

МИНИСТЕРСТВО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
Санкт-Петербургский государственный электротехнический университет

УДК [577.38 + 575.12] : 51-7

В.А.Карасев

ОБ АНТИСИММЕТРИЯХ КАНОНИЧЕСКИХ АМИНОКИСЛОТ

Деп. в ВИНТИ 23.03.2004 г. N 470-B2004.

Санкт-Петербург
2004

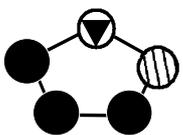
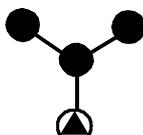
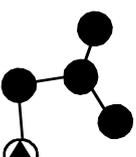
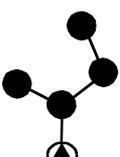
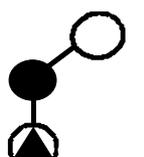
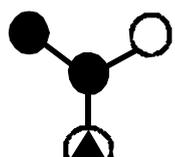
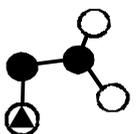
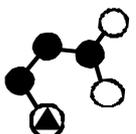
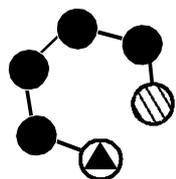
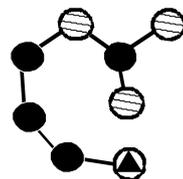
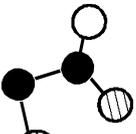
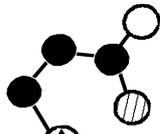
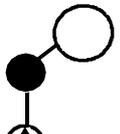
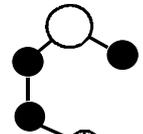
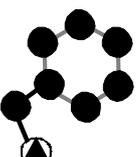
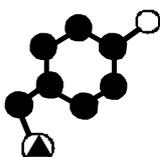
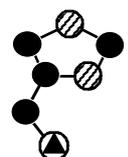
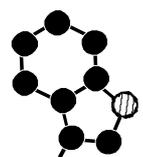
1. Введение.

Одной из нерешенных проблем современной молекулярной биологии является вопрос о природе канонического набора аминокислот. Только двадцать боковых цепей аминокислот, составляющих этот набор (табл.1), входят в состав белков и кодируются генетическим кодом [1-3]. Четверть века назад утверждалось [1], что ответить на этот вопрос пока нельзя. В настоящее время ситуация, практически, не изменилась, хотя отдельные аспекты этой проблемы частично решены.

Существует несколько подходов к классификации боковых цепей аминокислот. Наиболее часто применяемый и общепринятый подход рассматривает боковые цепи как радикалы, обладающие теми или иными физико-химическими свойствами [2]. Существует по крайней мере четыре группы боковых цепей аминокислот, представленные в таблице 1. Аминокислоты можно также классифицировать на основе учета протонодонорных и протоноакцепторных свойств боковых цепей [3]. Однако строгой химической классификации аминокислот провести не удастся. Большинство аминокислот, так или иначе, сочетают в себе свойства, принадлежащие другим группам.

В таблице 1 можно отметить две интересные особенности. Первая связана с тем, что большинство боковых цепей аминокислот представлена парами. Такими парами являются: Leu – Ile, Ser – Thr, Asp – Glu, Arg – Lys, Asn – Gln, Phe – Tyr. Аминокислоты Cys, Met и His, Trp, с некоторым приближением, также можно отнести к парным. Вторая особенность обусловлена наличием аминокислот, противоположных по своим свойствам. Наиболее четко она выражена для Asp⁻, Glu⁻ и Lys⁺, Arg⁺. Противоположные свойства можно найти и у других аминокислот. Эти особенности пока нашли рационального объяснения [2, 3].

Структура боковых цепей канонических аминокислот

Неполярные и слабо полярные аминокислоты			
			
Пролин (Pro)	Глицин (Gly)	Аланин (Ala)	Валин (Val)
			
Лейцин (Leu)	Изолейцин (Ile)	Серин (Ser)(-)	Треонин (Thr)(-)
Полярные аминокислоты			
			
Аспарагиновая кислота (Asp) (-)	Глутаминовая кислота (Glu)(-)	Лизин (Lys)(+)	Аргинин (Arg)(+)
Нейтральные и серосодержащие аминокислоты			
			
Аспарагин (Asn)	Глутамин (Gln)	Цистеин (Cys)(-)	Метионин (Met)
Циклические аминокислоты			
			
Фенилаланин (Phe)	Тирозин (Tyr)(-)	Гистидин (His)(+)	Триптофан (Trp)

Обозначения кружков: черные – углеродные атомы с различным числом атомов водорода (CH, CH₂, CH₃); белые малые – атомы кислорода (O, OH), белые большие – атомы серы (S, SH); заштрихованные – атомы азота (N, NH, NH₂, NH₃); с черным треугольником – α-углеродные атомы (CH).

Естественной системой аминокислот является генетический код. В рамках модели топологического кодирования цепных полимеров, частным случаем которых являются белки [3-5], кодируемые триплетами боковые цепи аминокислот рассматриваются в качестве физических операторов, воссоздающих закодированную структуру. При этом операторы связности (полярные аминокислоты) кодируются кодонами, содержащими во втором положении G и A, а операторы антисвязности (в основном, неполярные аминокислоты), кодируются триплетами с C и U во второй позиции.

Другой вариант классификации аминокислот, также основанный на генетическом коде, использует идею комплементарности аминокислот, кодируемых комплементарными триплетами [6 - 13]. Наиболее полно она была развита в работах Л.Б. Меклера и Р.Идлис [6]. Эти авторы считают, что аминокислоты, кодируемые антипараллельно комплементарными триплетами, должны быть комплементарны в структуре белка. Анализ аминокислот, связанных антипараллельной комплементарностью, привел к выделению среди них трех групп связности [6]. Однако между комплементарностью триплетов кода и контактами в белке кодируемых ими аминокислот нет однозначной связи, поэтому пока эта идея является скорее предметом веры ее авторов, чем доказанными представлениями.

Существует также классификация аминокислот, основанная на особенностях их функционирования в структуре белка. С позиции концепции систем сопряженных ионно-водородных связей (ССИВС) боковые цепи аминокислот подразделяются на два типа функциональных элементов – пассивные и активные [3, 14]. Пассивные элементы (Pro, Ala, и другие неполярные аминокислоты) создают гидрофобную среду для формирования и функционирования ССИВС. Активные элементы – аминокислоты в составе ССИВС, работают как функциональные модули.

Подробный анализ изложенных выше подходов дан в работах [3, 4]. В целом, однако, к проблеме систематизации аминокислот существует отношение как чему-то второстепенному и малозначимому.

Предложенные в 90-х годах прошлого века модели пространственной структуры генетического кода [15-17] описывают, в первую очередь, взаимоотношения дуплетов и триплетов кода. Однако, как нам представляется, система из двадцати канонических аминокислот может иметь и свою пространственную структуру, лишь косвенно связанную с кодом. Целью настоящей работы является попытка предложить модель структурной организации канонического набора аминокислот, которая была бы основана на физических категориях симметрии и антисимметрии.

2. Построение модели

2.1. Исходные предпосылки

Можно выделить следующие предпосылки к появлению этой модели:

1. *Совпадение числа канонических аминокислот белка с количеством вершин в додекаэдре.* Среди геометрических структур интерес представляют Платоновы тела – пять правильных выпуклых многогранников: тетраэдр, куб, октаэдр, икосаэдр и додекаэдр [18]. Одним из них является додекаэдр, имеющий 12 граней и 20 вершин. Число 20 привлекло наше внимание, поскольку оно совпадает с числом аминокислот в белке. Первые попытки построения модели структуры аминокислот на додекаэдре не привели, однако, к желаемому результату, поскольку существует большое число способов, на основе которых их можно было бы на нем расположить.

2. *Наличие у додекаэдра смежной структуры – икосаэдра.* Икосаэдр и додекаэдр являются парой, образование которой обусловлено

изоморфизмом их групп самосовмещений. Центры двадцати равносторонних треугольников, образующих грани икосаэдра, являются вершинами додекаэдра, а центры двенадцати правильных пятиугольников, составляющих грани додекаэдра, являются вершинами икосаэдра [19].

3. *Боковые цепи аминокислот связаны с процессом пространственно-временной самоорганизации белковой молекулы.* В работах [3,5] была предложена модель работы аминокислот в качестве физических операторов, осуществляющих первые этапы самоорганизации молекулы белка после ее синтеза. Эта модель находится в русле идеи о ко-трансляционном механизме свертывания белковой молекулы [20], находящей все больше опытных подтверждений. Для нас этот процесс важен тем, что он является направленным во времени и происходит от состояния с меньшей упорядоченностью к состоянию с более высоким порядком. Кроме того, на основе анализа структуры соответствий триплет – аминокислота было показано, что для воссоздания симметричных конформаций требуются разные по своей структуре боковые цепи аминокислот [3, 5]. Следовательно, идеи симметрии и асимметрии заложены уже в самих топологических основаниях генетического кода.

4. *Пространственные структуры, изоморфные булевым гиперкубам, могут использоваться для описания временных процессов.* В работе [21] был предложен метод описания временных процессов с помощью структур, изоморфных булевым гиперкубам, примененный к анализу математической структуры древнекитайского календаря. Детали метода будут рассмотрены ниже. Суть же состоит в том, что процесс перемещения по структуре булева гиперкуба от вершины с нулевыми значениями переменных к вершине с единичными значениями может служить моделью дискретных изменений во времени объекта и временным его описанием. Выделяемые при этом траектории были названы «меридиональными

циклами». Структура генетического кода изоморфна булевым гиперкубам B^4 и B^6 [22, 23], поэтому он является удобным объектом для применения этого подхода.

5. *Пространственной структурой меридиональных циклов дуплетного генетического кода является икосаэдр.* В результате анализа на структуре дуплетного генетического кода было выделено 12 меридиональных циклов, обладающих антисимметриями. Пространственной структурой, наиболее адекватно отразившей принципы антисимметрии этих циклов, оказался икосаэдр. А поскольку икосаэдр является фигурой, смежной с додекаэдром, то выявленные на икосаэдре принципы мы применили для организации аминокислот на додекаэдре.

2.2. Меридиональные циклы как модель для описания пространственно-временных процессов.

2.1.1. Введение понятия «меридиональный цикл»

Рассмотрим какой-либо процесс изменений структуры. Таким процессом может быть, например, формирование ближнего порядка в структуре белка, происходящее непосредственно после синтеза. Формальной моделью этого процесса может служить последовательное изменение конформаций 4х-звенного графа [4, 5]. На рис. 1 показаны три стадии таких изменений, начиная от состояния (а), в котором присутствуют только структурные ребра (аналоги ковалентных связей), к стадиям (б) и (в) с одним и двумя ребрами связности (аналогами нековалентных взаимодействий, фиксирующих конформацию белка) [3-5]. Эти изменения описываются соответствующими матрицами из шести переменных, расположенными под графами. Для белка этот процесс происходит, очевидно, во времени, поэтому в модель включены параметры t_1 и t_2 , соответствующие временам перехода стадий $a \rightarrow б$ и $б \rightarrow в$.

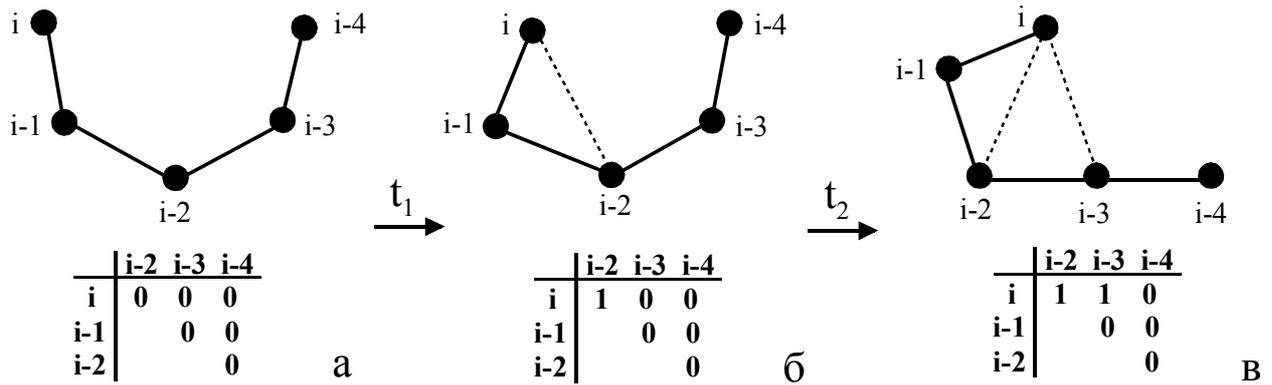


Рис. 1. Стадии изменения конформации 4х-звенного графа и их описание с помощью матриц из 6-ти переменных.

Последовательный переход 4х-звенного графа от несвязного (нулевого) состояния, описываемого матрицей с шестью нулями к полностью связному (единичному) состоянию, которое описывается матрицей из шести единиц, осуществляется за шесть стадий (рис.2,а и 2,б). Время перехода от одной стадии к другой может иметь величину t_i , и для отдельных стадий может быть записано как $t_1, t_2, t_3 \dots t_6$. Если эти интервалы времени обозначить как единичные, и поставить в соответствие $x_1 \rightarrow t_1, x_2 \rightarrow t_2, \dots, x_6 \rightarrow t_6$, то матричное описание структурных изменений (рис. 2,б) можно рассматривать и как описание этого процесса во времени.

После достижения единичного состояния может происходить обратный процесс возврата к нулевому состоянию. Однако следует обратить внимание, что по отношению к исходному состоянию этот процесс является антисимметричным. В матрицах слева (рис. 1,в) изменения происходят сверху вниз от нулевого значения к единичному, а в матрицах справа, наоборот, снизу вверх и от единичного значения к нулевому. В результате (схема 1) за время:

$$T = 2 \sum_{i=1}^{i=6} t_i \quad (1)$$

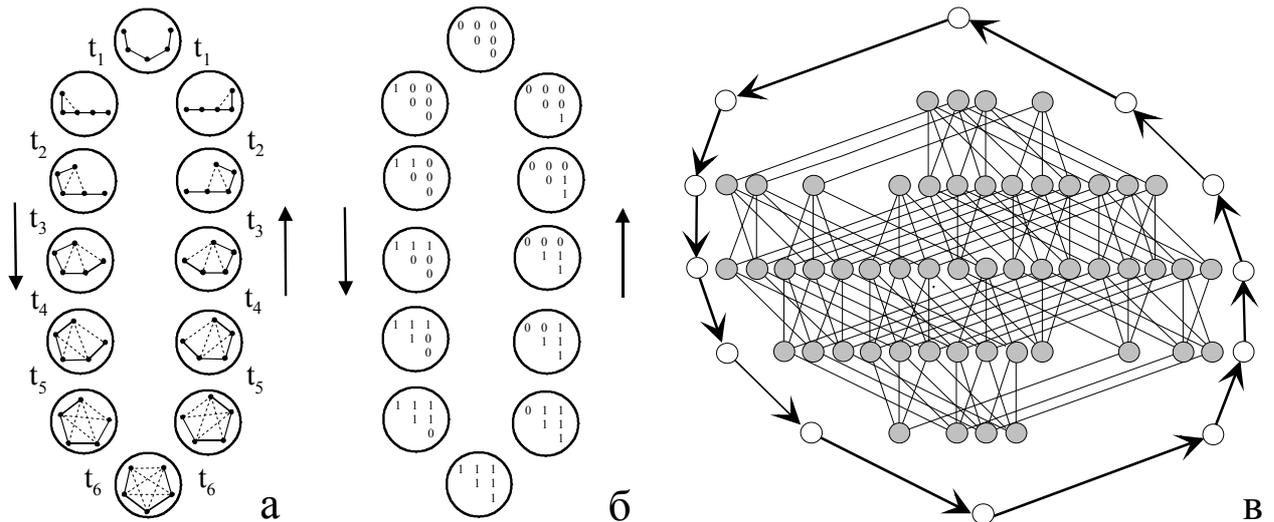


Рис. 2. Введение понятия меридионального цикла.

- а – полный цикл изменений конформации 4х-звнного графа;
- б – матричное описание цикла (структурное и временное);
- в – пример расположения меридионального на гиперкубе B^6 .

граф совершит полный цикл своих структурных изменений.

Для 4х-звенного графа возможно 64 конформации, описываемые булевым гиперкубом B^6 [4, 5]. То же можно сказать и об единичных интервалах времени. Процесс перехода на гиперкубе от нулевого состояния к единичному может совершаться различными путями. На рис. 2,в описанный выше цикл расположен по периферии гиперкуба B^6 . При этом, поскольку такой переход совершен, то все другие возможности перемещений прекращаются, что показано на рис. 2,в как разрыв связей выделенного цикла с остальными переменными гиперкуба. Возможны и иные маршруты формирования циклов на гиперкубе [18]. Общее их число для B^6 равно шестидесяти [21], однако все они будут проходить от верхней вершины к нижней в меридиональном направлении. Учитывая это обстоятельство, мы предложили называть их меридиональными циклами (сокращенно – М-циклы) [21]. В свете сказанного выше, М-циклы, могут служить как для описания дискретных стадийных изменений структуры,

так и для описания соответствующих этим стадиям дискретных интервалов времени.

2.1.2. Меридиональные циклы булева гиперкуба B^4 и дуплетного генетического кода

Известно, что вследствие вырожденности триплетов генетический код может быть представлен как дуплетный [24]. Пространственная структура последнего описывается булевым гиперкубом B^4 [21, 25]. Вариант этой структуры представлен на рис. 3. Интерес к структуре дуплетного кода определяется тем, что его структура является четырехмерной и может служить аналогом пространственно-временных отношений (три измерения соответствуют пространственным координатам и четвертое – время).

В вершинах этой структуры помещены дуплеты оснований, связанные единичными переходами. Вверху справа рядом с вершинами записаны булевы переменные гиперкуба, на основе принятого в работах [3-5] соответствия (схема 2) :

$$C = 00 \quad U = 01 \quad G = 10 \quad A = 11 \quad (2),$$

а внизу справа в трехбуквенном обозначении приведены соответствующие этим дуплетам аминокислоты.

Первоначальной задачей нашего исследования был анализ и систематизация М-циклов. Для общности, этот анализ мы сначала проведем непосредственно с переменными булева гиперкуба B^4 , а затем распространим найденные закономерности на дуплетный код.

Как видно на рисунке 3, гиперкуб B^4 содержит пять ярусов. В первом и пятом ярусе находится по одной вершине. В первом ярусе этой вершине соответствует число, состоящее из четырех нулей, в пятом число – из четырех единиц. При этом каждое число, занимающее одну из вершин, связано однобитовым переходом с четырьмя соседними вершинами. Числа

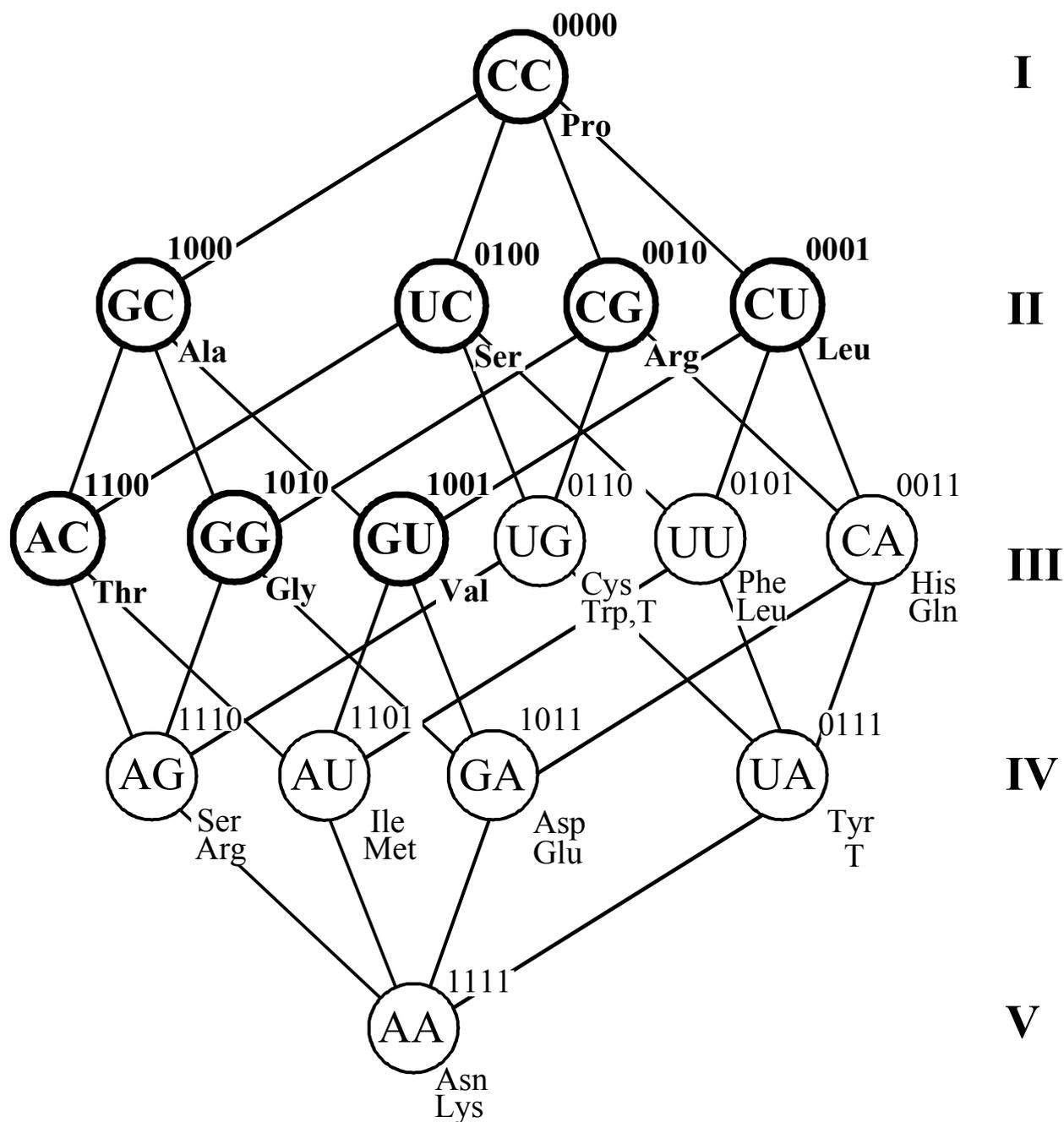


Рис. 3. Пространственная структура дуплетного генетического кода, изоморфная булевой гиперкубу $В^4$.

I – V – обозначения ярусов; T – терминация считывания.

на ярусах расположены справа и слева симметрично, например: 0001 и 1000, 0010 и 0100 и т.д. Во втором и четвертом ярусах эти числа занимают

четыре вершины гиперкуба, в третьем – шесть. Верхние восемь чисел, выделенные жирным шрифтом (они соответствуют восьми дуплетам, кодирующим по одной аминокислоте), могут быть трансформированы в нижние путем преобразования (схема 3):

$$0 \leftrightarrow 1 \quad (3)$$

Преобразование (3) называется преобразованием антисимметрии. Связанные им числа располагаются симметрично (группа симметрии C_2). Перемещение от числа 0000 до числа 1111 и симметричным маршрутом в обратном направлении составляет М-цикл. Мы провели рассмотрение всех вариантов М-циклов, которых оказалось 12. В таблице 2 показаны значения переменных, через которые проходят траектории циклов.

Из этой таблицы видно, что правая и левая ветви всех циклов проходят через симметричные значения переменных, например, для 1_1 : 0100 – 0010, 1100 – 0011, 1110 – 0111. В центральном столбце расположены 4 цикла, в которых числа верхней части одной ветви, включая центральную, антисимметричны числам нижней части противоположной ветви, например, для цикла 1: 1000 – 0111, 1100 – 0011, 0001 – 1110. В правом и левом столбцах располагаются антисимметричные структуры, у которых верхняя пара чисел анти-симметрична нижней паре, но перевернутой на 180° . Например, для цикла 1_1 верхняя пара представлена числами 0100 – 0010, а нижняя пара цикла 1_1 – числами 1101 – 1011. Сопоставим левое число из верхней пары с правым числом из нижней пары, а правое число из верхней пары – с левым числом из нижней пары: 0100 \leftrightarrow 1011 и 0010 \leftrightarrow 1101. Из сопоставления видно, что эти числа являются антисимметричными. Мы называем такую антисимметрию инверсной.

Кроме того, обнаружился еще один тип таких отношений, который можно назвать антисимметрией «антиподов». «Антиподами» мы называем такие пары М-циклов, у которых нет ни одной одинаковой пары чисел,

Состав переменных в меридиональных циклах булева гиперкуба B^4 .

Типы антисимметрии		
Инверсная (-)	Внутренняя антисимметрия	Инверсная (+)
0000 0100 0010 1100 0011 1110 0111 1111 1_1	0000 1000 0001 1100 0011 1110 0111 1111 1	0000 1000 0001 1100 0011 1101 1011 1111 1_1
0000 0100 0010 0110 0111 1110 1111 2_1	0000 1000 0001 1010 0101 1110 0111 1111 2	0000 1000 0001 1001 1011 1101 1111 1_2
Антисимметрия «антиподов»		
0000 1000 0001 1010 0101 1011 1101 1111 -1_1	0000 0100 0010 0101 1010 1101 1011 1111 -1	0000 0100 0010 0101 1010 0111 1110 1111 -1_1
0000 1000 0001 1001 1101 1011 1111 -2_1	0000 0100 0010 1100 0011 1101 1011 1111 -2	0000 0100 0010 0110 1110 0111 1111 -1_2

кроме пар начальных и конечных точек. Например, антиподами являются циклы в паре 1_1 и -1_1 . В самом деле, сравним пары чисел в этих циклах (схема 4):

1_1	-1_1
0000	0000
0100 0010	1000 0001
1100 0011	1010 0101
1110 0111	1011 1101
1111	1111

(4)

Действительно, ни одна пара чисел в этих циклах не совпадает. При этом они не являются и инверсно антисимметричными циклами, поскольку центральные пары не являются таковыми. «Антиподами» являются пары: 1 и -1 , 2 и -2 , 1_1 и -1_1 , 2_1 и -2_1 , 1^1 и -1^1 , 2^1 и -2^1 .

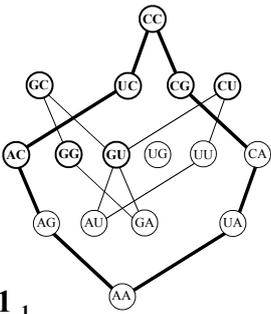
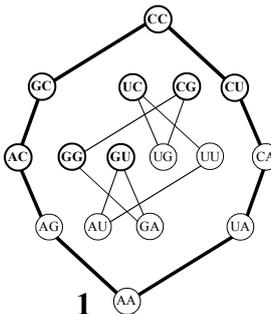
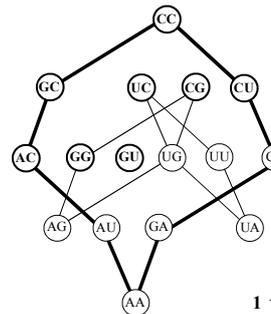
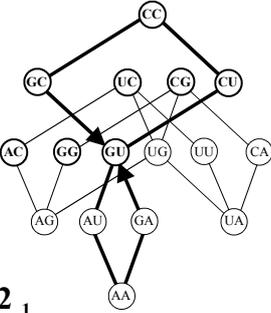
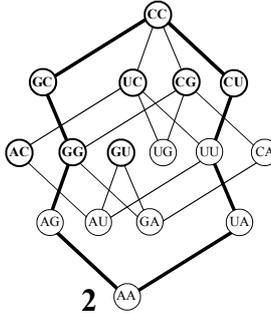
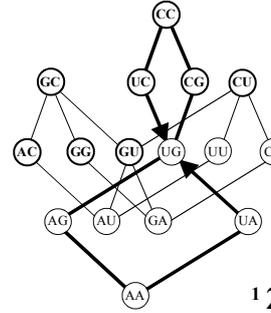
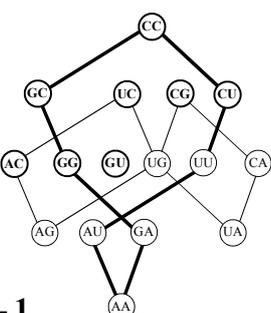
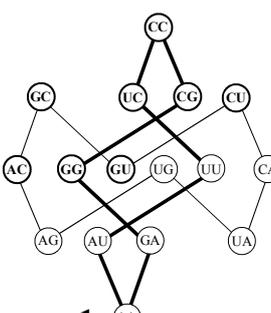
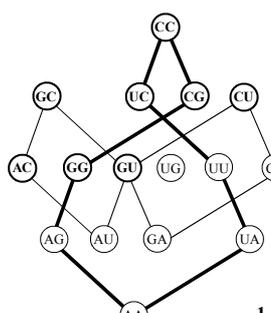
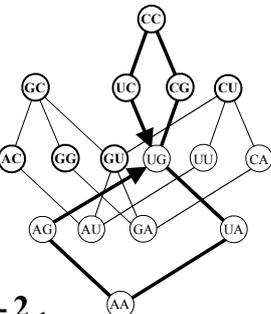
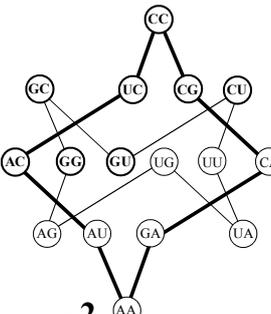
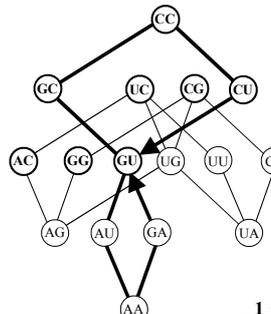
Перейдем теперь к рассмотрению M-циклов, выделенных на структуре дуплетного генетического кода. В таблице 3 показаны топологии полученных циклов, в которых пары чисел закодированы буквами, в соответствии со схемой 2. Они в точности воспроизводят численное описание полученных циклов. При этом симметричные значения кодируют буквы С и С, А и А, U и G, а антисимметричные – С и А, G и U. Последние пары соответствует правилу Румера [24].

В таблице 3 можно выделить три топологических группы:

- а) внутренне симметричные циклы $-1, 2, -1, -2$;
- б) полностью эквивалентные циклы $-2_1, {}^12, -2_1$ и $-{}^12$;
- в) циклы с низкой симметрией $1_1, {}^11, -1_1, -{}^11$.

Однако в данной таблице они расположены не в соответствии с принадлежностью к своей топологической группе, а на основе симметрии. Так, внутренне симметричные циклы, как и таблице 2, были помещены в центральные клетки. Пары эквивалентных и низко симметричных циклов находятся симметрично относительно центральных циклов (инверсная антисимметрия). Циклы-антиподы из второй группы заняли вторую и четвертую строки таблицы, а из третьей группы – первую и третью строки.

Топологии меридиональных циклов дуплетного генетического кода,
построенного на гиперкубе B^4 .

Типы антисимметрии		
Инверсная (-)	Внутренняя антисимметрия	Инверсная (+)
 <p>1₁</p>	 <p>1</p>	 <p>1¹</p>
 <p>2₁</p>	 <p>2</p>	 <p>2¹</p>
Антисимметрия «антиподов»		
 <p>-1₁</p>	 <p>-1</p>	 <p>-1¹</p>
 <p>-2₁</p>	 <p>-2</p>	 <p>-2¹</p>

Из таблицы 3 видно, что в пары циклов 2_1 и -2_1 , а также 1_2 и $^{-1}_2$ при внешнем сходстве, описывают разные маршруты. В М-циклах 2_1 и 1_2 маршруты проходят только по левой стороне, а затем только по правой, поэтому стрелки в общей точке находятся с разных сторон. В то же время, в цикле -2_1 маршрут переходит с левой стороны на правую, и поэтому стрелки в общей точке находятся слева, а в цикле $^{-1}_2$ маршрут переходит с правой стороны на левую, поэтому стрелки в точке находятся справа.

2.3. Икосаэдр как пространственная структура для М-циклов дуплетного генетического кода

Структурой, в вершинах которой можно расположить все М-циклы дуплетного кода, является, очевидно, икосаэдр, имеющий 12 вершин. В процессе расположения их на икосаэдре мы сохранили те отношения антисимметрии, которые были отмечены в таблицах 2 и 3 (рис. 4).

Полученный порядок расположения циклов на структуре икосаэдра может быть описан путем введения двух плоскостей антисимметрии. Первую плоскость можно провести перпендикулярно плоскости листа (отмечена на рисунке цифрой I), и назвать плоскостью инверсной антисимметрии. Меридиональные циклы, через которые проходит эта плоскость, являются внутренне антисимметричными, расположенными в центре табл.2 и 3. Плоскость инверсной антисимметрии разделяет циклы на две группы, в которых пары антисимметричных циклов связаны между собой преобразованием инверсии (вращением на 180°).

Вторую плоскость можно провести перпендикулярно первой и расположить в плоскости листа (отмечена цифрой II). На ней находится ось, при повороте вокруг которой на 180° вершины икосаэдра, расположенные перед этой плоскостью совмещаются с вершинами, находящимися под плоскостью. При этом происходит совмещение

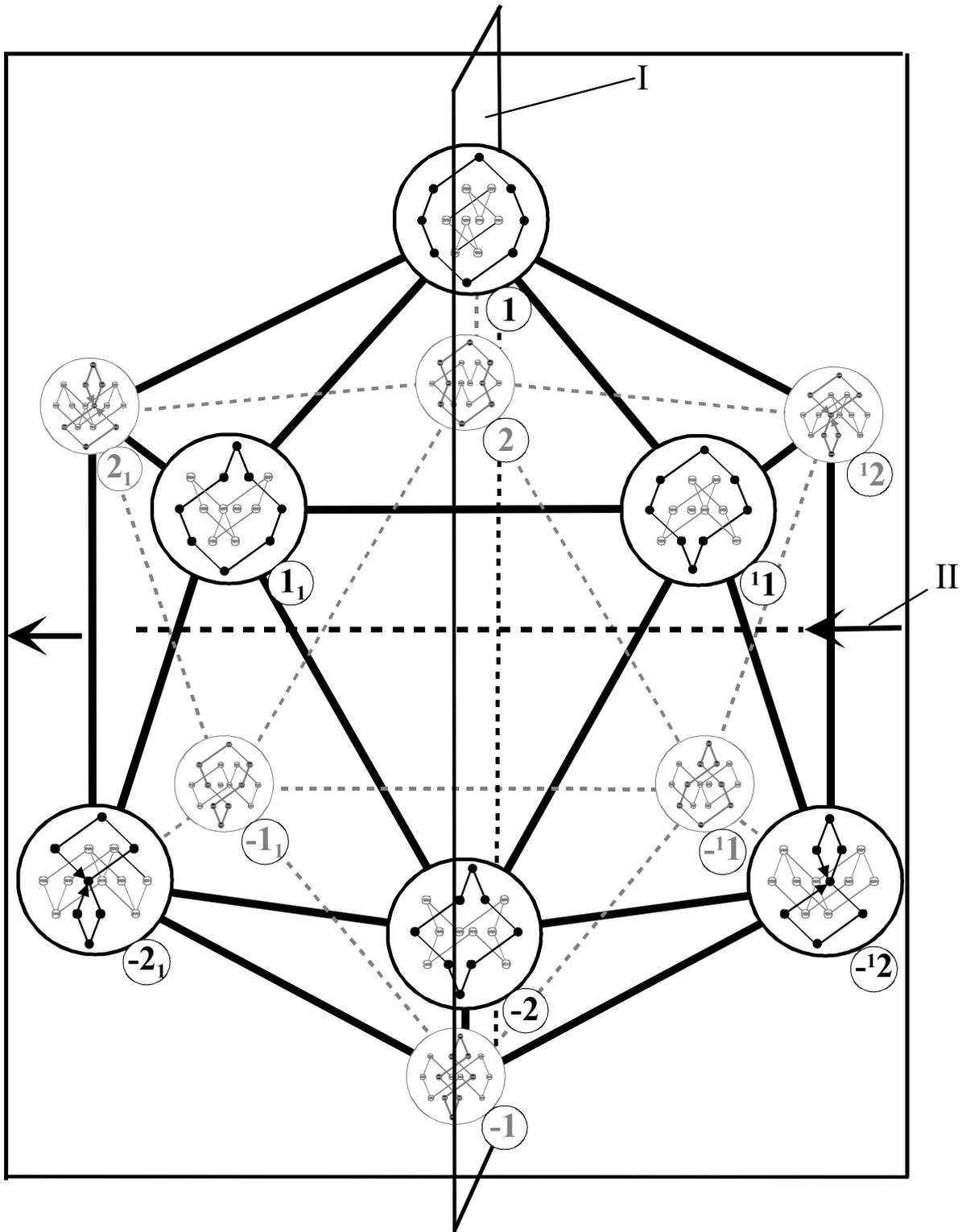


Рис. 4. Порядок расположения меридиональных циклов генетического кода на икосаэдре.

I - плоскость инверсной антисимметрии;

II - ось вращения и плоскость антисимметрии "антиподов".

меридиональных циклов «антиподов». Например, цикл 1 совпадает с циклом -1 , цикл 2 – с циклом -2 и т.д. Таким образом, эта плоскость обеспечивает преобразования группы меридиональных циклов в группу «антиподов» при ее вращении на 180° .

2.4. Додекаэдр как пространственная структура для системы аминокислот

2.4.1. Возможные групповые свойства у структур, расположенных в вершинах додекаэдра

Как мы упоминали в разделе 2.1., п.2, структурой, смежной с икосаэдром, является додекаэдр. В поисках принципов расположения 20 аминокислот в вершинах додекаэдра мы предприняли попытку перенести плоскости антисимметрии, выявленные для М-циклов на икосаэдре, на додекаэдр. Связь аминокислот с М-циклами можно увидеть в том, что последние, по определению, представляют собой пространственно-временные описания изменений структуры 4х-звенного графа, а боковые цепи аминокислот выступают в роли физических операторов, воссоздающих закодированные структуры 4х-звенного графа.

Мы выделили на додекаэдре (рис. 4) две плоскости антисимметрии – плоскость инверсной антисимметрии (I) и плоскость антисимметрии антиподов (II) и попытались представить, в соответствии с какими свойствами должны быть расположены структуры в вершинах додекаэдра. В плоскости инверсной антисимметрии должны располагаться «внутренне антисимметричные» структуры, и связанные с ними антисимметричные антиподы – это структуры А и В, а также $(-A)$ и $(-B)$. По бокам от каждой такой структуры должны находиться две пары инверсно антисимметричных структур. Например, для А это структуры A_1 и 1A , а также A_2 и 2A (индексы справа внизу и слева вверху). Кроме того, к этим

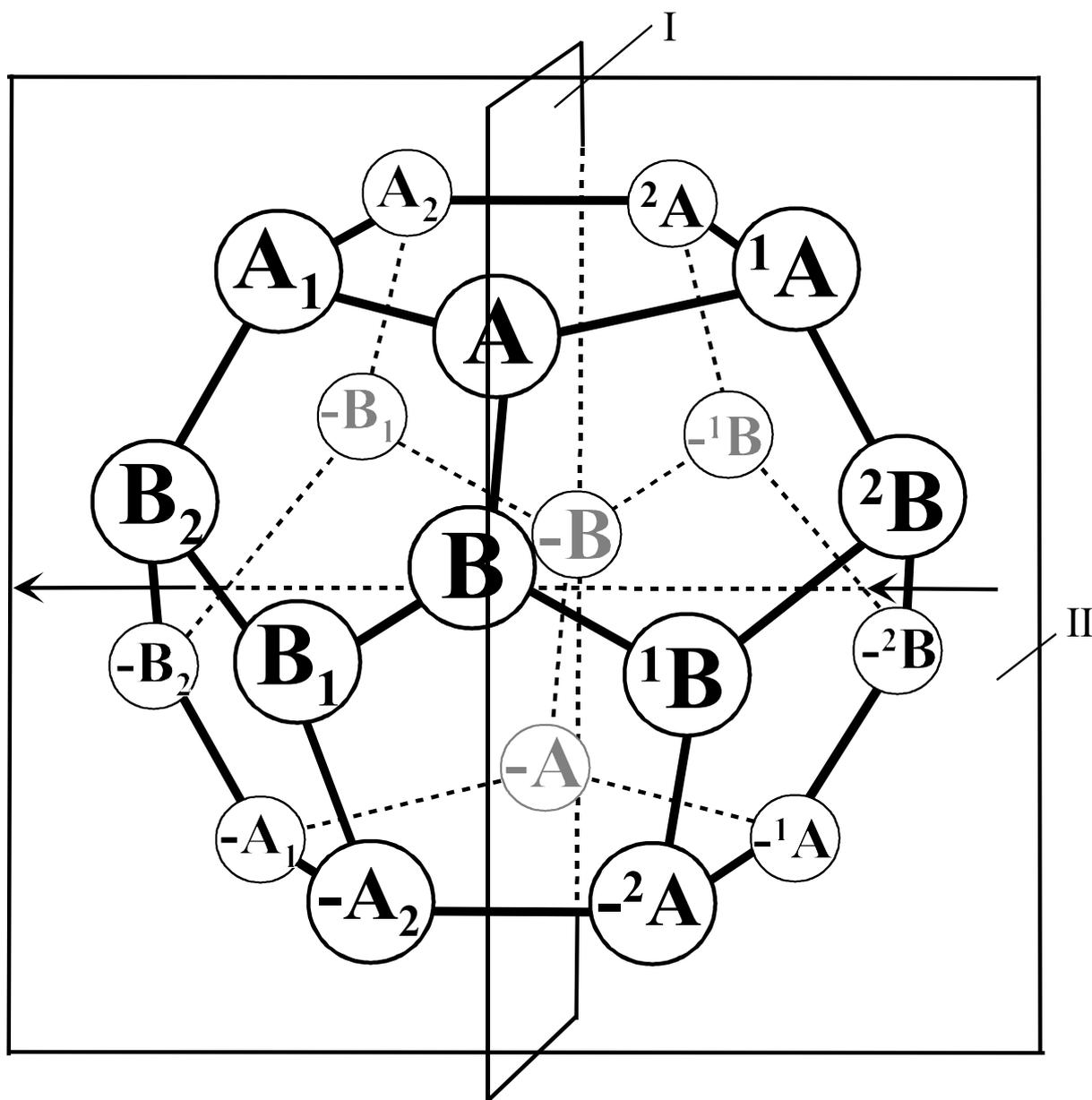


Рис. 4. Выделение возможных групп аминокислот для построения из них системы на додекаэдре.

I - плоскость инверсной антисимметрии;

II - ось вращения и плоскость антисимметрии "антиподов".

двум парам должны существовать еще две пары структур-антиподов ($-A_1$) и (${}^{-1}A$), а также ($-A_2$) и (${}^{-2}A$), расположенных по сторонам от структуры $-A$. Аналогичные рассуждения можно провести для структур B и $-B$. Таким

образом, если данная схема подходит для боковых цепей аминокислот, то в наборе из 20 аминокислот должны иметься четыре внутренне антисимметричных аминокислоты и восемь пар инверсно антисимметричных. В свою очередь, инверсно антисимметричные аминокислоты распадаются на пары антисимметричных антиподов (восемь пар). Прежде чем переходить к переносу этой схемы на аминокислоты, мы приняли во внимание еще ряд соображений, связанных с характером циклов, расположенных в вершинах икосаэдра.

2.4.2. Анализ М-циклов в треугольниках, образующих грани икосаэдра

В центрах граней икосаэдра формируются вершины додекаэдра. В вершинах треугольников, образующих грани икосаэдра, расположены М-циклы, состав которых может, в какой-то мере, определять характер структур, строящихся на додекаэдре. Поэтому анализ М-циклов может служить дополнительной подсказкой при выборе этих структур. Для удобства анализа все триады М-циклов приведены в таблице 5.

Из таблицы следует, что на структуре икосаэдра имеется две пары граней, на которых могут располагаться внутренне антисимметричные структуры (столбец «внутренняя антисимметрия»). В частности, триады, обозначенные буквами А и В, а также $-А$ и $-В$, содержат в своем составе пары инверсно антисимметричных циклов, влияние которых на выбор помещаемой на грани структуры будет взаимно уравновешено. Конкретно, в триадах А и В инверсными циклами являются 1_1 и 11 , а в триадах $(-А)$ и $(-В)$ – циклы (-1_1) и $(-^11)$. Вследствие этого структуры, детерминированные этими триадами, должны напоминать внешне оставшиеся непарные циклы. Так структура А, содержащая непарный цикл 1_1 , должна быть, по-видимому, циклической, а структура $(-А)$, антипод к ней (цикл -1_1) – наиболее сжатой по размеру (точечной). В применении к аминокислотам

Триады меридиональных циклов, локализованных в вершинах
треугольников, образующих грани икосаэдра

Типы антисимметрии				
Инверстная (-)	Внутренняя анти- симметрия		Инверсная (+)	
Антисимметричные «антиподы»				

это могут быть, соответственно, пролин и глицин. При этом, поскольку цикл 1 входит во все триады блока, обозначенного буквами А, то к нему, вероятно, можно приписать все циклические аминокислоты (фенилаланин, тирозин, гистидин и триптофан). В то же время, во все триады блока, обозначенного буквами (-А), входит цикл -1, поэтому к нему можно приписать аминокислоты, размер которых является наименьшим (серин, треонин, цистеин и метионин).

Структура В, вероятно, должна быть «с рожками» (цикл -2), а антипод ней ($-B$), содержащий цикл 2, должен иметь форму гантели. Такими структурами могут быть валин и аланин (см. табл. 1). Оставшиеся восемь аминокислот могут располагаться в центральной зоне додекаэдра. Из них две аминокислоты – глутаминовая и аспрагиновая, отрицательно заряженные, могут иметь в качестве антиподов положительно заряженные боковые цепи лизина и аргинина. Остаются лишь аспарагин и глутамин, антиподами к которым будут лейцин и изолейцин. Итак, кандидаты на роль «внутренне антисимметричных» аминокислот определились. Ясен также состав двух крайних и двух центральных блоков-антиподов.

Дополнительными подсказками могут служить следующие особенности циклов, расположенных на икосаэдре. Триады циклов, находящиеся в левом столбце (триады A_1, A_2, B_1, B_2), содержат М-циклы, у большинства которых центры тяжести смещены в нижнюю часть, а в правом столбце (триады ${}^1A, {}^2A, {}^1B, {}^2B$) – в верхнюю часть цикла. Можно предполагать, что это указывает на меньшую длину структур, располагающихся слева и большую – справа. Для антисимметричных антиподов, наоборот, в триадах, располагающихся слева (триады $-A_1, -A_2, -B_1, -B_2$), центры тяжести циклов смещены в верхнюю часть, а справа – в нижнюю часть циклов (триады $-{}^1A, -{}^2A, -{}^1B, -{}^2B$). Это может указывать на большую длину структур, располагающихся слева. При построении системы аминокислот мы приняли эти «подсказки» во внимание.

3. Результаты

3.1. Таблица антисимметрий боковых цепей аминокислот

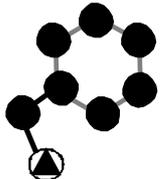
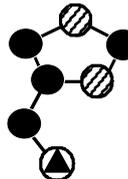
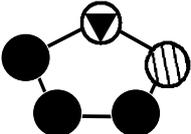
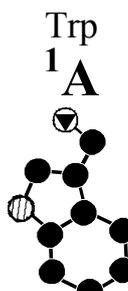
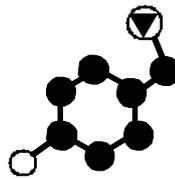
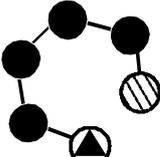
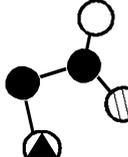
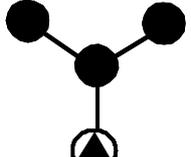
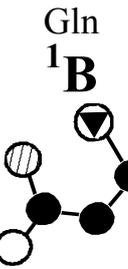
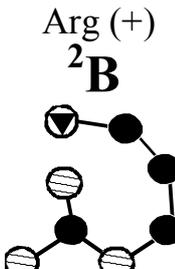
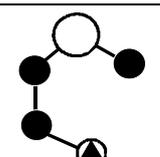
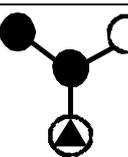
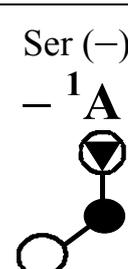
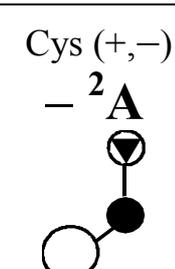
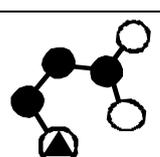
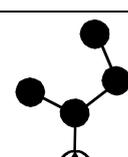
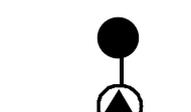
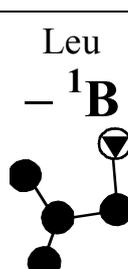
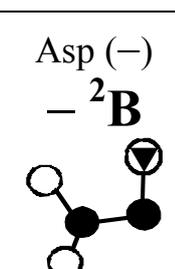
Приведенные выше представления об антисимметрии структур, связанных с М-циклами, мы применили к боковым цепям канонических

аминокислот. Результаты этой работы приведены в таблице 6. Буквенные индексы под названиями аминокислот соответствуют рис.4.

Из таблицы 6 видно, что все 20 боковых цепей аминокислот могут быть разделены на четыре группы (группы пролина, валина, глицина и аланина), выделенные на основе внутренне антисимметричных аминокислот, расположенных в центре. Справа и слева от центральных боковых цепей расположены инверсно антисимметричные аминокислоты. Положение α -углеродных атомов у них противоположное: слева – внизу, а справа – вверху. С учетом соображений, изложенных в предыдущем разделе, в группе пролина и валина слева расположены боковые цепи аминокислот меньшего размера (His, Phe, Asn, Lys), а справа, сходные по свойствам, но большего размера (Trp, Tyr, Gln, Arg). В группах глицина и аланина, антиподах к предыдущим группам, аминокислоты находятся в обратном порядке: слева большего размера (Thr, Met, Ile, Glu), а справа – меньшего (Ser, Cys, Leu, Asp). В результате пары аминокислот-«антиподов» в сумме обладают наименьшей молекулярной массой, в чем нетрудно убедиться. Сравните, например, His – Thr и Trp – Ser, Phe – Met и Tyr – Cys.

Отметим, что боковые цепи аминокислот-«антиподов» (табл. 6) имеют противоположные физико-химические свойства. Например, Lys (+), протоноакцептор, противостоит Glu (-), протонодонору, Arg (+) противостоит Asp (-), His (+) – Thr (-). Нейтральные доноро-акцепторы (Asn, Gln), являются антиподами неполярных аминокислот (Ile, Leu). Боковая цепь тирозина, проявляющего свойства донора-акцептора (+, -), является антиподом цистеина (+, -), а циклического Phe – ациклического Met. Любопытно, что боковые цепи Ile и Leu полностью идентичны (табл.6). Они отличаются лишь положением α -углеродного атома: у Ile он внизу, а у Leu – вверху.

Система антисимметрий боковых цепей канонических аминокислот

Типы антисимметрии				
Инверсная (-)	Внутренняя антисимметрия		Инверсная (+)	
Группа пролина				
 Phe A_2	 His(+) A_1	 Pro A	 Trp 1A	 Tyr (+,-) 2A
Группа валина				
 Lys (+) B_2	 Asn B_1	 Val B	 Gln 1B	 Arg (+) 2B
Антисимметричные «антиподы»				
Группа глицина				
 Met $-A^2$	 Thr (-) $-A^1$	 Gly $-A$	 Ser (-) $-^1A$	 Cys (+,-) $-^2A$
Группа аланина				
 Glu (-) $-B_2$	 Ile $-B_1$	 Ala $-B$	 Leu $-^1B$	 Asp (-) $-^2B$

3.2. Система канонических аминокислот, построенная на додекаэдре

Конечным результатом нашей работы явилось построение системы из 20 канонических аминокислот на додекаэдре (рис. 5). Как следует из рисунка, группы аминокислот пролина и глицина расположились на додекаэдре в виде двух циклов, а группы валина и аланина – в центре, с образованием сплошного кольца. Отмеченные в таблице 6 особенности сохранены. В плоскости I расположены «внутренне антисимметричные» аминокислоты – Pro, Val и их «антиподы» – Gly, Ala. Аминокислоты, связанные отношением инверсной антисимметрии, располагаются симметрично относительно плоскости I и перевернуты по отношению друг к другу. Например, His – Trp, Lys – Arg, Asn – Gln и т.д. Аминокислоты, расположенные перед плоскостью II (они более крупного размера), после поворота плоскости на 180° , совмещаются со своими «антиподами», расположенными за плоскостью II (они меньшего размера). Так, His совпадает с Thr, Phe с Met, Trp образует пару с Ser, Tyr с Cys. Противоположно заряженные аминокислоты также образуют пары: Lys – Glu, Arg – Asp, а неполярные гидрофильные аминокислоты образовали пары с неполярными гидрофобными: Gln – Leu, Asn – Ile. Однако, система аминокислот на додекаэдре имеет ряд дополнительных особенностей.

Во-первых, аминокислоты образуют замкнутые или почти замкнутые циклы. Так, все циклические аминокислоты (His, Phe, Tyr, Trp) сгруппировались вокруг пролина, а слабо полярные с атомами кислорода и серы (Ser, Cys, Met, Thr) – вокруг глицина. Аминокислоты, содержащие гетероатомы – N, O, S, расположились чуть выше и также сформировали два цикла. Левый цикл содержит Thr, Glu, Lys, Asn, Met, а правый – Ser, Asp, Arg, Gln, Cys. В циклах-«антиподах», расположенных в центральной области додекаэдра, преобладают нейтральные аминокислоты (Asn, Met, Gln, Val и Ile, Ala, Leu, Phe), за исключением, соответственно, Cys и Tyr.

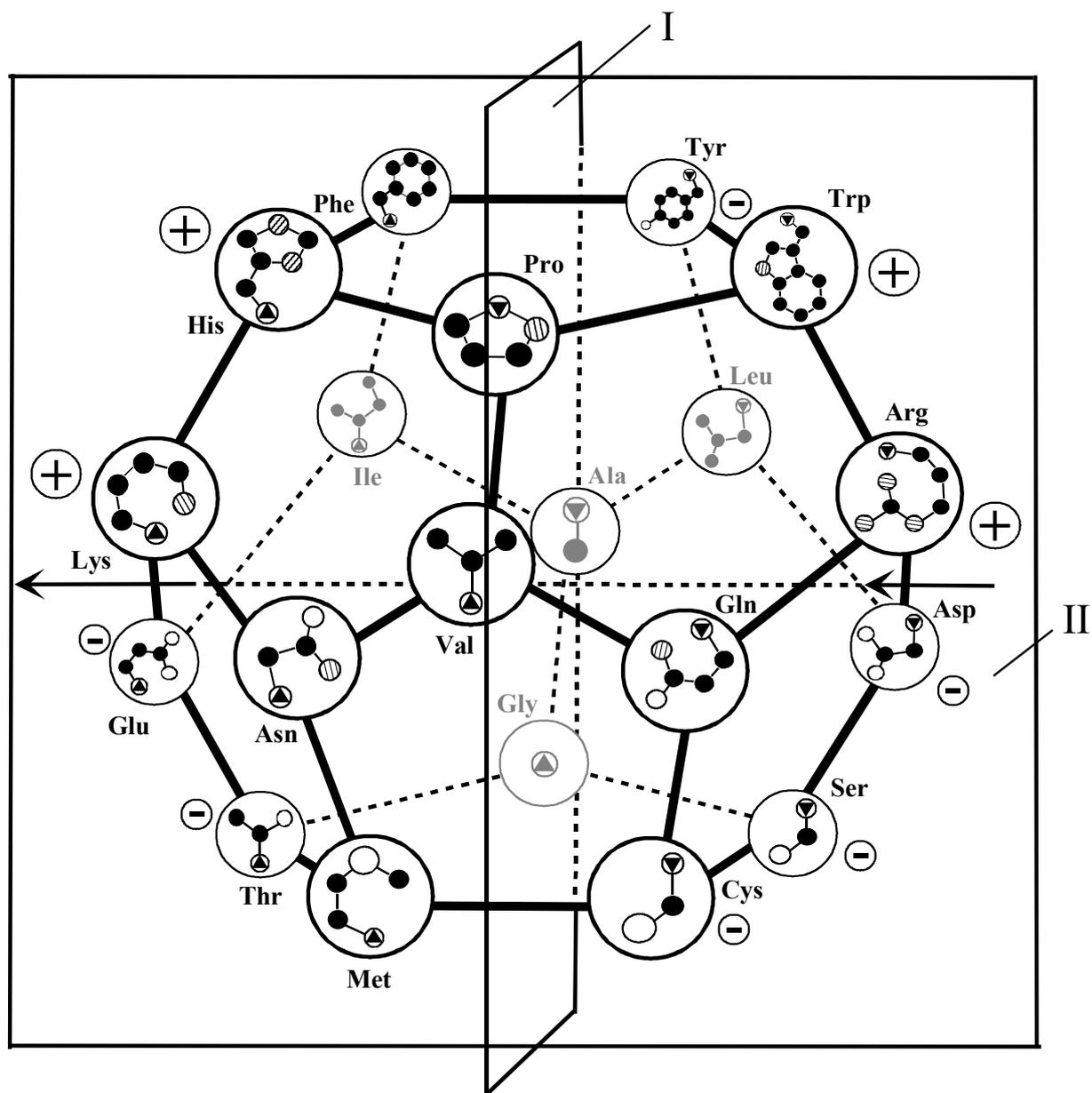


Рис. 5. Система аминокислот, связанная отношениями антисимметрии, построенная на додекаэдре

I - плоскость инверсной антисимметрии;

II - ось вращения и плоскость антисимметрии "антиподов".

Во вторых, при движении по периферии додекаэдра к центральной его области усиливается степень полярности боковых цепей. В верхней части это Pro, His, Lys и Pro, Trp, Arg, а в нижней – Gly, Thr, Glu и Gly, Ser, Asp.

В третьих, при перемещении от вершины к вершине наблюдаются наименьшие из возможных изменения в структуре аминокислот. При этом следует отметить, что такие изменения для инверсно антисимметричных аминокислот, в основном, совпадают, что хорошо видно из табл. 7.

4. Обсуждение

Можно задать справедливый вопрос: а что, в сущности, меняет построенная система боковых цепей аминокислот на додекаэдре? Очевидно, что классифицировать можно аминокислоты и без таких сложных построений. Тем не менее, полученный результат (рис.5) является скорее неожиданным, чем тривиальным. Во всяком случае, среди рассмотренных во введении подходов к классификации боковых цепей аминокислот [2,3, 6, 16] такой результат не просматривается.

Логично высказать предположение, что таких «классификаций» на додекаэдре можно построить много. Однако анализ показывает, что это не так. В самом деле, все циклические аминокислоты расположились в верхней части додекаэдра и их связь с циклической структурой пролина кажется очевидной. В то же время их антиподы – аминокислоты с короткой цепью (Ser, Thr, и т.д.) также образуют естественную группу, родоначальником которой может быть глицин. Аминокислоты вышеупомянутых групп можно отнести к боковым цепям 1-го ранга, связанных с самыми простыми аминокислотами – Pro и Gly. Расположение боковых цепей внутри этих групп относительно плоскости инверсной антисимметрии достаточно однозначно.

То же можно сказать и об аминокислотах центральной области додекаэдра (боковые цепи 2-го ранга), которые связаны с более сложными исходными боковыми цепями – Ala, Val. Принадлежность боковых цепей Lys и Arg, Asp и Glu, Asn и Gln, а также Ile и Leu к разным группам

Единичные изменения между аминокислотами, занимающими соседние вершины додекаэдра

Типы антисимметрии				
Инверсная (-)		Внутренняя антисимметрия	Инверсная (+)	
Phe +2CH ₂		Pro (-) CH ₂ ,		Tyr + 2CH ₂ , +OH
	His + CH ₂ , C→N		Trp + CH ₂ , +(CH ₂) ₄	
Lys (-) 2CH ₃ + (CH ₂) ₃ NH ₂	Asn (-) 2CH ₃ +H ₂ N-C=O	Val	Gln (-) 2CH ₃ + CH ₂ , +H ₂ N-C=O	Arg (-) 2CH ₃ , + CH ₂ , CH ₂ N-C=N NH ₂
Антисимметричные «антиподы»				
Met + CH ₂ , +CH ₂ S-CH ₃		Gly +CH ₃		Cys +CH ₂ SH
	Thr + CH ₂ , +CH ₂ OH		Ser +CH ₂ OH	
Glu + CH ₂ + O=C-OH	Ile + CH ₃ + CH ₂ CH ₃	Ala	Leu +CH ₃ CH ₂ CH ₃	Asp + O=C-OH

инверсной антисимметрии очевидна, также как и принадлежность пар Lys – Arg и Asp – Glu к разным группам «антиподов».

Наличие пентагональных циклов в использованной для построения системы аминокислот структуре додекаэдра указывает на ее связь с 4х-звенным фрагментом структуры белка. Этот фрагмент по форме близок к

пятиугольнику. Как было показано в работах [3-5], именно конформации этого фрагмента являются топологическим основанием генетического кода. От расположения на нем боковых цепей аминокислот будет зависеть, сложится ли этот фрагмент в α -спираль, β -структуру или что-либо иное. Не исключено, что использование этой структуры в дальнейшем позволит разработать новый алгоритм для предсказания структуры белков.

Представляет также интерес, нельзя ли с помощью полученной системы аминокислот или каким-либо иным образом получить систему соответствий триплет-аминокислота. Некоторые указания на природу этих соответствий были получены в рамках модели топологического кодирования белков [3-5], однако хотелось бы получить этот результат и независимым путем.

Заслуживает обсуждения вопрос, а всю ли информацию мы «извлекли» из полученного результата. Так, поскольку данная система возникла под идейным влиянием меридиональных циклов, связанных с описанием временных процессов [21], возникает вопрос: не являются ли эти боковые цепи системой модулей с различной временной задержкой? Факт построения данной системы аминокислот может указывать на принципиальную важность временных факторов для процессов самоорганизации белка. В самом деле: если под влиянием боковых цепей аминокислот происходят, например, процессы формирования вторичной структуры в двух разных участках белка, то синхронизация процессов будет способствовать тому, что в нужный момент пространственно-временные траектории этих участков пересекутся, что приведет к образованию структуры следующего, более высокого уровня.

На наш взгляд, в данной работе мы почти полностью ответили на главный вопрос, поставленный во Введении: почему именно 20 аминокислот и почему именно в таком наборе. Наш ответ звучит

следующим образом. Двадцать аминокислот были выбраны потому, что сложившаяся система физических операторов и функциональных модулей в основе своей имеет структуру, изоморфную додекаэдру, который имеет 20 вершин. Большого числа операторов на этой структуре просто не уместилось. Выявленные внутри данного множества групповые отношения, обусловленные преобразованиями инверсной антисимметрии и антисимметрии антиподов, также ограничивают число боковых цепей. По сути, мы имеем дело с набором из пяти боковых цепей, четырежды повторенным и преобразованным с помощью упомянутых выше операций.

Полученная система аминокислот заставляет задуматься и над проблемой построения на принципах антисимметрии альтернативных систем боковых цепей. Нам представляется, что такая возможность существует. Многочисленные ограничения, накладываемые структурой основной цепи, существенно сокращает количество альтернативных боковых цепей. Однако при разработке систем молекулярной электроники, основанных на различных типах основной цепи, проблема построения иной системы боковых цепей все же возникнет. Принципы, лежащие в основе системы канонических аминокислот, могут способствовать ее решению. В работе [26] была предположена возможность построения модели молекулярной электроники, в которой в качестве основной цепи выступает полинуклеотидная цепь, а физическими операторами служат модифицированные азотистые основания. Прототипом для построения системы модифицированных оснований может служить модель структуры канонических аминокислот на додекаэдре. Данная система, после дальнейших исследований, может оказаться основой и для решения системотехнических задач молекулярного уровня.

Хотя на данном этапе конкретных прикладных исследований пока не проводилось, представляет интерес опробовать полученную структуру к

анализу ряда проблем. Например, можно попытаться конструировать аналоги каких-либо известных соединений (к примеру, тетрапептидов, обладающих биологическими свойствами), на основе принципов антисимметрии. Не исключено, что данная система может оказаться полезной и при анализе конформаций белков в процессе их сворачивания. Преимущество данной системы аминокислот состоит в том, что она относительно независима от других систем, таких как генетический код, поскольку в белках существуют свои закономерности, определяемые структурными отношениями аминокислот. Будущее покажет, насколько данная классификация аминокислот окажется продуктивной.

Выражаю искреннюю благодарность проф. В.В.Лучинину за поддержку этой работы и полезные дискуссии, доцентам В.Е. Стефанову и А.Ю. Первицкому – за высказанные ценные замечания.

Литература

1. Шульц Г., Ширмер Р. Принципы структурной организации белков. – М.: Мир. – 1982. – 354 с.
2. Кнорре Д.Г., Мызина С.Д. Биологическая химия. – М.: Высшая школа. – 2000. – 480 с.
3. Карасев В.А. Генетический код: новые горизонты. – СПб: Тэсса. –2003. – 145 с.
4. Karasev V.A., Demchenko E.L., Stefanov V.E. Topological Coding of Polymers and Protein Structure Prediction.// *Mathematical Chemistry*. – 1999. – V. 6. – P.295-345.
5. Karasev V.A., Stefanov V.E. Topological nature of the genetic code. // *J.Theor.Biol.* – 2001. – V.209. – P.303-317.
6. Меклер Л.Б., Идлис Р.Г. Общий стереохимический генетический код – путь к биологии и универсальной медицине XXI века. // *Природа*. – 1993. – No.5. – С.29-63.
7. Гладышев Г.П. Теория комплементарности белковых структур. Препринт. М.: Инст. хим. физики. – 1970. – 12 с.
8. Zull J.E., Smith S.K. Is genetic code redundancy related to retention of structural information in both DNA strands?// *Trends Biochem. Sci.* –1990. – V.15. – P.257-261.
9. Polya G.M. Detection of informational complementarity in an amino acid sequence.// *Nature*. – 1975. – V.255. – P. 652-653.
10. Biro J. Comparative analysis of specificity in protein-protein interactions. Part III. Models of gene expression based on the sequential complementarity coding of some pituitary proteins.// *Med. Hypotheses*. – 1981. – V.7. – P.995-1007.

11. Blalock, J.E.; Smith, E.M. Hydropathic anti-complementarity of amino acids based on the genetic code.// *Biochem. Biophys. Res. Commun.* – 1984. – V.121. – P.203-207.
12. Root-Bernstein, R.S. Amino acid pairing. // *J.Theor.Biol.* – 1982. – V.94. – P.885-894.
13. Чипенс Г.И. Скрытая симметрия генетического кода и законы взаимодействия аминокислот.// *Биоорг. химия.* – 1991. – Т.17. – №10. – С.1335-1346.
14. Карасев В.А., Стефанов В.Е., Курганов Б.И. Надмолекулярные биоструктуры: организация, функционирование, происхождение. В кн.: *Итоги науки и техники, сер. Биолог. химия, т.31.* – М.: ВИНТИ. – 1989. – 199 с.
15. Карасев В.А., Лучинин В.В., Стефанов В.Е. Как построить биочип? // *Биотехнология.* – 1993. – №2. – С.3-15.
16. Karasev V.A., Luchinin V.V., Stefanov V.E. A model of molecular electronics based on the concept of conjugated ionic-hydrogen bond systems.// *Adv.Mater.Opt.Electron.* – 1994. – V.4. – P.203-218.
17. Карасев В.А., Лучинин В.В. Проблемы создания искусственных бионических микро- и наносистем. // *Известия вузов. Электроника.* – 1998. – №5. – С.53-68.
18. Харгиттай И., Харгиттай М. Симметрия глазами химика. М.:Мир. – 1989. – 494 с.
19. Гроссман И., Магнус В. Группы и их графы. М.: Мир. – 1971. – 248 с.
20. Kolb V.A., Makeyev E.V., Spirin A.S. // Co-translational folding of an eukaryotic multidomain protein in a prokaryotic translation system. *J. Biol Chem.* – 2000. – V.275. – P.16597-601.
21. Карасев В.А. О научном наследии С.Г. Сорокина. Метод выделения меридиональных циклов на структурах, изоморфных булевым

гиперкубам, и его использование для описания временных процессов.
Деп. в ВИНТИ, №237-B2004 от 10.02.04.

22. Jimenez-Montaña, M.A., de la Mora-Basañez, C.R. & Pöschel, Th. The hypercube structure of the genetic code explains conservative and non-conservative aminoacid substitutions in vivo and in vitro. // *BioSystems*. – 1996. – V.39. – P.117-125.
23. Карасев В.А., Сорокин С.Г. О топологической структуре генетического кода // *Генетика*. – 1997. – Т.33. – С.744-751.
24. Румер Ю.Б. Систематизация кодонов в генетическом коде // *Докл. АН СССР*. – 1968. – Т.183. – No.1. – С.225-226.
25. Яблонский С.В. Введение в дискретную математику. М.: Наука. – 1986. – 384 с.
26. Karasev V.A., Luchinin V.V., Stefanov V.E. Topological coding: towards new materials for molecular electronics // *Adv. Funct. Mater.* – 2002. – V.12. – No. 6-7. – P.461-469.