

УДК 577.150.2

ВАРИАБЕЛЬНЫЕ ПОЗИЦИИ И УЧАСТКИ АМИНОКИСЛОТНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ β -ЦЕПЕЙ ГЕМОГЛОБИНОВ МЛЕКОПИТАЮЩИХ ОТРЯДА ХИЩНИКОВ

Костецкий П.В.

ФГБУН Институт биоорганической химии им. академиков М.М.Шемякина и Ю.А.Овчинникова РАН,
117997 г. Москва, ул. Миклухо-Маклая 16/10, E-mail: pykost1940@ibch.ru

Резюме. Для выявления переменных позиций и участков в гомологичных аминокислотных (АК) последовательностях сравнивали АК-последовательности β -цепей гемоглобинов (Hb β) представителей отряда хищников *Carnivora*. При “групповом сравнении” Hb β -последовательностей 16-ти представителей двух подотрядов хищников, включающих 10 представителей *Caniformia* (псовые) и 6 - *Feliformia* (кошачьи) обнаружили 35 переменных позиций, тогда как при обычном парном сравнении число замен не превышает 26. Среди 35 переменных позиций имеется два отдельных набора из 12 и 10 позиций для псовых и кошачьих Hb β -последовательностей, соответственно, которые в основном отвечают за видовое разнообразие Hb β -молекул в подотрядах.

Большинство переменных позиций сосредоточено в 4 фрагментах, занимающих участки 1-17, 51-59, 74-85 и 121-131. В отдельных «групповых сравнениях» Hb β -последовательностей каждого из подотрядов характеристики и положение переменных участков различаются. Участок 51-59 наиболее переменен для подотряда *Caniformia*, тогда как участки 78-85 и 121-131 переменны только для подотряда *Feliformia*. При этом оказывается возможным предсказать структуру таксонспецифичного октапептида 78-85 NLDNLKGT инвариантного только для Hb β -молекул подотряда псовых *Caniformia*.

На множестве гомологичных АК-последовательностей, получаемых компьютерной перестановкой столбцов двумерного массива из 10 Hb β -последовательностей подотряда хищников *Caniformia*, показали, что достоверность неслучайного наличия участка из 17 позиций с 11 переменными позициями выше 95%. Достоверность наличия тандема с участием второго участка из 9 позиций с 6 переменными позициями тоже высока ($p > 99\%$). Уровни достоверности переменных участков Hb β -последовательностей отряда *Carnivora* ($p \sim 93\%$) и подотряда *Caniformia* ($p > 99\%$) свидетельствуют о высокой линейной неравномерности АК-замен в Hb β -молекулах.

Ключевые слова: профиль изменчивости, гомологичные последовательности, таксонспецифический пептид, видовое разнообразие, Монте-Карло

THE VARIABLE POSITIONS AND REGIONS OF HEMOGLOBIN β -CHAIN AMINOACID SEQUENCES OF THE *CARNIVORA* MAMMALIA

Kostetsky P.V.

Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry RAS, Moscow, E-mail:
pykost1940@ibch.ru

Summary. The hemoglobin β -chains (Hb β) aminoacid (AA) sequences of 16 *Carnivora* species (10 *Caniformia* and 6 *Feliformia*) were “group-compared” to locate the variable positions and regions of homologous amino acid sequences. It was found that this group of 16 Hb β -sequences has 35 variable position, but the pairwise comparison gives maximum 26 AA-changes. Among 35 variable positions there are two separate sets of 12 and 10 items which belong to *Caniformia* or *Feliformia*, and are responsible for the species variety of Hb β -molecules.

Most of the variable positions are located in four variable regions and occupied positions 1-17, 51-59, 74-85 and 121-131. But the location and peculiarities of variable regions of *Caniformia* and *Feliformia* (two suborders of *Carnivora*) Hb β -sequences are very different. The 51-59 region is variable (and active during evolution) only for *Caniformia* and the 74-85, 121-131 regions are variable only for *Feliformia* Hb β -sequences. Moreover there is the possibility to predict the taxonspecific peptide structure for the group of *Caniformia* Hb β -sequences – 78-85 NLDNLKGT. This fragment is constant for 10 *Caniformia* Hb β -sequences but each of 6 *Feliformia* Hb β -sequences contains three AA-changes.

Using the great number of artificial homologous AA-sequences - randomly permuted columns of *Caniformia* 10 Hb β -sequences twodimensional array - was possible to show that the presence of variable regions is not accidental. The probability of the by chance appearance of the region with 11 variable positions among 17 position is low (confidence level 5%). Moreover the combined presence of such variable region and the second 51-59 region with 6 variable positions is also not accidental ($p > 99\%$). The confidence levels of *Carnivora* ($p \sim 93\%$) and *Caniformia* ($p > 99\%$) Hb β -sequences demonstrate the high linear irregularity of the variable positions in Hb β -molecules.

Keywords: variability profile, homologous sequences, taxonspecific peptide, species variety, Monte-Carlo

Введение. С ростом числа доступных гомологичных АК-последовательностей белков, выполняющих одну и ту же функцию в различных организмах, растет число работ с анализом групп АК-последовательностей этих белков [1-5]. В результате формируют профили изменчивости, в которых видны вариабельные и консервативные позиции [6].

Некоторые авторы сравнивают гомологичные АК-последовательности, принадлежащие определенному таксону или нескольким близким таксонам [7-9].

Так, подотряд чешуйчатых *Serpentes* (змеи) представлен рядом семейств, для которых известны десятки АК-последовательностей фосфолипаз А2. Было установлено, что эти молекулы имеют при попарном сравнении до 50% АК-замен, и часть из них сосредоточена в вариабельных участках [9]. При сравнении нескольких АК-последовательностей β -цепей гемоглобинов (Hb β -последовательностей) позвоночных обнаружили, что из 146 позиций более 100 содержат АК-замены. При этом около 20 позиций имеют два или более АК-остатков, отличных от доминантного [10,11].

Показано, что профили изменчивости миоглобинов двух отрядов млекопитающих – китообразных и приматов – имеют различия. Эти различия наиболее заметны в вариабельном участке из 38 позиций, в котором число вариабельных позиций у приматов почти втрое выше, чем у китообразных. Авторы предположили [7], что наблюдаемые различия важны для изменения функциональных характеристик миоглобинов двух отрядов млекопитающих.

Целью данной работы явилось: а) выявление таксонных различий в количестве и местоположении вариабельных позиций и участков в Hb β -последовательностях подотрядов млекопитающих отряда хищников *Carnivora* и б) оценка достоверности вариабельных участков в сравниваемых группах Hb β -последовательностей методом Монте-Карло с использованием множества искусственных гомологичных АК-последовательностей.

Материалы и методы. В качестве объекта исследования были взяты Hb β -последовательности отряда млекопитающих хищников *Carnivora* (10 из подотряда псовых *Caniformia* и 6 из подотряда кошачьих *Feliformia*). Названия организмов и соответствующие им коды Hb β -последовательностей отвечают данным, содержащимся в банке белковых структур SwissProt.

Подотряд псовых *Caniformia*: *Canis familiaris* (собака) – HBB_*canfa*, *Procyon lotor* (енот) – HBB_*prolo*, *Meles meles* (барсук) – HBB_*melme*, *Odobenus rosmarus divergens* (морж) – HBB_*odoro*, *Vulpes vulpes* (лиса) – HBB_*vulvu*, *Leptonychotes weddelli* (тюлень *weddelli*) – HBB_*lepwe*, *Phoca vitulina* (тюлень обыкновенный) – HBB_*phovi*, *Mustela lutreola* (норка) – HBB_*mushu*, *Ailuropoda melanoleuca* (большая панда) – HBB_*ailme*, *Ailurus fulgens* (рыжая панда), – HBB_*ailfu*. Подотряд кошачьих *Feliformia*: *Felis catus* (домашняя кошка) – HBB_*felca*, *Panthera pardus saxicolor* (персидский тигр) – HBB_*panps*, *Panthera pardus orientalis* (амурский тигр) – HBB_*panpo*, *Paguma larvata* (пальмовая циветта) – HBB_*pagla*, *Lynx lynx* (рысь) – HBB_*lynly*, *Proteles cristata* (гиена) – HBB_*procr*.

Группы искусственных гомологичных АК-последовательностей получали методом Монте-Карло [9,12] случайной перестановкой столбцов двумерного массива выровненных Hb β -последовательностей. Так как в исходном и в случайных массивах элементы каждого столбца остаются неизменными, то и его новая позиция остается или вариабельной или инвариантной (символ «*» или « »). При этом во всех случаях результаты сравнения гомологичных АК-последовательностей могут быть представлены в виде «строк изменчивости» (Рис.1), составленных из постоянного количества символов «*» для позиций, содержащих АК-замены, и пробел « » для инвариантных позиций.

Поэтому для оценки достоверности наличия вариабельных участков выполняли численные эксперименты на 100000 или более случайных строках изменчивости, получаемых перестановкой символов исходной строки изменчивости при помощи компьютерного датчика случайных чисел. Случайную строку изменчивости анализируют на присутствие участка заданной длины, в котором количество наблюдаемых «*»-символов строки изменчивости – (Рис. 1) равно или превышает значение в исходном природном участке сравнения. Число случайных строк, в которых наблюдались участки с указанными свойствами, дает возможность оценивать частоту событий и тем самым определять достоверность наличия (или вероятность случайного возникновения) вариабельного участка в массиве сравниваемых природных Hb β -последовательностей.

При оценке достоверности одновременного присутствия двух вариабельных участков (тандема) определяли количество вариабельных позиций («*»-символов) в двух участках случайных строк изменчивости. Длины участков считали неизменными, а число «*»-символов было не меньше, чем в оцениваемых участках. Вариабельный участок или тандем считали достоверными, если вероятность их случайного появления в численном эксперименте не превышала уровень 5%. Достоверность тандема оценивалась среди событий с наличием более достоверного участка, и число таких событий должно быть не менее 100.

Для обнаружения позиций, где наблюдаются согласованные АК-замены [13], выполняли все парные сравнения N Hb β -последовательностей из L остатков ($L = 147$). Результатом сравнения для каждой АК-позиции являлась строка длиной $N*(N-1)/2$ (число парных сравнений) содержащая символы «0» (нет замен в парном сравнении) и «1» (есть замены). Полученные L (число всех АК-позиций) строк сравнивали попарно, фиксируя номера каждой из АК-позиций при идентичности строк. Подобный результат для двух АК-позиций считали необходимым признаком согласованности АК-замен.

Результаты и обсуждение. Отряд хищников *Carnivora* состоит из двух подотрядов: псовые *Caniformia* и кошачьи *Feliformia*. Очевидно, что возможно сравнение Hb β -

последовательностей трех групп: близкородственных таксонов *Caniformia* и *Feliformia* и объединенной группы *Carnivora*.

Среди Hbβ-последовательностей подотряда *Caniformia* наибольшее число различий наблюдается для пары *Procyon lotor* / *Vulpes vulpes* – 14 замен, а в подотряде *Feliformia* – для пары *Panthera pardus orientalis* / *Proteles cristata* – 18 замен. При сравнении Hbβ-последовательностей разных подотрядов наблюдается разброс значений от 14 до 26 замен (пары *Proteles cristata* / *Phoca vitulina* и *Panthera pardus saxicolor* / *Meles meles*, соответственно). При групповом сравнении всех 16 Hbβ-последовательностей отряда хищников число переменных позиций не может быть меньше 26, так как в остальных сравниваемых парах могут быть дополнительные позиции доступные для мутаций.

Для поиска числа и местоположения всех возможных переменных позиций в Hbβ-последовательностях отряда хищников выполняли групповое сравнение. Пример сравнения группы из 8 Hbβ-последовательностей представлен на Рис.1. Результат сравнения имеет простой вид нижней «строки изменчивости», составленной из пробелов для инвариантных позиций и «*»-символов для позиций, содержащих замены.

	10	20	30	40	50	60
HBB_canfa	-VHLT A E E KSLVSGLWGKVN V DEVGGEALGRLL I VYPW T Q R FFDSFGDL S TPDA V MS N AK					
HBB_prolo	-VHLT A DE K TA V TT L WGKVN V EEVGGEALGRLL V VYPW T Q R FFESFGDL S SAD A IMGN P AK					
HBB_ailme	MVHLT G E E KA A VTGLWSKVN V DEVGGEALGRLL V VYPW T Q R FFDSFGDL S TPDA V M N NP K					
HBB_lepwe	-VHLT A E E K S A V TALWGKVN V DEVGGEALGRLL V VYPW T Q R FFDSFGDL S SP N AIM S NP K					
HBB_felca	- G FLT A E E KGLV N GLWGKVN V DEVGGEALGRLL V VYPW T Q R FFESFGDL S SADAIM S NA K					
HBB_panpo	MS F LS A E E KNL V SGLWGKVN V DEVGGEALGRLL V VYPW T Q R FFQ S FGDL S SADAIM S NA K					
HBB_pagla	- G FLT A E E KGLV N GLWGKVN V DEVGGEALGRLL V VYPW T Q R FFQ S FGDL S SADAIM H NS K					
HBB_panps	- G FLT A E E KSLV N DLWSKVN V DEVGGEALGRLL V VYPW T Q R FFQ S FGDL S SADAIM G NG K					
	*** **	*	*	*	*	*** **
	70	80	90	100	110	120
HBB_canfa	VKAHGK K VLNSFSDGLKNL D N L KG T FAKLSELHCD K LHVDPENFKLLGNVLVCVLAHHFG					
HBB_prolo	VKAHGK K VLNSFSEGLKNL D N L KG T FAKLSELHCD K LHVDPENFRLLGNVLVCVLAHHFG					
HBB_ailme	VKAHGK K VLNSFSEGLKNL D N L KG T FAKLSELHCD K LHVDPENFKLLGNVLVCVLAHHFG					
HBB_lepwe	VKAHGK K VLNSFSDGLKNL D N L KG T FAKLSELHCD Q LHVDPENFKLLGNVLVCVLAHHFG					
HBB_felca	VKAHGK K VLNSFSDGL K NI D DL K GA F AKLSELHCDKLHVDPENFRLLGNVLVCVLAHHFG					
HBB_panpo	VKAHGK K VLNSFSDGL K NI D DL K GA F AKLSELHCDKLHVDPENFRLLGNVLVCVLAHHFG					
HBB_pagla	VKAHGK K VLNSFSDGL K H V DDL K GT F AKLSELHCDKLHVDPENFKLLGNVLVCVLAHHFG					
HBB_panps	VKAHGK K VLNSFSDGL K H I DDL K GT F AKLSELHCDKLHVDPENFKLLGNVLVCVLAHHFG					
	*	** *	*	*	*	*
	130	140				
HBB_canfa	KEFT P Q V QAAYQKV V AGVANALAHKY H		собака			
HBB_prolo	KEFT P Q V QAAYQKV V AGVANALAHKY H		енот			
HBB_ailme	KEFT P Q V QAAYQKV V AGVANALAHKY H		панда			
HBB_lepwe	KEFT P Q V QAAYQKV V AGVANALAHKY H		тюлень			
HBB_felca	H DF N PQ V QA A FQKV V AGVANALAHKY H		кошка			
HBB_panpo	H EF N PQ V QA A FQKV V AGVASALAHRY H		амурский тигр			
HBB_pagla	K EFT P Q V QAAYQKV V AGVASALAHRY H		гиена			
HBB_panps	N EFT P Q V QAAYQKV V AGVANALAHKY H		персидский тигр			
	** *	*	*	*	*	*

Рис. 1. Сравнение Hb β -последовательностей хищников (4 представителя подотряда *Caniformia* и 4 – *Feliformia*). Серым выделены вариабельные участки, а жирным шрифтом – позиции, вариабельные для одного подотряда и важные для видового разнообразия Hb β -молекул подотряда.

В строке изменчивости видно, что в 8 Hb β -последовательностях вариабельными являются 34 позиции, тогда как при парном сравнении наблюдаемое число замен не выше 24 (пара *Ailuropoda melanoleuca* / *Panthera pardus orientalis*). Отметим, что 22 позиции вариабельны только в пределах одного из подотрядов, и они, по-видимому, имеют большое значение для видового разнообразия Hb β -молекул подотрядов.

В численном эксперименте с использованием большого числа искусственных гомологичных АК-последовательностей (см. Материалы и методы исследования) показали, что вероятность случайного появления в любом месте участка из 17 позиций с 11 или более вариабельными позициями менее 6%. Более того, при наличии участка с такими характеристиками вероятность случайного тандема с появлением второго участка подобного 51-59 не превышает 4%.

Аналогично сравнивали АК-последовательности четырех групп (Таблица), каждая из которых принадлежала определенному таксону. Во всей группе отряда хищников из 16 Hb β -последовательностей имеется 4 вариабельных участка, а в группе Hb β -последовательностей подотряда псовых *Caniformia* имеются только два вариабельных участка, но их совместное присутствие отличается высокой достоверностью ($p > 99\%$). В группе сравнения подотряда кошачьих *Feliformia* присутствуют два вариабельных участка, содержащие по 7 и 5 вариабельных позиций, но достоверность их совместного присутствия около 91%.

Данные Таблицы для всего отряда хищников дают суммарные сведения о АК-заменах Hb β -последовательностей обоих подотрядов. Однако сравнение эволюционных изменений в каждом подотряде обнаруживает важные таксонные различия. Только N-концевой участок является вариабельным во всех случаях, тогда как остальные три вариабельных участка проявляют себя различно в двух подотрядах. Участок 51-59 является очень вариабельным для подотряда псовых и умеренно вариабельным для кошачьих – 6 и 2 вариабельных позиций, соответственно. Напротив, участок 121-131 является вариабельным только у кошачьих, но не у псовых – 5 и 1 вариабельная позиция, соответственно. Участок 78-85 является вариабельным только у кошачьих – 3 вариабельных позиций.

Таблица. Количество и локализация вариабельных позиций и участков при групповом сравнении Hb β -последовательностей млекопитающих отряда хищников *Carnivora* и подотрядов *Caniformia* и *Feliformia*. Нумерация позиций соответствует Рис. 1. Жирным шрифтом отмечены достоверные тандемы из двух вариабельных участков ($p > 95\%$).

Таксон группы сравнения / число Hb β -последовательностей	Максимум замен при попарном сравнении	Число переменных позиций в группе	Локализация переменных участков / число переменных позиций в участке
Отряд <i>Carnivora</i> / 8*	24	34	1-17 / 11; 51-59 / 6; 74-85 / 5; 121-131 / 5
Отряд <i>Carnivora</i> / 16	26	35	1-17 / 11; 51-59 / 6 **; 74-85 / 5; 121-131 / 5
Подотряд <i>Caniformia</i> / 10	16	23	1-17 / 8; 51-59 / 6
Подотряд <i>Feliformia</i> / 6	18	21	1-17 / 7; 121-131 / 5, 78-85 / 3

* Состав группы см. Рис.1. Состав других групп см. Материалы и методы исследования.

** Достоверность данного тандема - 93%.

Групповое сравнение 16 Hb β -последовательностей отряда *Carnivora* позволяет оценить общее число переменных позиций. Максимальное число замен – 26 (18%) - имеется в паре *Panthera pardus saxicolor* / *Meles meles*, тогда как число переменных позиций в 16-и Hb β -последовательностей хищников заметно выше – 35 (24%) (Таблица). Важно, что ряд позиций являются переменными только в каждом из подотрядов в отдельности – 12 и 10 для *Caniformia* и *Feliformia*, соответственно. (Рис. 1). Можно предположить, что в основном эти позиции отвечают за видовое разнообразие Hb β -молекул млекопитающих каждого из подотрядов хищников *Carnivora*.

Отметим, что в некоторых парах переменных позиций АК-замены заметно согласованы [13]. Например, в паре 7 & 22 происходят синхронные близкие функционально замены E / D и D / E, а в паре 51 & 55 - замены T / S и V / I (Рис 1).

При идентификации переменных участков в Hb β -молекулах представителей отряда *Carnivora* оказалось, что участок 78-85 проявляется по-разному у двух подотрядов, и число замен в подотрядах *Caniformia* и *Feliformia* составляет 0 и 3 (позиции 78,79,81 и 85), соответственно (Рис. 1 и 2).

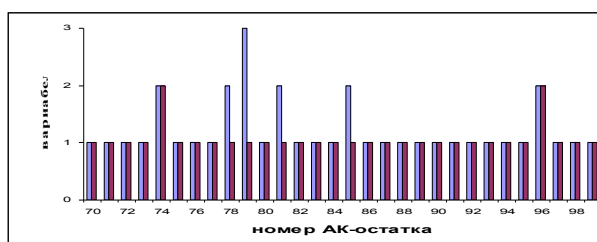


Рис. 2. Вариабельность фрагмента 70-99 Hb β -последовательностей 16 хищников отряда *Carnivora* (серый фон) и 10 подотряда *Caniformia* (черный фон). По оси ординат отложена вариабельность, определяемая как число разных АК-остатков, присутствующих в данной позиции.

Похожая ситуация отмечена для АК-последовательностей миоглобинов отрядов приматов и китообразных [7]. В одном из переменных участков длиной 38 позиций число переменных позиций существенно отличалось в двух сравниваемых группах – 16 и 6 позиций, соответственно, тогда как по всей длине Hb β -молекул это различие менее заметно – 38 и 29 переменных позиций.

Полученный результат по различию в числе переменных позиций в участке 74-85 Hb β -молекул двух подотрядов хищников (Рис. 1 и 2) полезен для предсказания таксонспецифичных пептидов [9]. Таким пептидом для Hb β -молекул подотряда *Caniformia*

является инвариантный октапептид 78-85 NLDNLKGT. Напротив, в каждой из 6 Hb β -последовательностей подотряда *Feliformia* на указанном участке имеется по 3 АК-замены.

Выводы. При сравнении Hb β -последовательностей отряда хищников (*Carnivora*) определили число и местоположение переменных позиций и переменных участков в группах Hb β -молекул подотрядов *Caniformia*), *Feliformia* и объединенной группы Hb β -молекул отряда *Carnivora*. В профилях изменчивости Hb β -молекул подотрядов *Caniformia* и *Feliformia* обнаружили переменные позиции, отвечающие за видовые отличия Hb β -молекул каждого из подотрядов. В Hb β -последовательностях отряда *Carnivora* участки 75-85 и 121-131 переменны только для подотряда *Feliformia*, а участок 51-59 – для подотряда *Caniformia*. При этом октапептид 78-85 NLDNLKGT является инвариантным и таксонспецифичным только для Hb β -молекул подотряда *Caniformia*.

С помощью искусственных гомологичных АК-последовательностей методом Монте-Карло показали высокую достоверность переменных участков в Hb β -молекулах отряда хищников.

Список литературы

1. Гомологи гена *AlkB* термофильных бактерий рода *Geobacillus* / Турова Т.П., Назина Т.Н., Михайлова Т.А., Родионова Т.А., Екимов А.Н., Машукова А.В., Полтораус А.Б. // Молекулярная биология – 2008. – V.42. - № 2. – С. 247-257.
2. Кожокина О., Ковалева Т. Сравнительный анализ аминокислотных последовательностей глюкоамилаз рода *Aspergillus* // Фундаментальные исследования – 2009.- № 8.- С. 19-21.
3. Порозов Ю.Б. Биоинформатика и средства компьютерного анализа и визуализации макромолекул // Саратовский научно-медицинский журнал – 2010. - № 2. – С. 273-276.
4. Lukashov V.V., Goudsmit J. Recent evolutionary history of HIV-1 subtype B // J. Mol. Evol. – 2003 – V. 56 – P. 645-647.
5. Identification of functionally conserved residues with the use of entropy-variability plots / Olivera L., Paiva P.B., Paiva A.C.M., Vriend G. // Proteins – 2003. – V. 52. – P. 544 -552.
6. PVS: a web server for protein sequence variability analysis tuned to facilitate conserved epitope discovery / Garsia-Boronat M., Diez-Rivero C., Reinherz E.L., Reche P.F. // Nucleic Acids Research – 2008 –V. 35 – W35-W41.
7. Naylor G.J.P., Gerstein M. Measuring shifts in function and evolutionary opportunity using variability profiles: a case study of the globins // J. Mol. Evol. – 2000. – V. 51. – P. 223-233.
8. Strommer J. The plant ADH gene family // Plant J. – 2011. – Vol. 66, N 1. – P. 128-142.

9. Kostetsky P., Arkhipova S., Vladimirova R. Conservative and variable regions of snake phospholipases A2 sequences: prediction of the taxonspecific peptides structure // J. Protein Chem. – 1991. – Vol. 10, N 6. – P. 593-601.
10. Dickerson R.E., Irving G. Hemoglobin, Structure, Function, Evolution, and Pathology // [Benjamin Cummings series in the life sciences. Frontiers in Physics](#), 1983, P. 1-176.
11. Кантор Ч., Шиммел П. Биофизическая химия в 3 т. // М., «Мир», 1984, Том 1, С. 1-336.
12. Соболев И.М. Метод Монте-Карло // М., «Наука», 1978, С. 1-64.
13. Marks D.S., Hopf T.A., Sander C.D. Protein structure prediction from sequence variation // Nat. Biotechnol. – 2012 – V.30. – P. 1072-80.