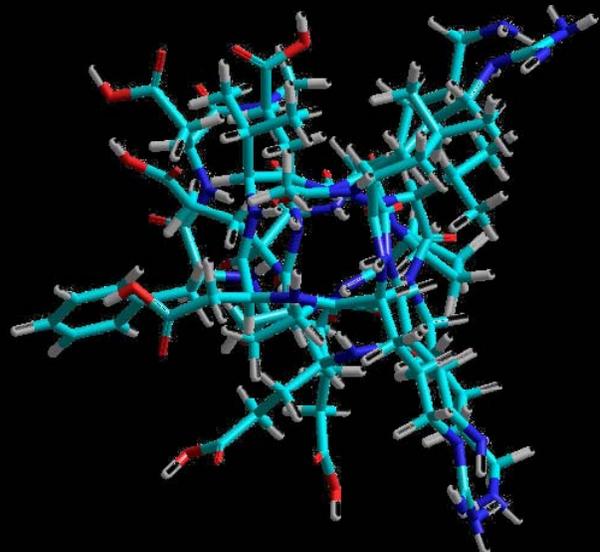
# Образование альфа-спирали внутри углеродной нанотрубки (моделирование рибосомального канала)



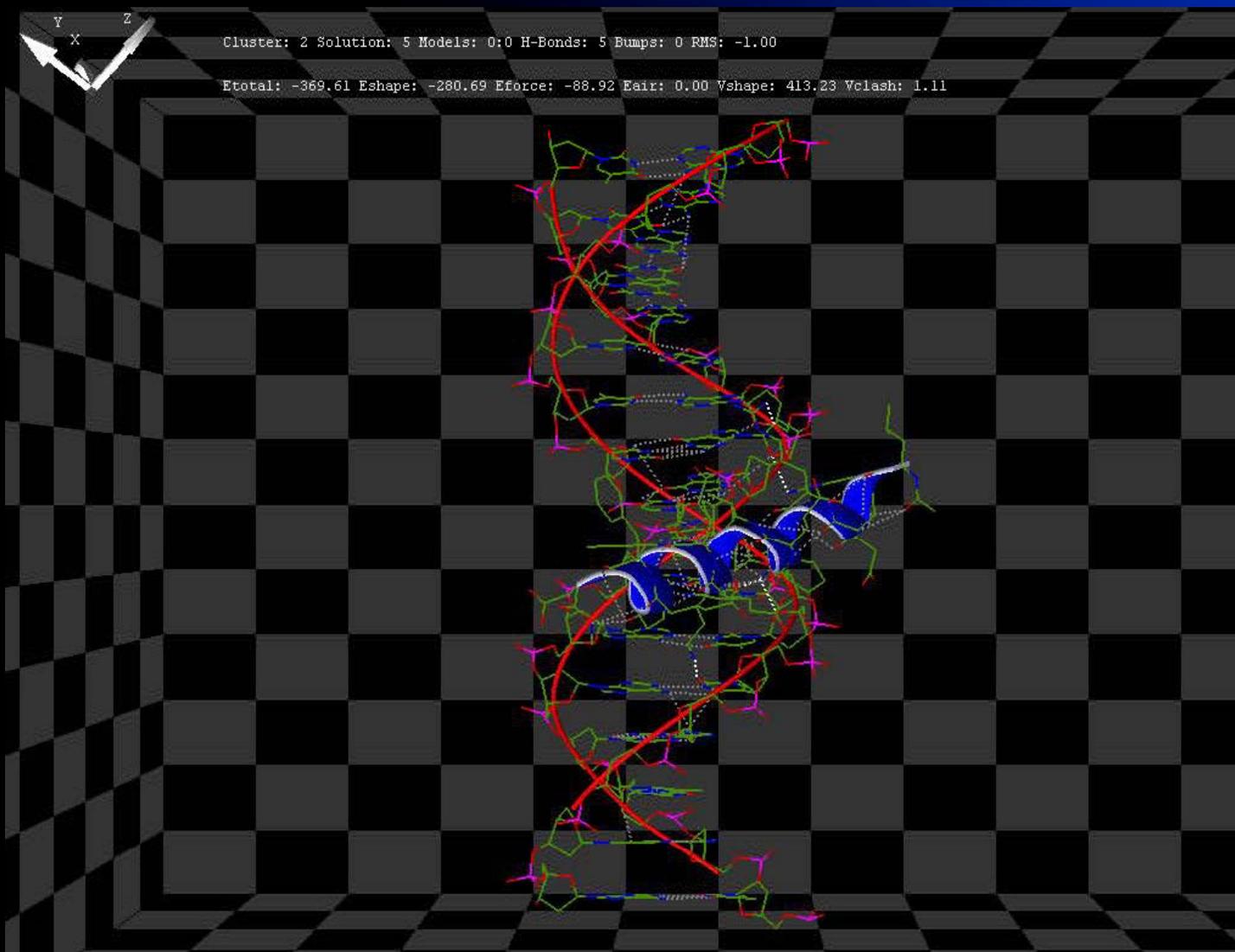
# Междоменный пептидный линкер может быть альфа-спиралью?

-Arg1-Asp2-Val3-Arg4-Gln5-Pro6-Glu7-Val8-Lys9-Glu10-Glu11-Lys12-Pro13-Glu14-Phe15-

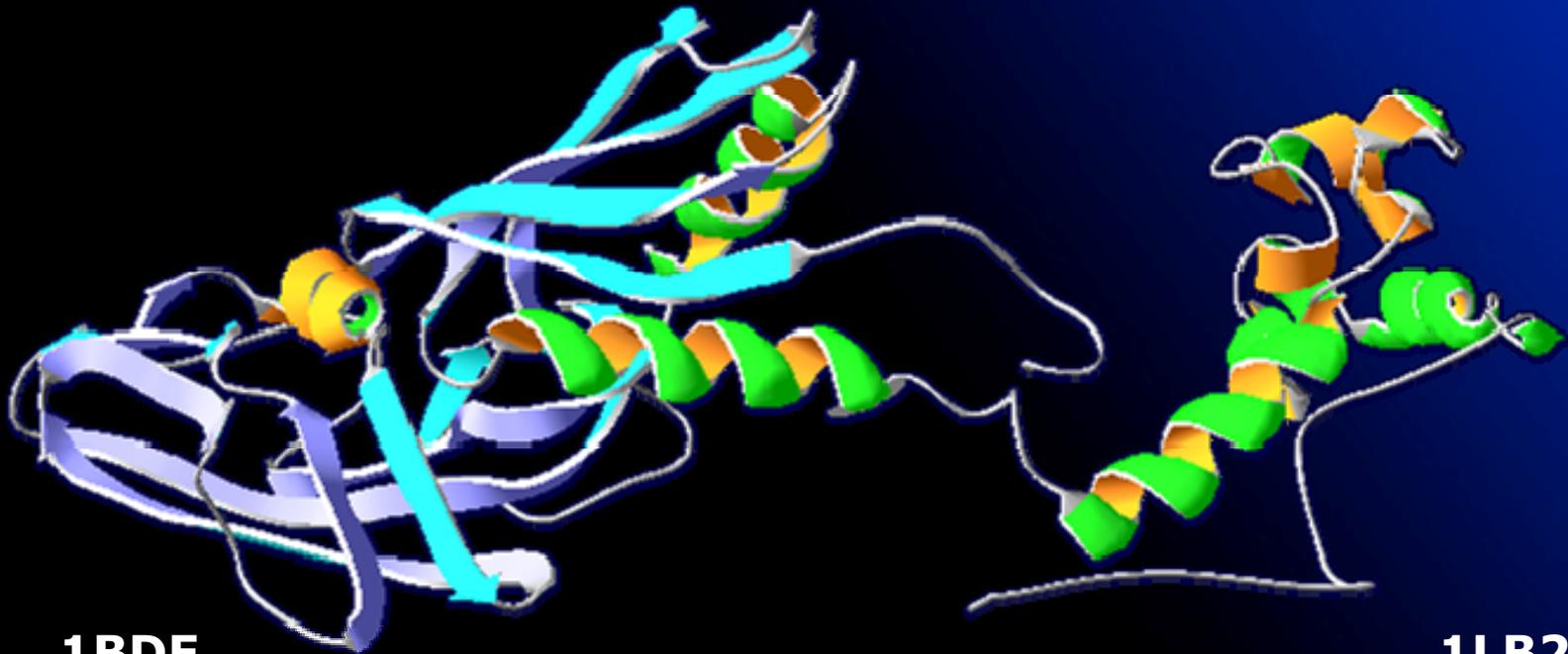
Конформер	Теплота образования	Дипольный момент	HOMO-LUMO
1. Альфа-спираль	-1371	38,2	8,15
2. Развернутая цепь	-1378	13,8	9,13
3. Смешанный тип	-1375	9,3	9,08



# Расчет «жесткого» докинга спирального линкера на модельной ДНК



# Реконструкция структуры альфа-субъединицы РНК-полимеразы



**1BDF**

**1LB2**

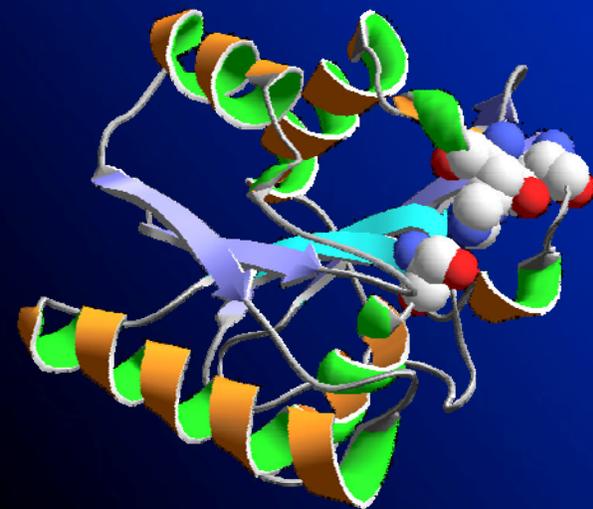
## Часть II.

# Расчеты в рамках исследований, имеющих прикладное значение (повышение термостабильности белков)

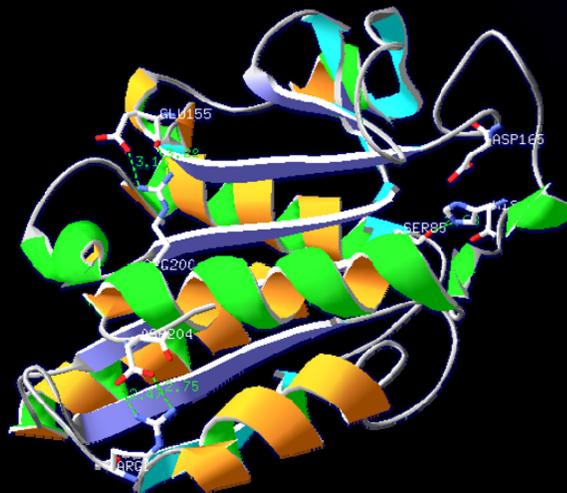
Пероксиредоксин 6



Липаза LipA



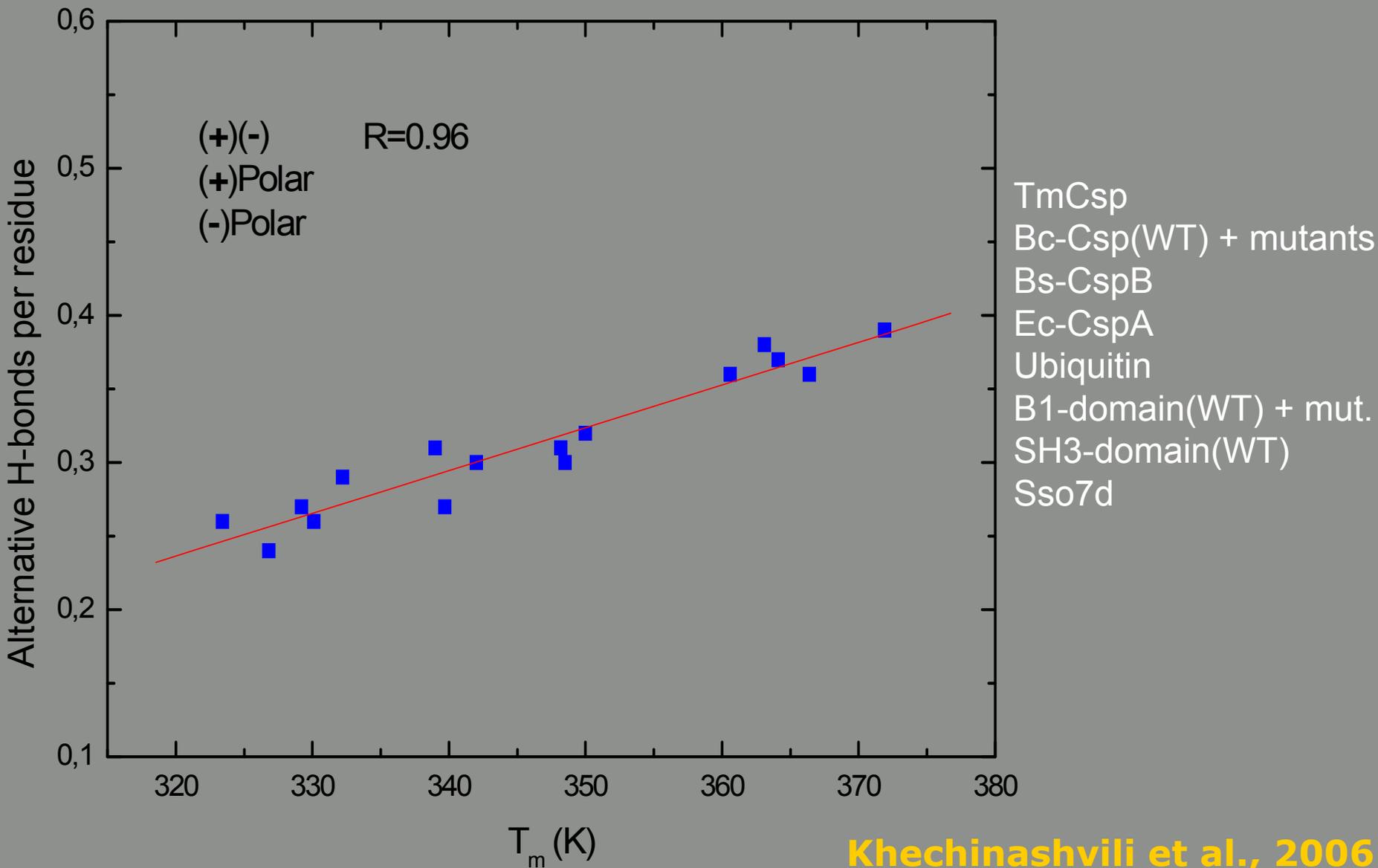
Липаза CLE



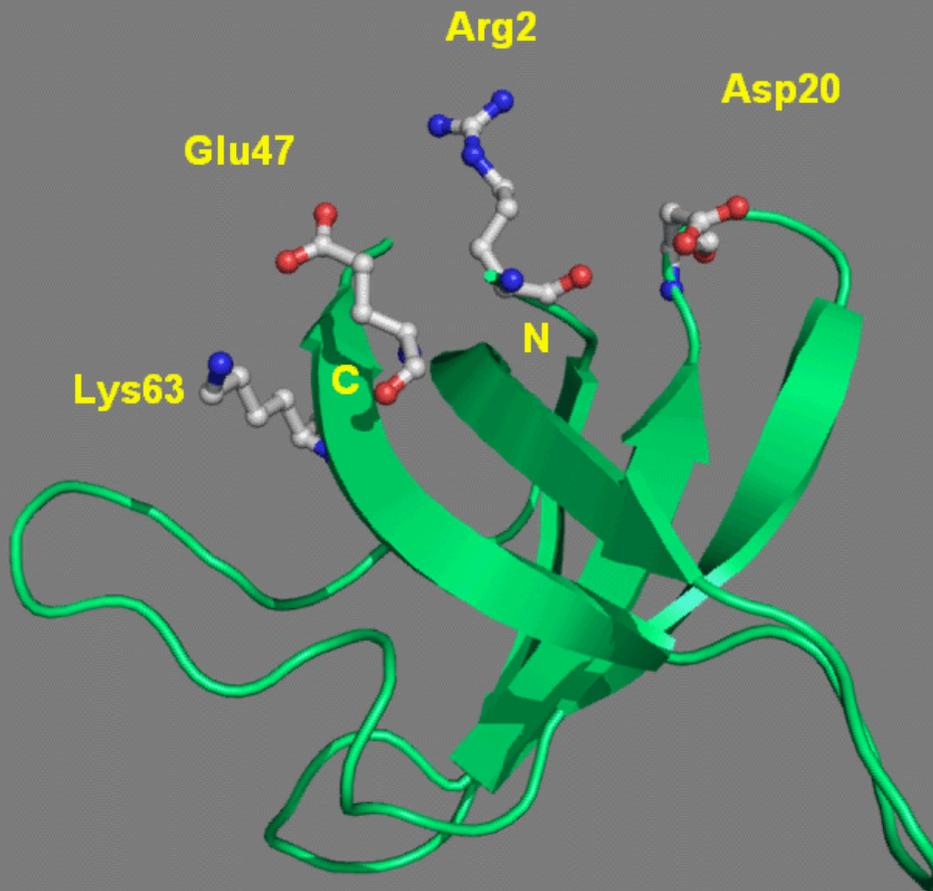
# Основные положения концепции альтернативного водородного связывания

- Регуляторный механизм термостабилизации белков обеспечивается динамическими свойствами аминокислотных остатков и их способностью образовывать **альтернативные** водородные связи.
- Непрерывный процесс разрушения и образования водородных связей придает динамические свойства белкам, необходимые для их функционирования.
- Чем больше альтернативных водородных связей тем выше термостабильность белка.
- Обладая таким механизмом, белку нет необходимости в перманентном сохранении жесткой структуры за счет максимально возможного числа прочных водородных связей - это может привести к потере динамических характеристик белка и, следовательно, к утрате его функциональных свойств.
- Движущей силой разворачивания белка является рост конформационной энтропии с температурой и стратегия белка состоит в том, чтобы с помощью оптимального количества водородных связей обеспечить структуре необходимую устойчивость, используя механизм энтропийно – энтальпийной компенсации.
- Локальные участки, ответственные за инициацию процесса разворачивания, располагаются в *петлях*, а также в N- и C-концевых участках полипептидной цепи.
- Предоставляется возможность для проведения целенаправленного мутагенеза с целью получения модифицированных белков-ферментов с повышенной термостабильностью и их использования в биотехнологических процессах.

Корреляция между количеством альтернативных водородных связей  
(в динамической модели) и термостабильностью белков.



# АЛЬТЕРНАТИВНОЕ ВОДОРОДНОЕ СВЯЗЫВАНИЕ



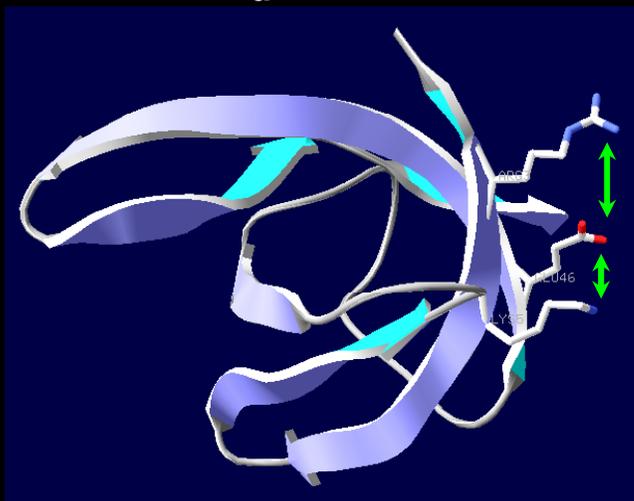
Khechinashvili et al., 2006

# Калибровка методики оценки эффективности точечных мутаций (для альтернативного водородного связывания)

**1C90**

(нативный белок)

$T_{tr}=350K$



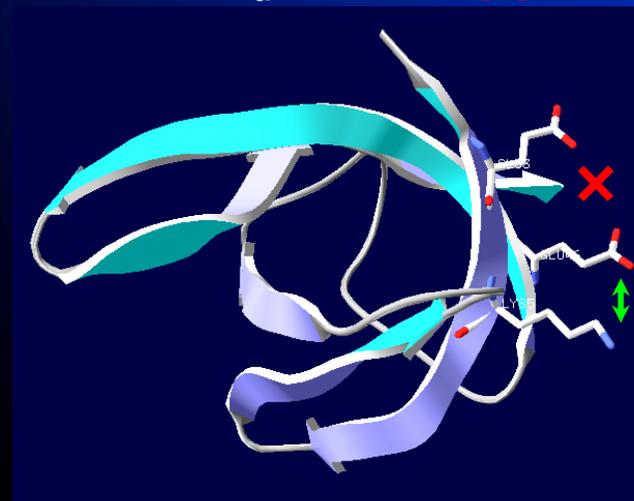
Длина  
MD-траектории  
составила  
1 нс

saltbr-ASP25-LYS7.dat  
saltbr-GLU12-LYS13.dat  
saltbr-GLU19-LYS7.dat  
saltbr-GLU21-LYS5.dat  
saltbr-GLU36-LYS39.dat  
saltbr-GLU42-LYS65.dat  
saltbr-GLU46-ARG3.dat  
saltbr-GLU46-LYS5.dat

**1I5F**

(мутант R3E)

$T_{tr}=332.2K$  (!)

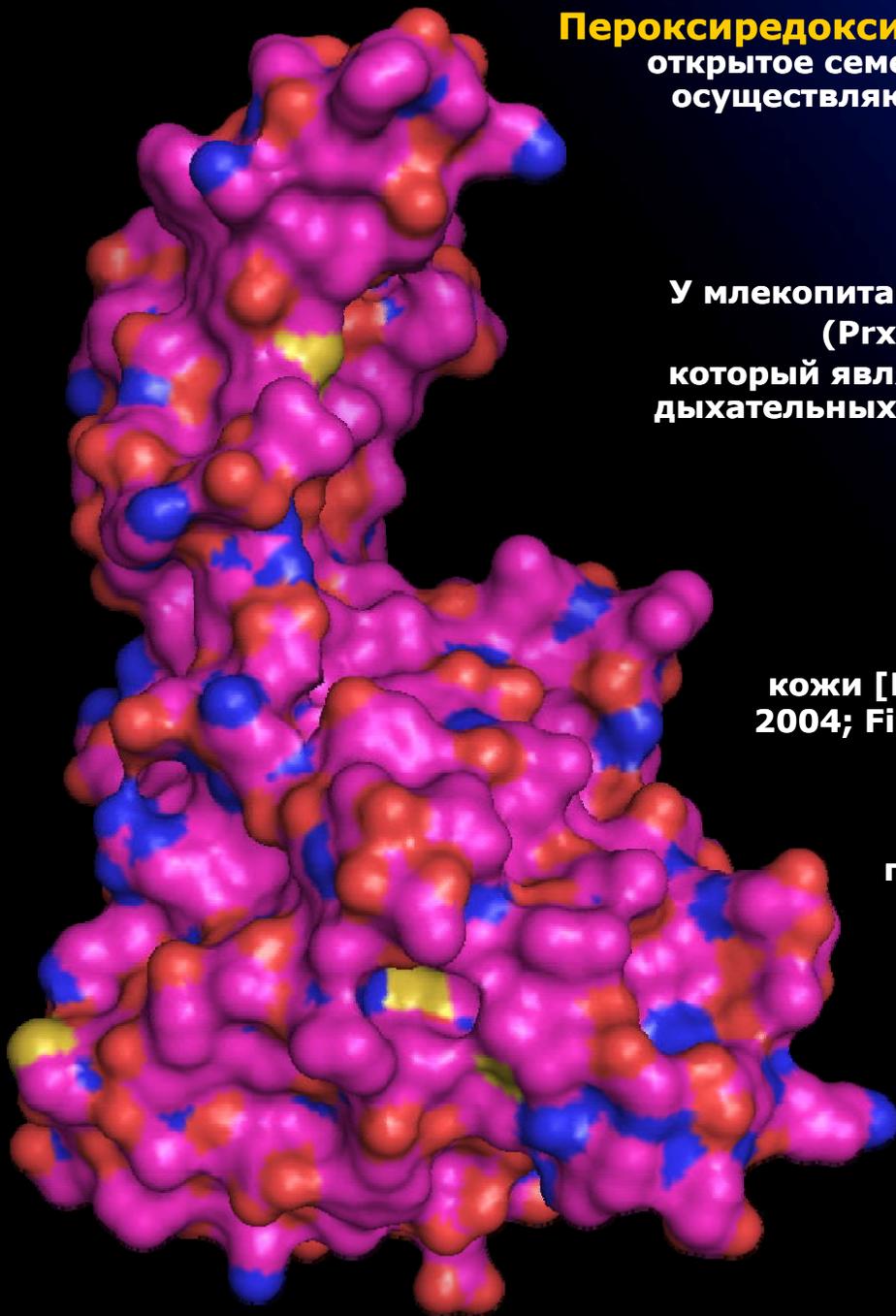


saltbr-ASP25-LYS7.dat  
saltbr-GLU12-LYS13.dat  
saltbr-GLU19-LYS7.dat  
saltbr-GLU21-LYS5.dat  
saltbr-GLU36-LYS39.dat  
saltbr-GLU42-LYS65.dat  
saltbr-GLU46-LYS5.dat



**Пероксиредоксин 6 человека**  
Белок, 204 а.к. остатка  
код PDB: **1PRX**  
(разрешение = 2 Å)

**Осуществлен поиск  
аминокислотных замен,  
которые могут существенно  
повысить термостабильность  
изучаемого белка.**



**Пероксиредоксины (Prx, КФ 1.11.1.15)** – сравнительно недавно открытое семейство селен-независимых пероксидаз, которые осуществляют ферментативную деградацию органических и неорганических пероксидов и пероксинитрита [Hofmann B. et al., 2002; Flohe L. et al., 2007].

У млекопитающих известно шесть типов пероксиредоксинов (Prx1 - Prx6), но особый интерес представляет Prx6, который является основной пероксидазой эпителия верхних дыхательных путей и легких и играет наиболее важную роль в антиоксидантной защите этих органов [Новоселов В.И. и др. 1999, 2003].

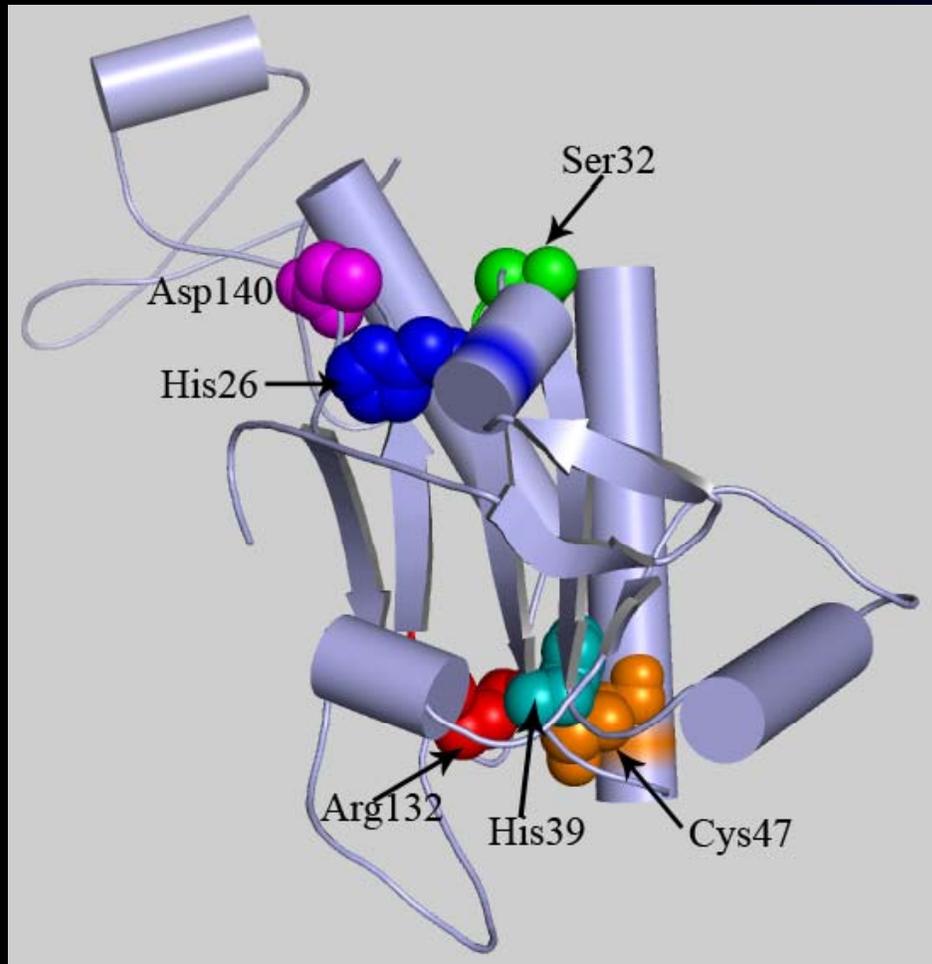
В ряде работ показана защитная роль Prx6 при различных патологиях: кожи [Kumin A. et al., 2006], легких [Kinnula V.L. et al., 2004; Fisher A.B. et al., 2004], глаз [Kubo E. Et al., 2008] и нервной системы [Power J. et al., 2008].

Терапевтическая активность Prx6 была продемонстрирована при лечении ран у крыс, после механических и термических травм кожи, а также после химических ожогов верхних дыхательных путей [Новоселов В.И. и др. 2000, 2003].

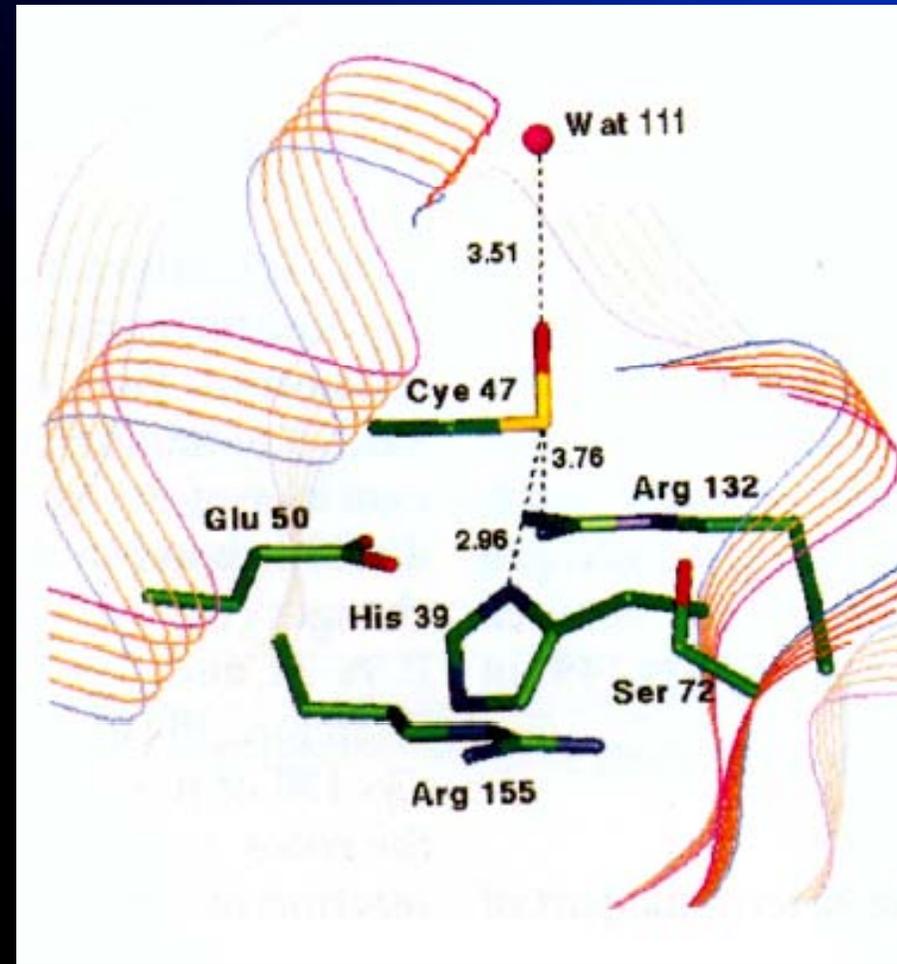
Недавно в ИБК РАН был проведен сравнительный физико-химический анализ Prx6 из различных организмов (человек, крыса, шпорцевая лягушка, дрозофила) в одних экспериментальных условиях [Шарапов М.Г., Равин В.К., 2009].

# PRX6 человека имеет два активных центра (внесение мутаций в них совершенно исключено)

Фосфолипазный центр: **His26, Ser32, Asp140**

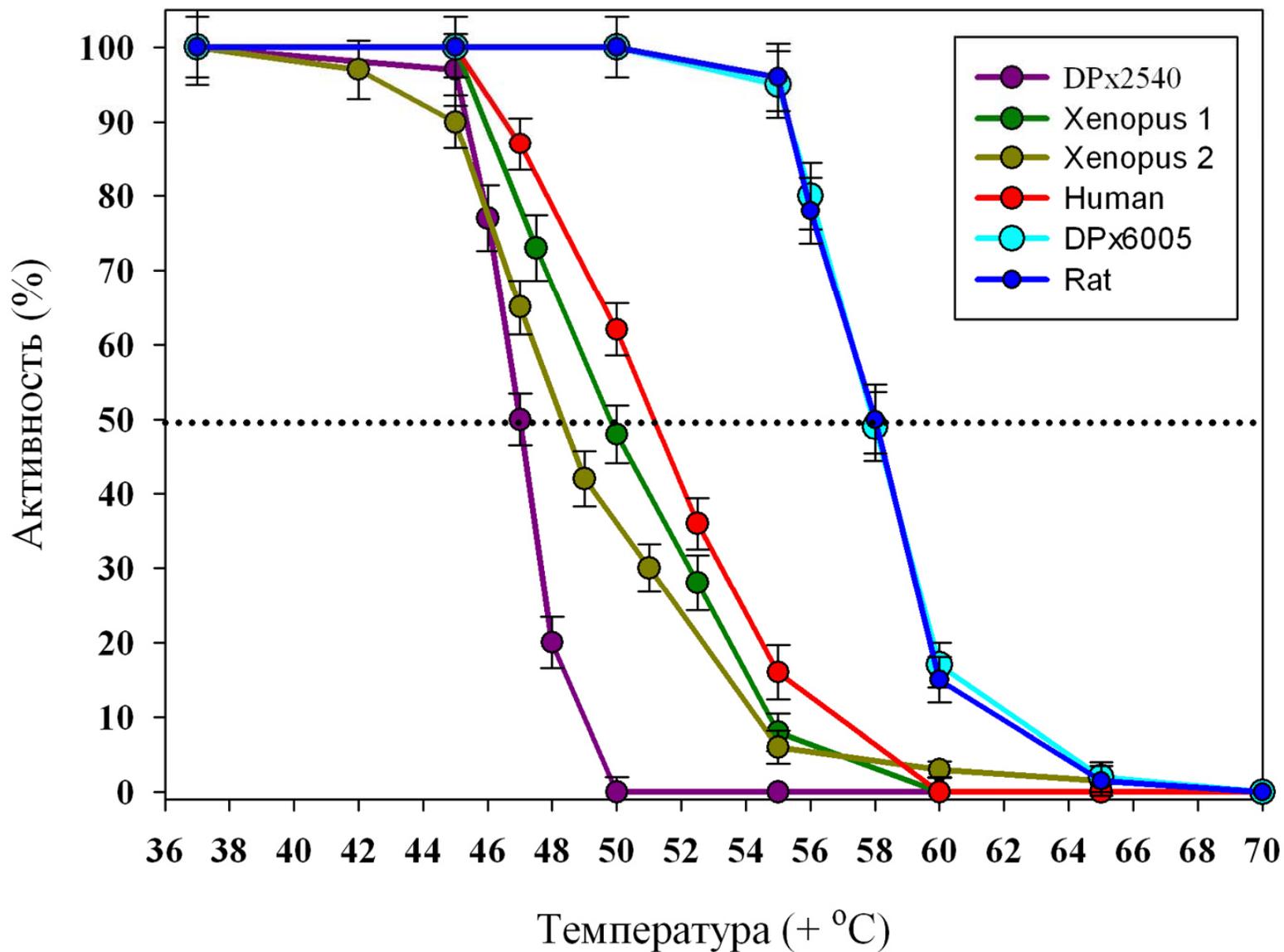


Пероксидазный центр: **His39, Cys47, Arg132**



Choi et al., 1998

# Термостабильность различных P<sub>rx</sub>6



Из материалов диссертации М.Шарапова, 2009.

# Гомология различных Prx6

	10	20	30	40	50	60	70	80	90
Human/1-224	1	MPGG - LLLGDVAPNF	EANTTVGRIRFHDFL	GD SWG I LFSHPRDF	TPVCTTELGRAAKLA	PEFAKRNVKLI	ALSIDSVEDH	LAWSKDINA	YNC 91
Rat/1-224	1	MPGG - LLLGDEAPNF	EANTTIGHIRFHDFL	GD SWG I LFSHPRDF	TPVCTTELGRAAKLA	PEFAKRNVKLI	ALSIDSVEDH	FAWSKDINA	YNG 91
Xenopus_278_/1-224	1	MPG - - ILLGERFPDF	EADTTIGRIKFHDFL	GD RWG V LFSHPRDY	TPVCTTELGR	CVKLAPEFKRN	VMIALSIDS	VEDHLGWSKD	INSYNC 90
Xenopus_309_/1-224	1	MPG - - ILLGDVFPNF	EADTTIGRIKFHDFL	GD SWG V LFSHPRDY	TPVCTTELGR	CVKLAPEFKRN	SMIALSIDS	VEDHLGWSKD	INSYNC 90
DPx6005/1-222	1	MSGKALNIGDQFPNF	TAETSEGRIDFYDWM	QD SWA I LFSHPADF	TPVCTTELSRVAAL	IPEFQKRGVK	PIALSICDP	VESHKGWIED	IKSFG - 91
DPx2540/1-220	1	- - - - MR LGQTVPNF	EADTTKGP I KFH EWQ G	NSWV V LFSHPADF	TPVCTTELGR	I AVHQPEFAKRN	TKCLAH SVD	DALNSHVDW	VNDIKSYCL 87

	100	110	120	130	140	150	160	170	180
Human/1-224	92	E E P T E K L P F P I I D D R N R E L A I L L G M L D P A E K D E K G M P V T A R V V F V F G P D K K L K L S I L Y P A T T G R N F D E I L R V V I S L Q L T - A E K R V A T P V D W K	182						
Rat/1-224	92	A A P T E K L P F P I I D D K D R D L A I L L G M L D P A E K D E K G M P V T A R V V F I F G P D K K L K L S I L Y P A T T G R N F D E I L R V V D S L Q L T - A S N P V A T P V D W K	182						
Xenopus_278_/1-224	91	D E P T E T L P F P I I A D P K R D L A V Q L G M L D P D E K D M Q G M P V T A R C V F I I G P D K K M K L S I L Y P A T T G R N F D E I L R V V D S L Q L T - A V H N V A T P V D W K	181						
Xenopus_309_/1-224	91	D E P T E T L P F P I I A D P K R E L A V Q L G M L D P D E K D M Q G M P V T A R C V F I I G P D K K M K L S I L Y P A T T G R N F D E I L R V V D S L Q L T - A V H N V A T P V D W K	181						
DPx6005/1-222	92	- - K L S S F D Y P I I A D D K R E L A L K F N M L D K D E I N A E G I P L T C R A V F V V D D K K L R L S I L Y P A T T G R N F D E I L R V I D S L Q L T - Q T K S V A T P A D W K	180						
DPx2540/1-220	88	D - I P G D F P Y P I I A D P T R D L A V S L G M L D E E Q K K D P E V G K T I R A L F I I S P D H K V R L S M F Y P M S T G R N V D E I L R T I D S L Q L T D R L K V V A T P A N W T	178						

	190	200	210	220
Human/1-224	183	D G D S V M V L P T I P E E E A K K L F P K G V F T K E L P S G K K Y L R Y T P Q P -	224	
Rat/1-224	183	K G E S V M V L P T L P E E E A K Q L F P K G V F T K E L P S G K K Y L R Y T P Q P -	224	
Xenopus_278_/1-224	182	P G D R V M V P P T V P E E E A R K I F T C G V F T K E L P S G K K Y L R Y T P Q P Q	224	
Xenopus_309_/1-224	182	P G D R V M V P P N V P E E E A S K I F T C G V F T K E L P S G K K Y L R Y T A Q P Q	224	
DPx6005/1-222	181	Q G G K C M V L P T V K A E D V P K L F P D G I E T I E L P S G K S Y L R I T P Q P -	222	
DPx2540/1-220	179	P G T K V M I L P T V T D E E A H K L F P K G F D K V S M P S G V N Y V R T T D N Y -	220	

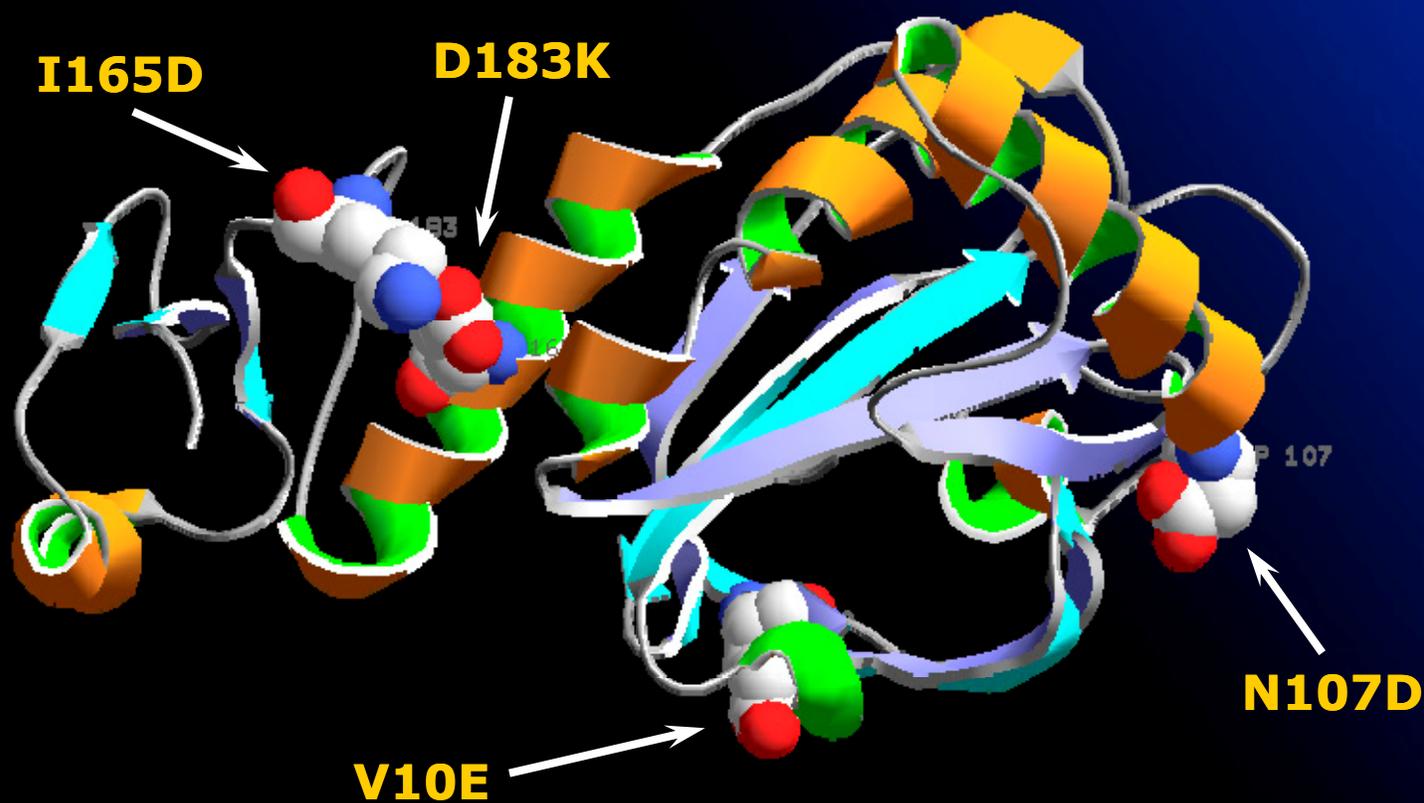
Prx6 из других организмов	Rattus norvegicus	Xenopus laevis (309)	Xenopus laevis (278)	Drosophila melanogaster (DPx6005)	Drosophila melanogaster (DPx2540)
Идентичность С Prx6 Human (%)	<b>91,5</b>	80,4	79	60	53,7

Из материалов диссертации М.Шарапова, 2009.

**Результат расчета траектории молекулярной динамики  
(AMBER96, CUDA-GROMACS, 10 ns, 300K, accurate water)**



**Итогом изучения структуры пероксиредоксина б человека и моделирования динамики её изменений стало обоснование важности для повышения термостабильности фермента внесения всего 4 (из 19-ти) мутаций**



# Липиды: один из основных классов загрязнителей

Весьма эффективно липиды расщепляются биокатализаторами-ферментами, многие из которых имеют бактериальное происхождение (это плюс), но низкую термостабильность (это минус). Осознанное увеличение термостабильности липаз теперь стало достижимым!



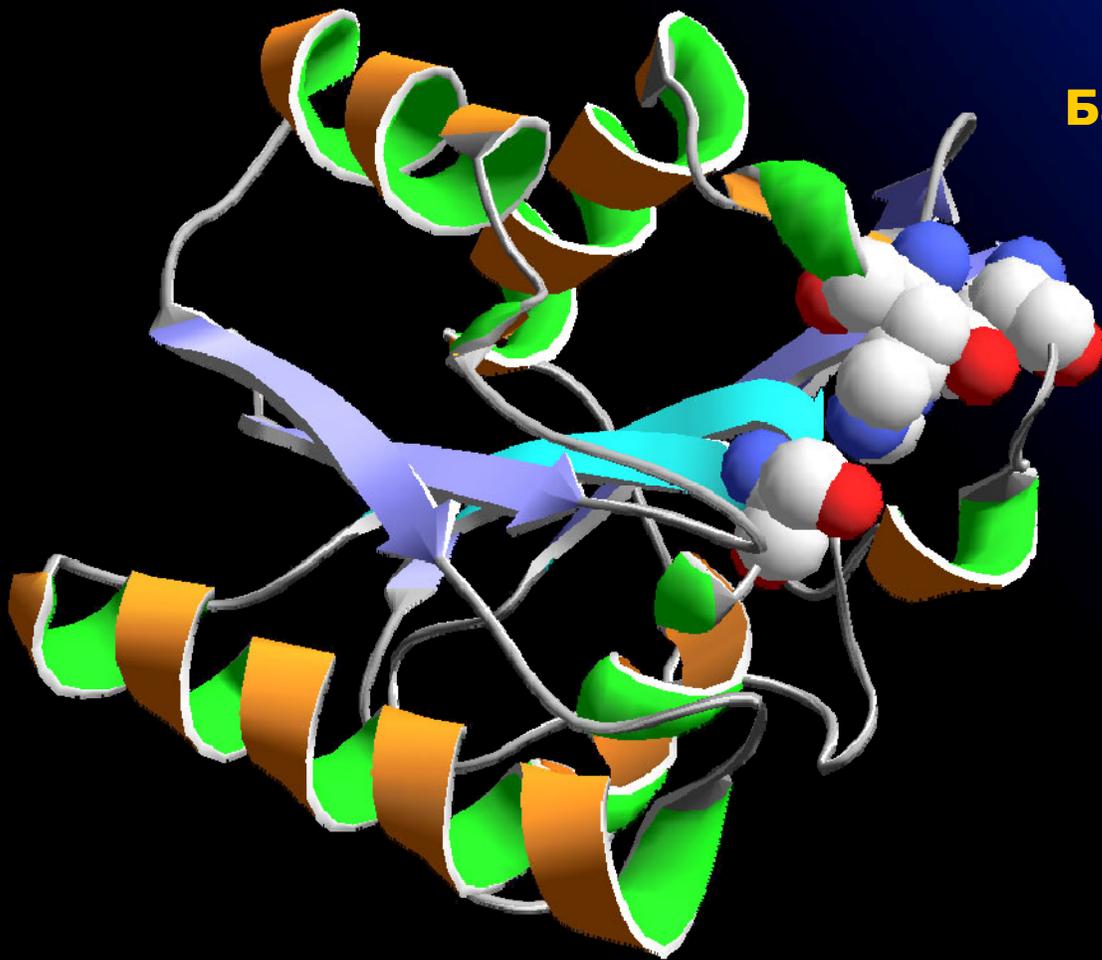
Повышенная  
скорость  
реакции

Удобство  
хранения

Высокая  
термостабильность  
фермента

Инактивация  
паразитных  
энзимов

Экономия на  
охлаждении



**Бактериальная липаза LipA**

Белок, 181 а.к. остаток

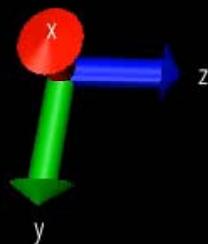
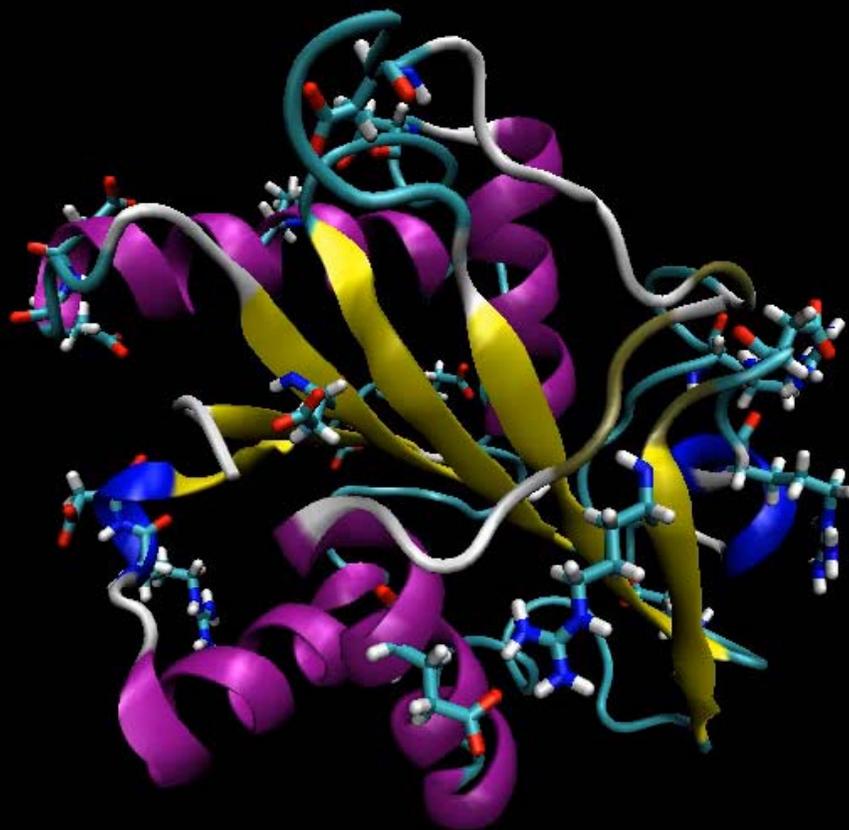
Код PDB: **2QXU**

(разрешение = 1.5 Å)

**Активный центр: триада  
Ser77, Asp133, His156.**

**Существующие  
модификации фермента  
пока не запатентованы**

Результат расчета траектории молекулярной динамики  
(AMBER96, CUDA-GROMACS, 10 ns, 300K, accurate water)



## 4. Практическая часть.

**Спасибо за внимание!**

# **Соавторы:**

**Н.Н. Хечинашвили**

**В.М. Комаров**

**В.И. Новосёлов**

**А.А. Самченко**

**М.Г. Шарапов**

**Ю.А. Пуртов**