



УДК 547.962/458:543.63

КОНФОРМАЦИОННЫЙ АНАЛИЗ ГЛИКОПЕПТИДНЫХ
ФРАГМЕНТОВ
ГРУППОСПЕЦИФИЧЕСКИХ ВЕЩЕСТВ КРОВИ

Липкин Г. М., Аванов А. Я., Кочетков Н. К.

*Институт органической химии им. Н. Д. Зелинского
Академии наук СССР, Москва*

Проведен теоретический конформационный анализ небольших гликопептидов, содержащих остатки треонина, гликозилированные дисахаридом $\text{Gal}\beta 1 \rightarrow 3\text{GalNAc}\alpha 1$. Показано, что взаимодействия углеводной цепи с пептидным остовом допускают ограниченный набор значений углов вращения в боковых цепях. Из рассмотрения модельных гликопептидов следует, что в группоспецифических веществах крови может реализоваться ближний порядок плотной упаковки олигосахаридных цепей. Предложена пространственная модель группоспецифических веществ крови, основным структурным элементом которой является β -структурная шпилька.

Группоспецифические вещества крови (ГВК), представляющие собой гликопротеины с молекулярной массой порядка 10^5 – 10^6 [1, 2], определяют иммунологическую специфичность как человека, так и высших животных. Состав пептидной и углеводной частей ГВК весьма своеобразен. Все они характеризуются высоким содержанием треонина, серина и пролина, которое достигает в среднем 60–70%. Однако в отдельных фрагментах ГВК оно может быть и существенно выше. Так, Гудвином и Уоткинс [3] при обработке проназой гликопротеина из кисты человека с групповой специфичностью были выделены фракции, в которых сумма трех указанных остатков достигает 90%. Например, в высокомолекулярной фракции I соотношение следующее: Thr – 47,3%, Ser – 24,8%, Pro – 17,3%. Эти же авторы установили, что в пептидной цепи ГВК имеются участки, состоящие в основном из остатков серина и треонина, которые практически полностью гликозилированы.

Известно, что углеводные цепи соединены с пептидным остовом молекулы O-гликозидными связями через остатки серина и треонина [4]. В среднем в ГВК степень O-гликозидного замещения близка к 60–70%, однако в образце, исследованном в работе [3], были замещены гидроксильные группы всех остатков серинов и треонина. Боковые углеводные цепи ГВК близки по составу, но различаются длиной (3–12 остатков [2, 5]). Связываясь с пептидным остовом ссуществляется через остаток α -D-N-ацетилгалактозамина, который всегда соседствует с остатком β -D-галактозы. Дисахаридное звено $\text{Gal}\beta 1 \rightarrow 3\text{GalNAc}\alpha 1$ * инвариантно во всех ГВК и представляет собой кор углеводных цепей. На невозстанавливаемом конце дисахаридного кора находятся иммунодоминантные остатки, отделенные от галактозного остатка кора цепью, включающей в себя ряд дисахаридов: $\text{Gal}\beta 1 \rightarrow 3\text{GalNAc}$ или $\text{Gal}\beta 1 \rightarrow 4\text{GalNAc}$.

Данные о макромолекулярной структуре ГВК весьма скудны. Ряд исследователей полагают, что ГВК имеет клубкообразную конфигурацию [1]. Однако анализ препарата из слизистой оболочки желудка свиньи методом акваметрического титрования [6] показал, что только небольшая часть потенциальных центров гидратации взаимодействует с молекулами воды. На основании этих данных сделан вывод, что ГВК имеют компактную структуру.

* Все аминокислоты – L-конфигурации, моносахариды – D-конфигурации. Дисахаридное звено в изображении молекулы гликопротеина в ряде случаев обозначено символом СНО.

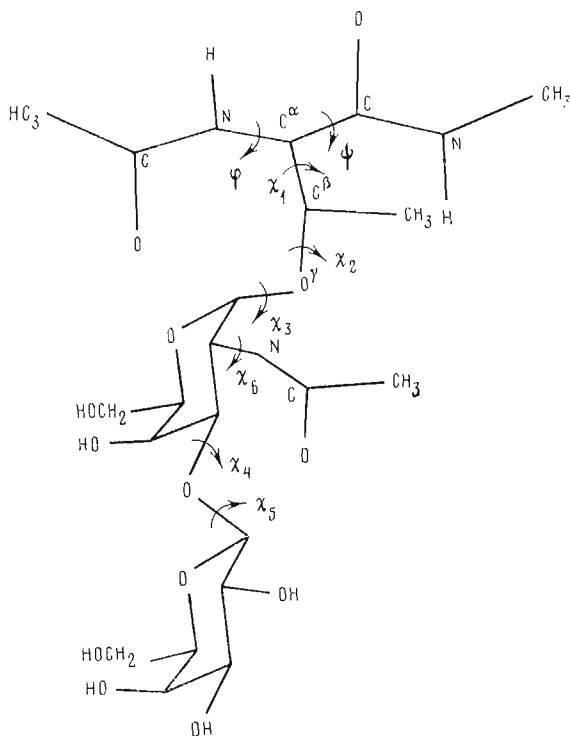


Рис. 1. Модель гликопептида $\text{CH}_3\text{CO-Thr}(\leftarrow 1\alpha\text{GalNAc3}\leftarrow 1\beta\text{Gal})\text{-NHCH}_3$ с указанием углов вращения

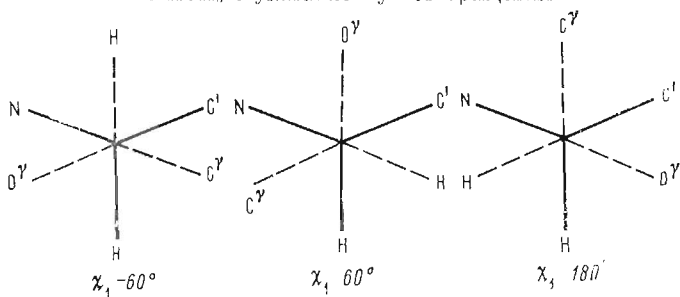


Рис. 2. Ньюменовские проекции положений боковой цепи остатка треонина

С целью поиска подходов к оценке конформационных особенностей сложной молекулы ГВК в данной работе предпринят теоретический конформационный анализ небольших гликопептидов. В работе рассматривались гликопептиды, содержащие остатки треонина, гликозилированные дисахаридными последовательностями $\text{Gal}\beta 1 \rightarrow 3\text{GalNAc}\alpha 1$, которые присутствуют во всех олигосахаридах (кор углеводной цепи). Анализ невалентных взаимодействий между кором и остатками пептидного остова позволяет выявить влияние таких взаимодействий на конформации и положения кора, которые, очевидно, в существенной степени определяют и пространственные ориентации всех боковых углеводных цепей ГВК. В пептидной части учитывались только остатки треонина, которые обладают меньшей конформационной свободой по углам вращения φ ($\text{C}^\alpha\text{-N}$), ψ ($\text{C}^\alpha\text{-C}'$) и χ ($\text{C}^\alpha\text{-C}^\beta$) по сравнению с остатками серина [7], что несколько уменьшает объем расчета.

Ближние взаимодействия в структуре ГВК были рассмотрены на традиционной в конформационном анализе пептидов модели — метиламиде N-ацетил- α -аминокислоты. В данном случае такой моделью является гликопептид $\text{CH}_3\text{CO-Thr}(\leftarrow 1\alpha\text{GalNAc3}\leftarrow \beta\text{Gal})\text{-NHCH}_3$. Так как в первичной последовательности ГВК имеются участки, состоящие только из гликози-

лированных остатков серина и треонина, представляло интерес рассмотреть взаимодействия между соседними такими остатками. Это сделано на примере расчета конформаций гликопептида $\text{CH}_3\text{CO-Thr}(\text{CHO})\text{-Thr}(\text{CHO})\text{-NHCH}_3$. Нами также проведен анализ контактов боковых цепей двух O-гликозелированных остатков, разделенных каким-либо другим остатком. Такие последовательности также широко представлены в ГВК [2, 3]. Расчет отмеченных взаимодействий проведен на модели с тремя аминокислотными остатками — $\text{CH}_3\text{CO-Thr}(\text{CHO})\text{-Ala-Thr}(\text{CHO})\text{-NHCH}_3$. Кроме указанных моделей в работе также обсуждаются результаты расчета избранных конформаций ряда других олигопептидов, а также O-гликозелированного политреонина.

Конформационный анализ гликопептидов проведен с использованием эмпирических энергетических функций. При оценке потенциальной энергии конформеров учитывались невалентные и электростатические взаимодействия, водородные связи и торсионная энергия. Функции невалентных взаимодействий взяты из работы Скотта и Шераги [8]. Заряды на атомах пептидных групп и углеводных остатков приведены в работах [9, 10]. Для эффективной диэлектрической проницаемости использовано значение, равное 10 [11]. Образование водородной связи описывалось потенциалом, предложенным в работе [9]. В функции торсионного вклада барьер вращения вокруг алифатической связи C—C равен 3 ккал/моль, для остальных связей в рассматриваемых молекулах он не превышает 1 ккал/моль [9, 10].

Конформационные параметры остатка треонина, гликозелированного дисахаридом, показаны на рис. 1. Это углы вращения φ ($\text{C}^\alpha\text{-N}$) и ψ ($\text{C}^\alpha\text{-C}'$) в основной и χ_1 ($\text{C}^\alpha\text{-C}^\beta$), χ_2 ($\text{C}^\beta\text{-C}'$) в боковой цепях треонина, а также углы вращения вокруг гликозидных связей χ_3 (C1-O), χ_4 (O-C3) и χ_5 (C1-O) дисахаридного звена, представленного в боковой цепи гликопептида. Кроме того, учтено вращение N-ацетиламидной группы N-ацетилгалактозамина (угол χ_6 (C2-N)). Остальные углы вращения в углеводном остатке приняты постоянными. При определении значений углов вращения использована международная номенклатура [12], по которой исходной является *цис*-ориентация соответствующих связей (см. также [13]). На Ньюеновских проекциях (рис. 2) даны положения боковой цепи треонина при значениях угла χ_1 -60° , 60° и 180° . Отметим, что в белках представлены именно такие расположения атомов O' и C' относительно основной цепи, а не мысленно возможные их зеркальные отображения. Для геометрических параметров пептидных цепей использованы средние значения, приведенные в работе [9], а для углеводных остатков — также средние значения, полученные Арноттом и Скоттом [14].

$\text{CH}_3\text{CO-Thr}(\leftarrow 1\alpha\text{GalNAc3}\leftarrow 1\beta\text{Gal})\text{-NHCH}_3$. В теоретическом конформационном анализе метиламида N-ацетилтреонина [7] отмечена взаимозависимость допустимых углов вращения в основной цепи (φ и ψ) от положения боковой цепи. Это обстоятельство учитывалось при поиске оптимальных конформаций гликопептида. На первом этапе расчета определялись углы вращения χ_4 и χ_5 дисахаридного звена $\text{Gal}\beta 1\rightarrow 3\text{GalNAc}\alpha 1$. Затем определялись параметры χ_2 и χ_3 путем построения зависимости потенциальной энергии всей молекулы от этих углов при условии, что углы φ , ψ и χ_1 остатков треонина и углы χ_4 , χ_5 дисахаридного звена равны оптимальным значениям. На конечном этапе проводилась минимизация потенциальной энергии всей молекулы.

На конформационных картах φ — ψ аминокислотных остатков, имеющих боковые радикалы, выделяются три области разрешенных значений углов, обозначенные буквами *R*, *B* и *L* [11]. В данной работе для основной цепи треонина приняты во внимание конформации, соответствующие областям *R* и *B* конформационных карт [7]. Конформации *L* ($\varphi\sim 60^\circ$, $\psi\sim 60^\circ$) в работе не учитывались, так как остатки треонина в этом состоянии в белках не встречаются [15].

На потенциальной поверхности дисахарида выделяются два минимума, для которых значения угла χ_5 (C1-O) почти равны ($\sim 60^\circ$), а угла χ_4 (O-C3) незначительно отличаются друг от друга (рис. 3). Оба мини-

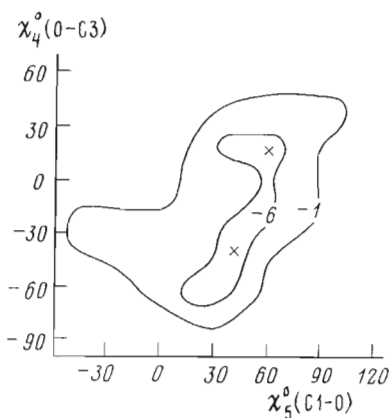


Рис. 3

Рис. 3. Конформационная карта $\chi_5(C1-O) - \chi_4(O-C3)$ для дисахарида Gal $\beta 1 \rightarrow 3$ GalNAc $\alpha 1$ (знаком \times помечены положения минимумов)

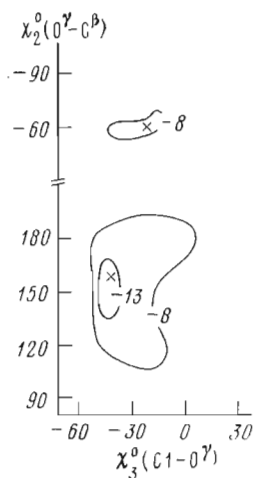


Рис. 4

Рис. 4. Сечение $\chi_2(C^{\beta}-O^{\beta}) - \chi_3(O^{\gamma}-C1)$ потенциальной поверхности гликопептида $CH_3CO-Thr(SHO)-NHCH_3$, основная цепь которого находится в состоянии R : $\varphi -64^\circ$, $\psi -44^\circ$, $\chi_1 -60^\circ$, $\chi_4 20^\circ$, $\chi_5 60^\circ$, $\chi_6 180^\circ$

мума изоэнергетичны. Внутримолекулярные водородные связи это звено не образует. Отметим, что в дисахариде, в котором N -ацетилгалактозамин замещен галактозой Gal $\beta 1 \rightarrow 3$ Gal $\alpha 1$, возможно образование водородной связи Gal(O2...HO3)Gal.

Из построения сечений потенциальной поверхности гликопептида $\chi_2 - \chi_3$ следует, что при конформациях пептидного остова типа B ($\psi \sim 120^\circ$) имеется три локальных минимума χ_2 , χ_3 160, -40° ; 100, -60° ; -60° , -20° , из которых третий является высокоэнергетическим. В том случае, когда пептидный остов принимает конформацию R ($\psi \sim -60^\circ$), наиболее реальна только одна пара значений χ_2 и χ_3 : 160 и -40° (рис. 4). Таким образом, углы вращения χ_2 и χ_3 в боковой цепи гликозилированного треонина являются «жесткими» конформационными параметрами. Угол χ_3 , равный $\sim -60^\circ$, соответствует *gauche*-ориентации агликона N -ацетилгалактозамина относительно связи $C1-O5$ кольца, которая также является оптимальной по критерию экзоаномерного эффекта [16].

Форма сечений $\chi_2 - \chi_3$ весьма чувствительна к положению ацетамидной группы N -ацетилгалактозамина. Из двух допустимых по невалентным взаимодействиям ее ориентаций, при которых связь $N-H$ располагается в *цис*- и *транс*-положениях относительно связи $C2-H$ кольца ($\chi_6 0^\circ$ и 180°), более предпочтительна (на ~ 4 ккал/моль) ориентация со значением $\chi_6 180^\circ$.

Таким образом, невалентные взаимодействия боковой цепи O -гликозилированного треонина с пептидным остовом допускают ограниченный набор значений углов вращения $\chi_1 \dots \chi_5$.

В основной цепи метиламида N -ацетилтреонина как в области B , так и области R значений углов φ , ψ имеется по два локальных минимума, в которых угол φ близок -120° и -70° [7]. Оптимальные конформации основной цепи в зависимости от значений угла φ далее обозначаются буквами b , B , r и R . Таким образом, конформации с углами вращения $\varphi -120^\circ$, $\psi 120^\circ$, $\chi_1 60^\circ$, $\chi_2 160^\circ$, $\chi_3 -40^\circ$, $\chi_4 20^\circ$, $\chi_5 60^\circ$ отвечает идентификатор b_{211} (цифровые идентификаторы приведены в табл. 1).

Энергии оптимальных конформаций области R рассматриваемого гликопептида даны в табл. 2. Учитывая, что все конформеры, в которых углы χ_2 , χ_3 соответствуют состоянию 3, являются высокоэнергетическими, а состояния 1 и 2 по углам χ_4 , χ_5 характеризуются близкими энергетическими и геометрическими параметрами, выделим те конформации области R ,

Цифровые идентификаторы, соответствующие значениям углов $\chi_1 \dots \chi_5$ (град) в локальных минимумах

Идентификатор	χ_1	χ_2, χ_3	χ_4, χ_5
1	-60	160, -40	20, 60
2	60	100, -60	-40, 40
3	180	-60, -20	

Таблица 2

Энергии оптимальных конформеров области R молекулы $\text{CH}_3\text{CO}-(\beta\text{-D-Gal-(1 \rightarrow 3)-}\alpha\text{-D-GalNAc)-Thr-NHCH}_3$ ($\text{CH}_3\text{CO-Thr(CHO)-NHCH}_3$)

Идентификаторы углов вращения				E , ккал/моль	Идентификаторы углов вращения				E , ккал/моль
Φ, Ψ	χ_1	χ_2, χ_3	χ_4, χ_5		Φ, Ψ	χ_1	χ_2, χ_3	χ_4, χ_5	
R	1	1	1	-19,9	r	2	1	1	-23,2
	1	1	2	-19,2		2	1	2	-21,6
	1	3	1	-16,6		2	3	1	-9,0
	1	3	2	-16,4		2	3	2	-8,0
r	1	1	1	-20,6	R	3	1	1	-18,5
	1	1	2	-18,8		3	1	2	-17,4
	1	3	1	-17,0		3	3	1	-*
	1	3	2	-16,7		3	3	2	-
R	2	1	1	-22,5	r	3	1	1	-17,1
	2	1	2	-22,3		3	1	2	-16,5
	2	3	1	40,0		3	3	1	-
	2	3	2	40,5		3	3	2	-

* Проверка соответствует очень высоким значениям энергии.

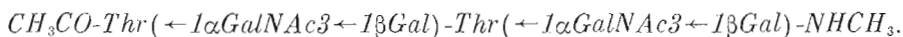
Таблица 3

Углы вращения (град) в низкоэнергетических конформациях гликопептида $\text{CH}_3\text{CO-Thr(CHO)-NHCH}_3$

Идентификатор	E , ккал/моль	Φ	Ψ	χ_1	χ_2	χ_3	χ_4	χ_5	χ_6
r_{11}	-20,6	-113	-55	-50	151	-41	28	49	-162
R_{11}	-19,9	-71	-48	-49	149	-42	28	49	-164
r_{21}	-23,2	-100	-49	61	177	-35	26	58	-180
R_{21}	-22,5	-72	-33	51	162	-41	27	57	-167
r_{31}	-17,0	-108	-37	-177	153	-45	23	51	-163
R_{31}	-18,5	-57	-35	-178	154	-42	23	51	-164
b_{11}	-22,3	-118	148	-48	170	-33	27	55	-167
b_{12}	-22,0	-130	146	-56	96	-46	26	47	-166
B_{11}	-19,8	-66	120	-51	170	-35	25	55	-168
B_{12}	-18,5	-78	121	-61	96	-44	26	48	-171
b_{21}	-20,8	-133	166	59	164	-41	26	55	-172
b_{22}	-20,8	-132	160	65	103	-40	22	51	-162
b_{31}	-21,8	-141	181	-177	170	-36	27	52	-175
b_{32}	-21,8	-143	174	-178	79	-61	26	52	-178

углы вращения в которых будут приниматься в качестве нулевых приближений при расчете более сложных молекул: r_{311} , R_{111} , r_{211} , R_{211} , r_{311} и R_{311} . Углы вращения в этих предпочтительных вариантах даны в табл. 3. Здесь же приведены параметры отобранных таким же образом конформеров области B (b_{111} , b_{121} , B_{111} , B_{121} , b_{211} , b_{221} , b_{311} , b_{321}), которые учитываются при расчете гликопептидов более сложной структуры*.

* Поскольку углы χ_4 и χ_5 во всех рассматриваемых конформациях областей R и B равны, третий цифровой индекс в таблицах и в дальнейшем тексте опущен.



При переходе к структуре с двумя соседними остатками треонина с гликозилированными боковыми цепями прежде всего отметим, что из четырех принимаемых во внимание форм дипептидов *BB*, *BR*, *RR* и *RB* в двух — *BB* и *BR* — конформации основной цепи развернуты, а в формах *RR* и *RB* свернуты, с одинаковой ориентацией концевых пептидных групп относительно центрального пептидного звена [17]. Этим обстоятельством объясняется близость энергии конформеров с одинаковыми положениями боковых цепей в парах *RR*, *RB* и *BB*, *BR*. Поэтому в табл. 4 и 5 даны энергии конформеров только форм *RR* и *BB*. Геометрические и энергетические параметры предпочтительных конформаций рассматриваемой молекулы приведены в табл. 6.

Самой низкоэнергетической является свернутая конформация $R_{11}R_{11}$ (рис. 5), в которой возможны существенные дисперсионные взаимодействия соседних углеводных цепочек (энергия ~ -11 ккал/моль). Это характерно и для конформации $R_{11}B_{11}$ (см. табл. 6). Оптимальная упаковка боковых цепей достигается в случае, когда они находятся в одном и том же конформационном состоянии ($\chi_1 -60^\circ$) и когда углы ϕ равны $\sim -55^\circ$, т. е. значению, характерному для правой α -спирали полипептидов [18]. Для второго остатка недопустимы ориентации боковой цепи остатков с $\chi_1 60$ и 180° (табл. 4).

Для развернутых конформаций типа *BB* не наблюдается столь существенной энергетической дифференциации, которая имеет место для формы *RR* (табл. 5), так как тенденция к взаимодействиям соседних боковых цепей в этом случае выражена значительно слабее. Только при равенстве угла $\chi_1 -60^\circ$ у первого остатка и 60° у второго дисперсионные взаимодействия углеводных цепей могут составлять $-6,0$ ккал/моль (варианты $b_{11}b_{21}$ и $b_{12}b_{21}$, табл. 6). Тем не менее низкоэнергетические конформации типа *BB* проигрывают $R_{11}R_{11}$ только 2,5 ккал/моль. Это обусловлено тем, что при развернутой форме пептидного остова возможны дополнительные взаимодействия боковых цепей с пептидными звеньями, которые уменьшают энергетическую разницу конформаций форм *BB*(*BR*) и *RR*(*RB*).

$CH_3CO-Thr(\leftarrow \alpha 1GalNAc\beta\leftarrow 1\beta Gal)-Ala-Thr(\leftarrow 1\alpha GalNAc\beta\leftarrow 1\beta Gal)-NHCH_3$. Для этой структуры, содержащей четыре пептидных звена, рассмотрены следующие формы основной цепи: *RRB* (*RRR*), *BBB* (*BBR*), *RBB* (*RBR*) и *BRB* (*BRR*). Для центрального остатка аланина учтены 4 конформационных состояния в областях *B* и *R*: *B*, *b*, *R*, *r*.

В конформациях свернутой формы *RRR*(*RRB*) углеводные цепи концевых остатков треонина не контактируют между собой, поэтому в них отсутствуют дополнительные по сравнению с внутримолекулярными взаимодействия. Напротив, между O -гликозилированными боковыми цепями соседних остатков треонина в конформациях *RR* может возникнуть плотная упаковка (см. предыдущий раздел).

В рассматриваемом гликопептиде по межостаточным взаимодействиям наибольший интерес представляют конформации развернутой формы *BBB* (рис. 6). Из расчета следует, что в структурах типа *BBB* оптимальное расположение углеводных цепей имеет место при идентичности конформаций первого и третьего остатков, включая боковые цепи. В такой ситуации олигосахариды ориентированы параллельно друг другу. Кроме того, остатки аминокислот в данном гликопептиде имеют тенденцию принимать значения углов ϕ , ψ , лежащие в области -130 , 150° (состояние *b*), которые близки углам вращения в регулярной β -структуре. В оптимальных вариантах $b_{11}bb_{11}$ (рис. 7), $b_{21}bb_{21}$, $b_{31}bb_{31}$ энергия дисперсионных взаимодействий дисахаридных звеньев первого и третьего остатков составляет ~ -7 ккал/моль (табл. 6).

Конформации $b_{11}Bb_{11}$, $b_{21}Bb_{21}$ и $b_{31}Bb_{31}$, в которых у остатка аланина углы ϕ близки к -70° , высокоэнергетичны из-за невалентных отталкиваний боковых цепей. Поэтому введение остатков пролина, для которых допустимые значения угла ϕ находятся вблизи области от -70 до -60° [19], в развернутые β -структурные фрагменты ГВК должно сопровождаться конформационными изменениями в пептидном остове.

Энергия (ккал/моль) оптимальных конформаций формы
R—R гликопептида CH₃CO-Thr(CHO)-Thr(CHO)-NHCH₃

Остаток 1	Остаток 2					
	R ₁₁	r ₁₁	R ₂₁	r ₂₁	R ₃₁	r ₃₁
R ₁₁	-52,5	—	-46,0	-49,0	-41,0	-40,0
r ₁₁	—	—	—	-49,0	-41,0	-40,0
R ₂₁	—	—	—	—	—	—
r ₂₁	—	—	—	—	—	—
R ₃₁	—	—	—	—	-41,7	—
r ₃₁	—	—	—	—	—	-39,1

Таблица 5

Энергии (ккал/моль) оптимальных конформаций формы
B—B гликопептида CH₃CO-Thr(CHO)-Thr(CHO)-NHCH₃

Оста- ток 1	Остаток 2							
	b ₁₁	b ₁₂	B ₁₁	B ₁₂	b ₂₁	b ₂₂	b ₃₁	b ₃₂
b ₁₁	-49,3	-47,6	-48,8	-43,4	-49,8	-48,2	-45,2	-45,0
b ₁₂	-47,1	-46,8	-48,1	-42,9	-49,9	-44,9	-45,0	-45,0
B ₁₁	-42,8	-44,2	-43,3	-44,3	-46,9	-43,8	-42,3	-41,0
B ₁₂	-41,1	-43,3	-42,8	-44,2	-48,5	-42,8	-41,5	-40,0
b ₂₁	-49,3	-49,2	-45,2	-43,6	-45,3	-44,7	-46,6	-45,1
b ₂₂	-49,1	-47,1	-47,5	-42,2	-47,9	-46,1	-46,6	-46,0
b ₃₁	-49,8	—	—	—	—	—	-41,2	—
b ₃₂	—	—	—	—	—	—	—	-40,2

Таблица 6

Параметры низкоэнергетических конформаций гликопептида
CH₃CO-Thr(CHO)-Thr(CHO)-NHCH₃

Конфор- мер	E, ккал/моль	E* _{s₁-s₂} , ккал/моль	Остаток 1			Остаток 2		
			φ	ψ	χ ₁ **	φ	ψ	χ ₁
R ₁₁ R ₁₁	-52,5	-10,7	-69	-39	-55	-61	-47	-50
R ₁₁ B ₁₁	-52,3	-12,0	-71	-44	-48	-66	120	-51
b ₁₁ b ₂₁	-49,8	-5,7	-142	157	-54	-124	161	45
b ₁₂ b ₂₁	-49,8	-5,9	-135	156	-64	-127	163	47
b ₁₁ b ₁₁	-49,2	-2,8	-145	147	-50	-122	148	-52
b ₃₁ b ₁₁	-49,9	-2,9	-138	-167	-169	-129	136	-57
b ₁₁ r ₂₁	-50,7	-4,3	-143	153	-50	-113	-51	59

* E_{s₁-s₂} — энергия взаимодействий боковых цепей остатков 1 и 2.

** Углы вращения χ₂—χ₆ не приведены, Они близки к значениям, указанным в табл. 2. Углы вращения φ, ψ и χ₁ даны в градусах.

Для конформеров формы BBB с разной направленностью углеводных боковых цепей отметим, что комбинации углов χ₁ 60 или 180° у первого остатка и χ₁ -60° у третьего невозможны, так как приводят к неблагоприятным контактам дисахаридов. Действительно, при χ₁¹ 180° и χ₁³ -60° боковые цепи направлены навстречу друг другу (см. рис. 6), в результате происходит недопустимое сближение концевых невозстанавливающих моносахаридных остатков. Поэтому, хотя конформации гликопептида b₁₁bb₂₁ и b₁₁Bb₂₁ являются низкоэнергетическими, на более длинных участках ГВК они вряд ли представлены, так как при чередовании углов вращения остатков триопина χ₁ -60° и χ₁ 60° возникают запрещенные конформации трипептидных звеньев, например b₂₁bb₁₁. Это относится и к комбинациям, в которых χ₁ равны -60 и 180°. Конформации, в которых первый и третий

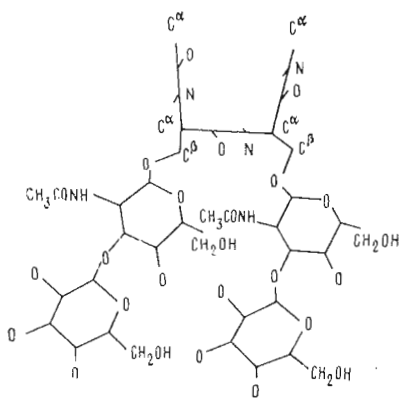


Рис. 5

Рис. 5. Модель конформации R_{11}, R_{11} гликопептида $\text{CH}_3\text{CO-Thr(CHO)-Ala-Thr(CHO)-NHCH}_3$

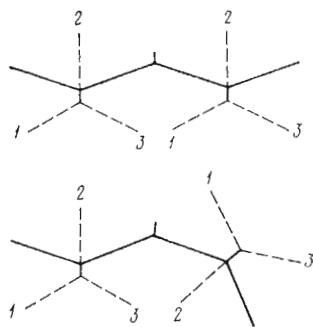


Рис. 6

Рис. 6. Схематическое изображение конформеров форм *BBV* и *BRB*. Ориентация боковых цепей показаны пунктиром. Цифры соответствуют идентификаторам угла вращения χ_1

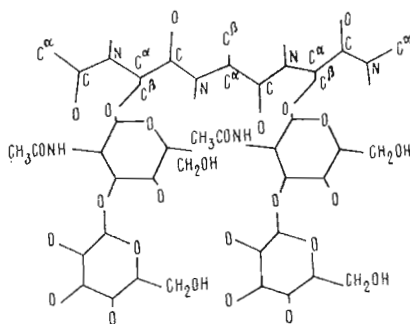


Рис. 7

Рис. 7. Модель конформации b_{11}, bb_{11} гликопептида $\text{CH}_3\text{CO-Thr(CHO)-Ala-Thr(CHO)-NHCH}_3$

остатки находятся в состояниях 2 и 3 или 3 и 2, нельзя признать удовлетворительными из-за отсутствия благоприятных взаимодействий между боковыми цепями.

Таким образом, если на фрагментах ГВК, богатых углеводными цепями, образуются развернутые тяжи, то наиболее оптимальна параллельная укладка тех углеводных цепей, которые находятся по одну сторону от плоскости пептидного остова.

Пространственные формы гликопептида *BRB* и *RBB* представляют собой сочетания развернутых и свернутых конформаций дипептидных звеньев. В них пептидный остов образует ковшеобразную структуру (рис. 6). Среди конформаций типа *BRB* самыми низкоэнергетическими являются те, у которых первый и третий остатки треонина имеют состояния 1 и 2 соответственно. При условии, что остаток аланина принимает конформацию *R* с $\varphi \sim -60^\circ$ (варианты b_{11}, Rb_{21} , b_{12}, Rb_{21} , табл. 7), энергия взаимодействий боковых углеводных цепей может колебаться от -4 до -6 ккал/моль. В остальных конформациях формы *BRB* эффективные ван-дер-ваальсовы контакты углеводов отсутствуют. Комбинации значений χ_1 концевых остатков 60 , -60 и 180 , 60° недопустимы. Для формы *RBB* также характерно отсутствие существенных взаимодействий боковых цепей первого и третьего остатков. Только в одном варианте R_{11}, BB_{11} , в котором углы χ_1 равны -60° , соответствующая энергия дисперсионных взаимодействий составляет $-4,7$ ккал/моль. Таким образом, в данном гликопептиде конформации типа *BRB* или *RBB* не обладают какими-либо энергетическими преимуществами перед структурами типа *BBB*.

Последовательное чередование форм *BRB* и *RBR* трипептидных звеньев образует форму $\dots RBRBR \dots$ полипептидной цепи, в которой для гликозилированных остатков наиболее приемлема ориентация со значением угла $\chi_1 -60^\circ$ (см. табл. 6).

Анализ конформационных состояний гликопептидов $\text{CH}_3\text{CO-Thr(CHO)-Thr(CHO)-NCH}_3$ и $\text{CH}_3\text{CO-Thr(CHO)-Ala-Thr(CHO)-NHCH}_3$, а также более сложных модельных соединений, о которых будет сказано ниже, позво-

Параметры низкоэнергетических конформаций гликопептида
 $\text{CH}_3\text{CO-Thr(CHO)-Ala-Thr(CHO)-NHCH}_3$

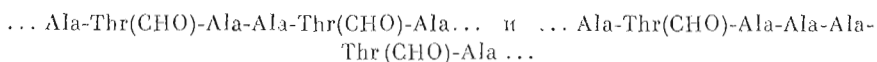
Конформер	E	$E^*_{s_1-s_3}$	Остаток 1			Остаток 2		Остаток 3		
	ккал/моль		φ	ψ	χ_1	φ	ψ	φ	ψ	χ_1
$b_{11}bb_{11}$	-52,3	-7,0	-118	150	-51	-136	162	-105	148	-51
$b_{12}bb_{12}$	-50,5	-4,2	-131	150	-56	-129	156	-123	146	-72
$b_{21}bb_{21}$	-52,3	-6,6	-133	154	54	-139	149	-134	160	59
$b_{31}bb_{31}$	-54,5	-6,7	-141	172	-171	-140	144	-140	-177	-172
$b_{11}bb_{21}$	-50,3	-4,9	-130	148	-47	-131	135	-143	162	54
$b_{11}Rb_{21}$	-55,1	-5,9	-118	149	-51	-59	-56	-130	160	59
$b_{11}RB_{11}$	-52,1	-1,8	-118	148	-50	-60	-58	-69	120	-45
$b_{21}Rb_{21}$	-53,1	-1,0	-133	154	53	-66	-60	-134	160	58
$R_{11}BB_{11}$	-50,3	-4,7	-71	-44	-49	-66	120	-66	120	-51
$R_{11}Bb_{21}$	-50,3	-1,3	-71	-49	-49	-80	112	-141	161	54
$r_{21}Bb_{21}$	-51,1	-1,3	-100	-37	55	-73	119	-134	160	59
$R_{11}RB_{11}$	-50,6	-0,5	-73	-50	-48	-62	-55	-67	149	-45

* $E_{s_1-s_3}$ — энергии взаимодействия боковых цепей гликозилированных остатков треонина.

ляет сделать некоторые предположения о пространственной структуре фрагментов ГВК с большим содержанием остатков треонина и серина.

Расчет показал, что наиболее эффективные стабилизирующие взаимодействия между углеводными цепями соседних гликозилированных остатков серина и треонина могут возникнуть в том случае, когда в основной цепи двух остатков представлено состояние RR (или RB). Однако этот локальный стабилизирующий эффект не может реализоваться на длинных участках ГВК, так как α -спиральная конформация ($\dots RRR \dots$) для цепи с большим содержанием гликозилированных остатков мало реальна.

Хорошо известно, что в α -спиральной структуре полипептидной цепи между собой взаимодействуют остатки соседних витков, и прежде всего остатки с номерами вдоль цепи n и $n+3$, а также n и $n+4$ (см., например, [18]). Поэтому нами проведен расчет оптимальной α -спиральной конформации полипептидов, в которых в указанных положениях находятся гликозилированные остатки треонина:



При минимизации потенциальной энергии варьировались углы вращения: как в основной, так и в боковых цепях. Из расчета следует, что во второй структуре при самой оптимальной упаковке боковых цепей на поверхности α -спирали между дисахаридными звеньями β -D-Gal(1 \rightarrow 3)- α -D-GalNAc возникают недопустимые ван-дер-ваальсовы контакты.

Итак, плотная упаковка боковых цепей расположенных рядом гликозилированных остатков может возникнуть только на небольших участках (2–3 остатка), в основной цепи которых представлены состояния $R-R$ или $R-B$. Вместе с тем на нерегулярных участках глобулярных белков звенья в конформациях $R-R$ или $R-B$ (а также $B-L$ при наличии остатков глицина) образуют в очень многих случаях β -изгибы [19, 20]. Эти β -изгибы, на которых происходит изменение направления хода пептидной цепи, в белках являются столь же важными компонентами структуры, как α -спиральные и β -структурные участки. В β -изгибах высока вероятность присутствия остатков, которыми обогащены ГВК: серина и пролина [20]. Естественно допустить, что для полипептидного остова ГВК характерна тенденция к образованию β -изгибов, которые дополнительно стабилизируются взаимодействиями углеводных цепей.

В том случае, если в ГВК представлены упорядоченные структурные элементы, ими скорее всего являются β -структурные участки. Расчет раз-

вернутой конформации изолированного гликозилированного политреонина с углами вращения основной цепи, близкими таковым в антипараллельной β -структуре, показал, что она возможна и все конформационные и энергетические закономерности, полученные при анализе модельных гликопептидов, остаются в силе. В оптимальном варианте имеет место адекватное параллельное расположение боковых углеводных цепей при остатках с номерами n и $n+2$ вдоль пептидного остова. Хотя энергия дисперсионных взаимодействий, приходящаяся на дисахарид Gal β 1 \rightarrow 3GalNAc α 1, в развернутой конформации несколько выше, чем в случае конформации RRR гликозилированного трипептида, образование β -структурных фрагментов весьма вероятно, так как отсутствуют ограничения на длину цепи. Отметим, что гомополипептиды серина и треонина, остатков, столь характерных для ГВК, образуют именно β -структуры [21].

Однако в случае ГВК традиционные β -структуры, образуемые, например, фибриллярными белками, невозможны. Это связано со своеобразием ориентации углеводных цепей гликозилированных остатков. Углеводные боковые цепи направлены не перпендикулярно к плоскости развернутого пептидного остова, а приблизительно параллельно ей. Поэтому такие цепи, которые уже в случае дисахарида Gal β 3 \rightarrow 3GalNAc α 1 по своим размерам превышают расстояние между пептидными тяжами β -структуры, должны испытывать значительные невалентные возмущения внутри β -структурного слоя.

Для фрагментов ГВК, насыщенных гликозилированными остатками, вполне приемлема конфигурация β -структурной «шпильки». Такая «шпилька» представляет собой два развернутых, соединенных водородными связями, тяжа, между которыми находится, как правило, β -изгиб. Потенциальную предрасположенность последовательности ГВК к образованию развернутых конформаций и β -изгибов, обусловленную взаимодействиями углеводных цепей, можно объединить в модели, в которой важным структурным элементом ГВК является β -шпилька. Этой модели не противоречит высокое содержание пролина в ГВК. Этот остаток наряду с глицином обладает максимальной потенциальной способностью встраиваться в β -изгибы глобулярных белков [20] и способствует их образованию. Вместе с тем его присутствие не исключено и в развернутых фрагментах β -шпильки. Пролин по пространственным причинам не может находиться во внутренних областях β -структуры. Однако на внешних, граничных пептидных тяжах он обнаружен. Например, в доменах иммуноглобулинов имеются два антипараллельных β -структурных образования, в которых остатки пролина находятся в граничных тяжах 4-1, 4-4, 3-1 и 3-3 [22]. В изолированной β -шпильке оба развернутых фрагмента являются «внешними» и, таким образом, допускают встраивание остатков пролина. Вследствие значительного содержания пролина, а также серина β -структурные тяжи в ГВК, вероятно, невелики. Приемлемыми для них являются размеры, обнаруженные в β -шпильках глобулярных белков. Так, в α -химотрипсине и карбоксипептидазе А они составляют в среднем 7-9 остатков.

Остатки треонина по сравнению с остатками серина являются более конформационно жесткими [15]. Кроме того, из статистического анализа последовательностей белков следует, что для треонина высока тенденция находиться в β -структуре [23]. Можно предположить, что в ГВК фрагменты с повышенным содержанием гликозилированных остатков треонина образуют развернутые тяжи β -шпилек. Не исключено, например, что гликопептид Pro-Thr(CHO)-Thr(CHO)-Thr(CHO)-Pro-Ser(CHO)-Thr(CHO), выделенный из ГВК Гудвином и Уоткинс [3], в нативной структуре белка принимает именно эту конформацию. Последовательности, содержащие преимущественно остатки серина и пролина, например Ser(CHO)-Ser(CHO) или Pro-Ser(CHO), с высокой степенью вероятности образуют β -изгибы.

Из анализа молекулярных моделей следует, что в полипептиде с большим числом гликозилированных остатков, образование β -шпильки допустимо только при определенном расположении углеводных цепей, при котором они направлены в стороны, противоположные водородным связям в шпиль-

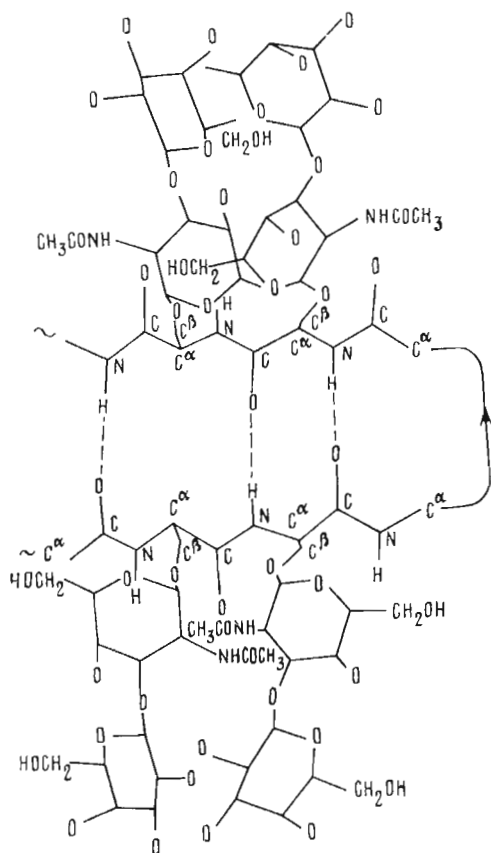


Рис. 8. Молекулярная модель β -структурной шпильки для пептидной цепи с гликозилированными остатками

ке (рис. 8). Конформации углеводных цепей при аминокислотных остатках, которые в шпильке находятся один против другого и образуют между собой водородные связи, должны быть одинаковы. В структуре, приведенной на рис. 8, у соседних вдоль цепи остатков углы вращения χ_1 ($C^\alpha - C^\beta$) равны 60 и -60° . Расчет гликозилированного политреонина в конформации $b_{11}b_{21}b_{11}b_{21} \dots$ показал, что она является одной из самых предпочтительных благодаря эффективным взаимодействиям углеводных цепей у остатков с номерами n и $n+2$. Таким образом, в β -структурной шпильке, включая β -изгиб, пептидный состав изолирован плотно упакованными олигосахаридами, невостанавливающие концы которых находятся в контакте с внешней средой.

Если β -структурные шпильки располагаются более или менее регулярно одна над другой, то образуется компактная структура гликопротеина с углеводами во внешней области. Приблизительная оценка размеров такой структуры (с учетом ван-дер-ваальсовых радиусов атомов и длины цепи ГВК) показывает, что ее ширина может составлять $60-70 \text{ \AA}$, а высота $\sim 300 \text{ \AA}$. Таким образом, в данной модели ГВК имеет форму удлиненного стержня. Это предположение согласуется с экспериментальной оценкой осевого отношения в молекуле ГВК [6].

Другим структурным компонентом ГВК с плотной упаковкой углеводных цепей могли бы быть фрагменты в зигзагообразной конформации $\dots RBRBR \dots$:



Как следует из расчета гликозилированного политреонина в конформации $R_{11}B_{11}R_{11}B_{11}R_{11} \dots$, частичный вклад остатка в общую энергию такой же, как в случае структуры $\dots bbb \dots$. Существенную роль в стабилизации

такой структуры играют дисперсионные взаимодействия между углеводными цепями соседних остатков на каждом звене *RV*. Для них допустима взаимная ориентация, показанная на рис. 5. Углы вращения ϕ и ψ в пептидной цепи таковы, что в любое ее положение могут встраиваться остатки пролина.

Таким образом, конформационный анализ модельных гликопептидов приводит к заключению, что ГВК могут образовывать пространственные структуры с ближним порядком плотной упаковки олигосахаридных цепей.

ЛИТЕРАТУРА

1. *Watkins W. M.* In: *Glycoproteins* / Ed. Gottschalk A. BBA Library, 1972, v. 5, part B, p. 830–891.
2. *Derevitskaja V. A.* *Pure and Appl. Chem.*, 1981, v. 53, № 4, p. 89–106.
3. *Goodwin S. D., Watkins W. M.* *Eur. J. Biochem.*, 1974, v. 47, № 2, p. 371–382.
4. *Дервицкая В. А., Кара-Мурза С. Г., Кочетков Н. К.* Докл. АН СССР, 1965, т. 163, № 3, с. 650–653.
5. *Derevitskaja V. A., Arbatsky N. P., Kochetkov N. K.* *Eur. J. Biochem.*, 1978, v. 86, № 2, p. 423–437.
6. *Шерман Ф. Б., Хургии Ю. И.* В кн.: Конформационные изменения биополимеров в растворах. Тбилиси: Мецниереба, 1980, с. 146.
7. *Kreissler M. A., Arkhipova S. F., Lipkind G. M., Popov E. M.* *J. chim. phys.*, 1974, v. 71, № 6, p. 907–912.
8. *Scott R. A., Scheraga H. A.* *J. Chem. Phys.*, 1966, v. 44, № 8, p. 3054–3068.
9. *Попов Е. М., Дашевский В. Г., Липкинд Г. М., Архипова С. Ф.* Молекулярн. биология, 1968, т. 2, № 4, с. 612–621.
10. *Rees D. A., Smith P. J.* *J. Chem. Soc., Perkin II*, 1975, № 8, p. 836–840.
11. *Липкинд Г. М., Архипова С. Ф., Попов Е. М.* Ж. структурн. химии, 1970, т. 11, № 1, с. 121–126.
12. IUPAC — IUB Commission on Biochemical Nomenclature. *J. Mol. Biol.*, 1970, v. 52, № 1, p. 1–17.
13. *Липкинд Г. М., Кочетков Н. К.* Биоорг. химия, 1980, т. 6, № 12, с. 1817–1829.
14. *Arnott S., Scott W. E.* *J. Chem. Soc., Perkin Trans. II*, 1972, № 3, p. 324–335.
15. *Липкинд Г. М., Попов Е. М.* Молекулярн. биология, 1971, т. 5, № 5, с. 667–679.
16. *Стоддарт Дж.* Стереохимия углеводов. М.: Мир, 1975.
17. *Липкинд Г. М., Архипова С. Ф., Попов Е. М.* Молекулярн. биология, 1970, т. 4, № 3, с. 331–338.
18. *Lipkind G. M., Popov E. M.* *Int. J. Peptide Protein Res.*, 1973, v. 5, № 2, p. 371–379.
19. *Попов Е. М., Липкинд Г. М.* Молекулярн. биология, 1971, т. 5, № 4, с. 624–636.
20. *Chou P. Y., Fasman G. D.* *J. Mol. Biol.*, 1977, v. 115, № 2, p. 135–175.
21. *Vohak Z., Katchalski E.* *Biochemistry*, 1963, v. 2, № 2, p. 228–237.
22. *Poljak R. J., Amzel L. M., Phizackerly R. P.* *Progr. Biophys. Molec. Biol.*, 1976, v. 31, p. 67–93.
23. *Chou P. Y., Fasman G. D.* *Biochemistry*, 1974, v. 13, № 2, p. 211–222.

Поступила в редакцию
24.XI.1981

CONFORMATIONAL ANALYSIS OF GLYCOPEPTIDE FRAGMENTS OF BLOOD GROUP SUBSTANCES

LIPKIND G. M., AVANOV A. Ya., KOCHETKOV N. K.

*N. D. Zelinsky Institute of Organic Chemistry, Academy
of Sciences of the USSR, Moscow*

Theoretical conformational analysis has been carried out of small glycopeptides containing threonine residues glycosylated by a Gal β 1 \rightarrow 3GalNAc α 1 disaccharide. It is shown that interactions of carbohydrate side chains with the peptide backbone permit a limited set of rotation angles in the side chains. As follows from analysis of model glycopeptides, near-neighbour order of packing of carbohydrate chains can be realized in blood group substances. The spatial model of blood group substances is proposed, a general structural element of which is a β -hairpin.