

УДК: 51-76, 519.252

Применение метода Тагучи для изучения структуры пептидного лиганда в составе сложного комплекса

**Данилкович А.В.^{1,2}, Туробов В.И.¹, Соболев Е.В.^{3,4},
Удовиченко И.П.^{*1,2}**

¹*Филиал института биоорганической химии им. академиков М.М.Шемякина и Ю.А.Овчинникова Российской академии наук, Пуццино, Россия*

²*Пуццинский государственный естественно-научный институт, Пуццино, Россия*

³*Институт математических проблем биологии РАН – филиал Федерального государственного учреждения "Федеральный исследовательский центр Институт прикладной математики им. М.В. Келдыша Российской академии наук", Пуццино, Россия*

⁴*EMBL Hamburg, Гамбург, Германия*

Аннотация. Для моделирования структуры пептидного лиганда в составе комплекса белков главного комплекса гистосовместимости и Т-клеточного рецептора (PDB ID 2Z31) был использован метод Тагучи, позволивший сократить число экспериментов, требуемых для анализа вклада различных аминокислотных остатков по каждой позиции активного участка молекулы лиганда. В соответствии с методом Тагучи исследован ряд прототипов молекул лиганда в виде пространственных моделей, которые были использованы для молекулярного докинга с оптимизацией структуры комплексов по критерию минимизации внутренней молекулярно-механической энергии. Получена аминокислотная последовательность пептида, которая демонстрирует уменьшение энергии комплекса по сравнению с природным лигандом, что демонстрирует применимость метода Тагучи для оптимизации структуры пептидов с целью разработки новых молекул лигандов.

Ключевые слова: метод Тагучи, оптимизация структуры, пространственная модель, первичная структура.

ВВЕДЕНИЕ

Оптимизация структуры биологических молекул с целью направленного изменения их свойств играет важную роль в фундаментальных исследованиях и лежит в основе разработки новых лекарственных средств. Существует простой подход к решению конкретных задач оптимизации процессов и структур способом перебора всех возможных вариантов, т.н. «метод проб и ошибок». Несмотря на ясность, такой метод не может практически использоваться при анализе сложных систем со многими параметрами. Так для оптимизации структуры шестичленного пептида с учетом возможных комбинаций 20 аминокислотных остатков, необходимо синтезировать и проверить в биологических экспериментах 20⁶ молекул. В фарминдустрии широко используется метод SBDD (Structure-Based Drug Design), в частности 3D-QSAR, основанный на изучении комплексов лигандов с рецепторами [1]. Метод доказал свою

*iudovichenko1@yandex.ru

эффективность при работе с «малыми молекулами», однако для оптимизации структуры пептидов неэффективен.

Для решения проблемы оптимизации структуры пептидных лигандов, нами предложен подход с использованием метода Тагучи, основанный на использовании понятия ортогональных массивов (ОА) [2, 3]. Данный подход широко используется с середины прошлого века в промышленности, где позволяет, в частности, оптимизировать план экспериментов, призванных определить характеристики ключевых параметров, при которых обеспечиваются требуемые свойства объекта оптимизации. Особенностью метода Тагучи является возможность уменьшить количество экспериментов, необходимых для оптимизации множества факторов, в совокупности влияющих на результат [4]. Метод эффективно применяется в различных технических областях, в том числе, при производстве интегральных микросхем, в химической промышленности, машиностроении и электротехнике [5].

В биологии, несмотря на сложный многофакторный характер процессов, метод Тагучи начал ограниченно применяться только с середины 1990 гг. [6–8], в частности для оптимизации условий при рефолдинге белков [9]. Для оптимизации структуры пептидной природы его использование не описано.

Целью данной работы было показать применимость метода Тагучи на примере оптимизации структуры пептидного лиганда, входящего в состав сложного белкового комплекса, в котором он взаимодействует с главным комплексом гистосовместимости и Т-клеточным рецептором.

МЕТОДЫ

Описание метода Тагучи

Если исход эксперимента зависит от k параметров (факторов), каждый из которых может принимать множество значений размерности s , то для проверки всех возможных комбинаций необходимо выполнить s^k экспериментов. Идея метода Тагучи состоит в том, чтобы заменить проверку всех возможных комбинаций значений полного набора параметров на проверку всех возможных комбинаций из подмножеств значений меньшей размерности t . Если $t = s$, то это означает полную проверку всех возможных комбинаций. При $t = 2$ проверяются только все возможные комбинации каждой пары параметров. Такой подход значительно уменьшает количество экспериментов. С другой стороны, есть риск пропустить синергетический эффект трех факторов, если он есть.

Минимальное количество экспериментов в методе Тагучи определяются с помощью ортогонального массива, столбцы которого означают исследуемые параметры, строки – отдельные эксперименты, а элементы – уровни значений параметров в каждом эксперименте.

Ортогональный массив, ОА (N, k, s, t, v) – матрица размера $N \times k$, элементы которой суть числа $1, 2, \dots, s$, обладающая тем свойством, что в каждой ее подматрице размера $N \times t$ любой из s^t возможных t -мерных векторов-строк, имеющих координатами эти числа, встречается в качестве строк этой подматрицы точно v раз. Из определения следует, что $N = vs^t$. Другими словами, если выбрать t столбцов, то в них найдутся все комбинации чисел от 1 до s ровно v раз.

Ортогональные массивы обладают специальными свойствами, что уменьшает количество необходимых экспериментов:

1. Каждый столбец содержит значения соответствующего параметра для всех планируемых экспериментов. При этом все значения параметра встречаются одинаковое количество раз. Это называется балансирующим свойством ортогональных массивов.

2. Все значения каждого параметра будут использованы в планируемых экспериментах.

3. Каждый столбец взаимно ортогонален любому другому, т.е. скалярное произведение векторов соответствующих значениям весов равно нулю. Веса назначаются путем равномерного распределения значений между -1 и 1 .

Последнее свойство позволяет оценивать эффект от отдельных параметров. Для этого нужно усреднить целевую характеристику, полученную во всех экспериментах, для каждого уровня значения заданного параметра. Таблица размером $s \times k$, столбцы которой соответствуют параметрам, строки – уровням параметров, а элементы – соответствующим суммам целевых характеристик, называется таблицей ответа.

Перед проведением опытов, значения уровней параметров заменяются в ортогональном массиве реальными значениями этих параметров. Каждый опыт проводится с соответствующим набором значений параметров, и целевая характеристика фиксируется в каждом отдельном эксперименте. Наилучшее значение целевой характеристики определяют первый оптимальный набор значений параметров.

Выбор значений факторов

Для использования метода Тагучи в качестве параметров (факторов), были выбраны позиции аминокислотных остатков в первичной последовательности, а в качестве значений параметров (уровней) – типы аминокислотных остатков. Использование метода Тагучи позволяет оценить индивидуальный вклад аминокислотного остатка по каждой позиции в составе структуры лиганда путем сравнения значений соответствующих факторов.

В качестве значений факторов для использования в методе Тагучи были выбраны пять аминокислот из пяти групп подобия, сформированных на основе окружения аминокислотных остатков в первичных структурах 168808 белков [10]. Характерные представители каждой группы составили стартовый набор значений факторов, использованных при моделировании пептидного лиганда. Выборка аминокислотных остатков {R, S, N, V, P} из разных групп характеризуется индексом подобия от -0.055 до 0.271 (табл. 1). Размер данной выборки позволяет исследовать применимость метода Тагучи для дизайна пептидного лиганда.

Таблица 1. Величина подобия аминокислотных остатков*

	R	S	N	V	P
R	1.000	0.127	0.183	-0.037	-0.052
S	0.127	1.000	0.271	0.129	0.194
N	0.183	0.271	1.000	-0.055	0.063
V	-0.037	0.129	-0.055	1.000	-0.034
P	-0.052	0.194	0.063	-0.034	1.000

* По результатам анализа первичных структур белков [10].

Для анализа шести факторов (позиция аминокислотных остатков), принимающих пять значений (тип аминокислотного остатка), была сгенерирована матрица Тагучи АО (25, 6, 5, 2, 1) в программе Design-Expert v.7.1.5, которая использовалась для исследования 25 структур пептидов (экспериментов) с целью выбора для каждой позиции в цепи аминокислотного остатка с минимальной внутренней молекулярно-механической (ММ) энергией. В случае простого перебора всех возможных комбинаций аминокислотных остатков потребовалось бы проведение 15625 экспериментов (5^6).

Молекулярный докинг и определение внутренней молекулярно-механической энергии

При создании пространственной модели экспериментальной структуры лиганда в комплексе с рецепторами использовали рентгеноструктурные данные (PDB ID 2Z31)

комплекса фрагмента основного белка миелина с главным комплексом гистосовместимости (H-2-II α , H-2-II β) и T-клеточным рецептором (TCR α , TCR β). Для построения трехмерных структур лиганда в составе комплекса использовали программу Swiss-PdbViewer v.4.1.0, в которой реализована автоматическая функция выбора ротамера, обладающего минимальной ММ энергией по каждой позиции аминокислотного остатка [11]. Для молекул лигандов, находящихся в растворе или на поверхности рецептора и контактирующих с растворителем, принято учитывать энергию сольватации ΔG [12, 13], но когда пептидный лиганд экранирован от растворителя сложным комплексом, как в нашем случае – двумя рецепторами H-2-II $\alpha\beta$ /TCR $\alpha\beta$ (рис. 2), использование ММ энергии в качестве характеристики структуры приемлемо.

Структуры комплексов пептидов с рецепторами минимизировали, осуществляя три последовательные серии оптимизации по 500 итераций каждая с промежуточными «встрясками» между ними. Учитывалось электростатическое взаимодействие, и минимизация продолжалась до достижения ΔE между двумя последовательными итерациями менее 1.0 кДж/моль (рис. 1). Минимальные ММ энергии всех модельных структур были достигнуты в искусственной среде GROMOS96 43B [14], реализованной в программе Swiss-PdbViewer v.4.1.0 [11].

Результаты минимизации ММ энергии комплексов пептидов с рецепторами использовали при заполнении матрицы Тагучи (табл. 2) для последующего анализа.

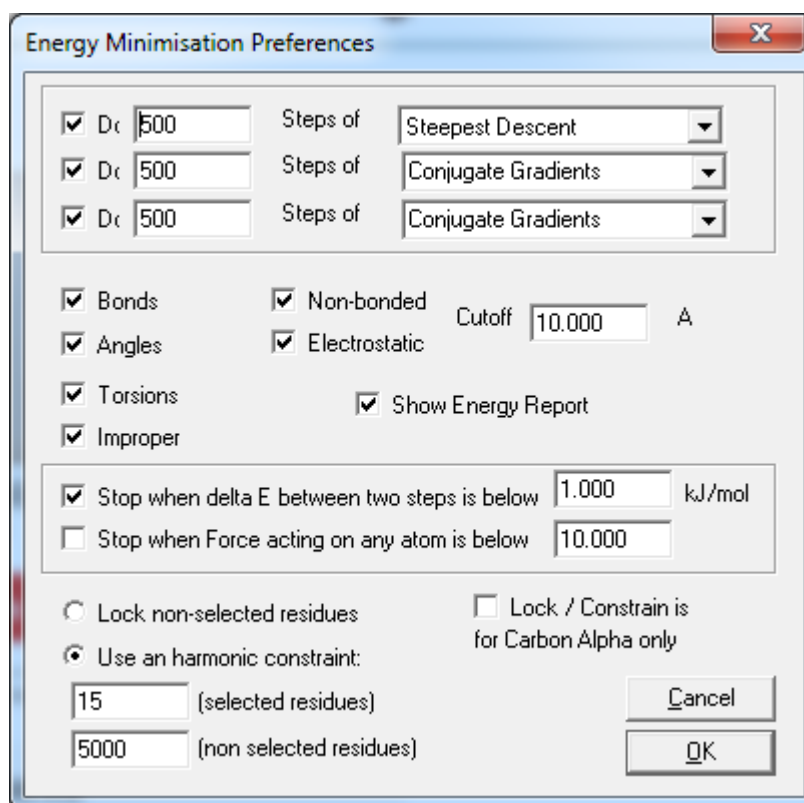


Рис. 1. Параметры программы Swiss-PdbViewer v.4.1.0, использованные для минимизации молекулярно-механической энергии комплексов пептидов с рецепторами.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Первичная структура пептидного лиганда в составе комплекса 2Z31 представлена аминокислотными остатками, находящимися в специфическом туннеле, сформированном субъединицами главного комплекса гистосовместимости и T-клеточного рецептора [15]. Поскольку основная информационная составляющая молекулы нативного лиганда представлена С-концевой областью пептида, для анализа

использовались модели лиганда длиной шесть аминокислотных остатков. В качестве оценки эффективности взаимодействия компонентов комплекса можно использовать величину ММ энергии в состоянии глобального минимума, что является отдельной задачей, и на первом этапе исследований был использован режим минимизации внутренней ММ энергии, реализованный в программе Swiss-PdbViewer v.4.1.0 (рис. 1). В соответствии с матрицей Тагучи АО (25, 6, 5, 2, 1) в результате молекулярного докинга были получены соответствующие пространственные модели комплекса с различными вариантами молекул лиганда и определены ММ энергии каждого комплекса (табл. 2). Оптимизация параметров программы Swiss-PdbViewer v.4.1.0 для минимизации ММ энергии позволила добиться высокой воспроизводимости результатов при использовании различных начальных состояний комплексов. Минимальная ММ энергия для комплекса с шестичленным фрагментом основного белка миелина H-SQYRPS-ОН составила -38456 кДж/моль.

Таблица 2. Матрица Тагучи АО (25, 6, 5, 2, 1), использованная для выбора структуры лигандов и экспериментов по определению ММ энергии комплексов, где в качестве значений факторов (уровней) были использованы аминокислотные остатки R, S, N, V, P, соответственно. В качестве факторов служили последовательные позиции в структуре пептидов, начиная с N-конца

Номер эксперимента	Фактор 1	Фактор 2	Фактор 3	Фактор 4	Фактор 5	Фактор 6	ММ энергия (кДж/моль)
1	S	V	P	R	S	N	-39734
2	N	S	V	R	N	P	-40216
3	P	P	V	N	S	R	-40016
4	N	N	P	S	V	R	-40274
5	V	P	N	R	V	S	-40264
6	N	R	N	P	S	V	-40538
7	S	N	V	P	R	S	-40204
8	V	N	R	V	S	P	-40490
9	R	R	R	R	R	R	-41470
10	P	R	P	V	N	S	-40644
11	R	S	S	S	S	S	-40283
12	R	N	N	N	N	N	-41121
13	N	P	S	V	R	N	-40666
14	R	P	P	P	P	P	-39973
15	S	P	R	S	N	V	-40595
16	V	S	P	N	R	V	-40477
17	R	V	V	V	V	V	-40184
18	P	S	R	P	V	N	-40683
19	V	V	S	P	N	R	-40513
20	N	V	R	N	P	S	-40869
21	P	N	S	R	P	V	-40821
22	S	S	N	V	P	R	-40662
23	P	V	N	S	R	P	-40926
24	V	R	V	S	P	N	-40591
25	S	R	S	N	V	P	-40669

* Нумерация столбцов, как позиций аминокислотных остатков, начинается с N-конца пептида.

Полученные значения ММ энергии были использованы для оценки вклада каждой из пяти аминокислот в соответствующей позиции (табл. 3). В результате была определена последовательность аминокислотных остатков пептида H-PRRNR-ОН, теоретически имеющая потенциал для низкой минимальной ММ энергии в составе комплекса. Для подтверждения этой гипотезы полученный пептид был использован для минимизации ММ энергии в тех же условиях, что и остальные пептиды. ММ энергия комплекса с ним составила -41609 кДж/моль, показав наименьшее значение из рассчитанных нами в этой работе для различных пептидов (таблицы 2 и 6).

Таблица 3. Анализ результатов, полученных согласно матрице Тагучи. Представлены средние значения ММ энергии (кДж/моль) для каждого значения (уровня) каждого фактора. Для каждого фактора выделены уровни с минимальным значением энергии, которые составляют оптимальный набор

Значение	Фактор 1	Фактор 2	Фактор 3	Фактор 4	Фактор 5	Фактор 6
R	-40606	-40782	-40821	-40501	-40749	-40587
S	-40373	-40464	-40590	-40534	-40212	-40453
N	-40513	-40582	-40702	-40630	-40618	-40559
V	-40467	-40445	-40242	-40529	-40415	-40523
P	-40618	-40303	-40220	-40382	-40583	-40455

Таблица 4. Медианы ММ энергии (кДж/моль) для каждого значения (уровня) каждого фактора

Значение	Фактор 1	Фактор 2	Фактор 3	Фактор 4	Фактор 5	Фактор 6
R	-40283	-40664	-40683	-40264	-40666	-40513
S	-40595	-40477	-40666	-40591	-40283	-40283
N	-40538	-40490	-40662	-40669	-40595	-40666
V	-40490	-40513	-40204	-40644	-40274	-40538
P	-40683	-40264	-40274	-40513	-40662	-40490

Таблица 5. Минимальные значения ММ энергии (кДж/моль) для каждого значения (уровня) каждого фактора

Значение	Фактор 1	Фактор 2	Фактор 3	Фактор 4	Фактор 5	Фактор 6
R	-40283	-40644	-40683	-40264	-40666	-40513
S	-40595	-40477	-40666	-40591	-40283	-40283
N	-40538	-40490	-40662	-40669	-40595	-40666
V	-40490	-40513	-40204	-40644	-40274	-40538
P	-40683	-40264	-40274	-40513	-40662	-40490

Полученный результат был подтвержден сравнением с тремя тестовыми выборками структур пептидов, выведенными на основании средних значений факторов (таблицы 3 и 6), медиан значений факторов (таблицы 4 и 6) и минимальных значений факторов (таблицы 5 и 6).

Данные сравнительного анализа подтверждают основной вывод, сделанный по результатам использования метода Тагучи, с определением пептида H-PRRNRR-OH, имеющего потенциал для низкой минимальной ММ энергии комплекса.

Таблица 6. Результаты минимизации ММ энергии (кДж/моль) комплексов пептидов, структуры которых сформированы на основании средних, медиан и минимальных значений факторов в порядке их убывания

Выборка по средним значениям ММ		Выборка по медианным значениям ММ энергии		Выборка по минимальным значениям ММ энергии	
Пептид	ММ энергия	Пептид	ММ энергия	Пептид	ММ энергия
H-PRRNRR-OH	-41609	H-PRRNRR-OH	-41443	H-RRRRRR-OH	-41470
H-RNNSNN-OH	-40967	H-SVSVPV-OH	-40021	H-PNNNNN-OH	-41200
H-NSSVPV-OH	-40273	H-NNNSNR-OH	-41313	H-NVSSPP-OH	-40341
H-VVVRVP-OH	-40547	H-VSPSP-OH	-40121	H-SSPPVS-OH	-40002
H-SPPSS-OH	-39946	H-RPVRVS-OH	-40619	H-VPVVS-OH	-40206

Комплекс пептида H-PRRNRR-OH с α - и β -цепями главного комплекса гистосовместимости II класса (H-2) и α - и β -цепями T-клеточного рецептора (TCR) после минимизации ММ энергии представлен на рисунке 2.

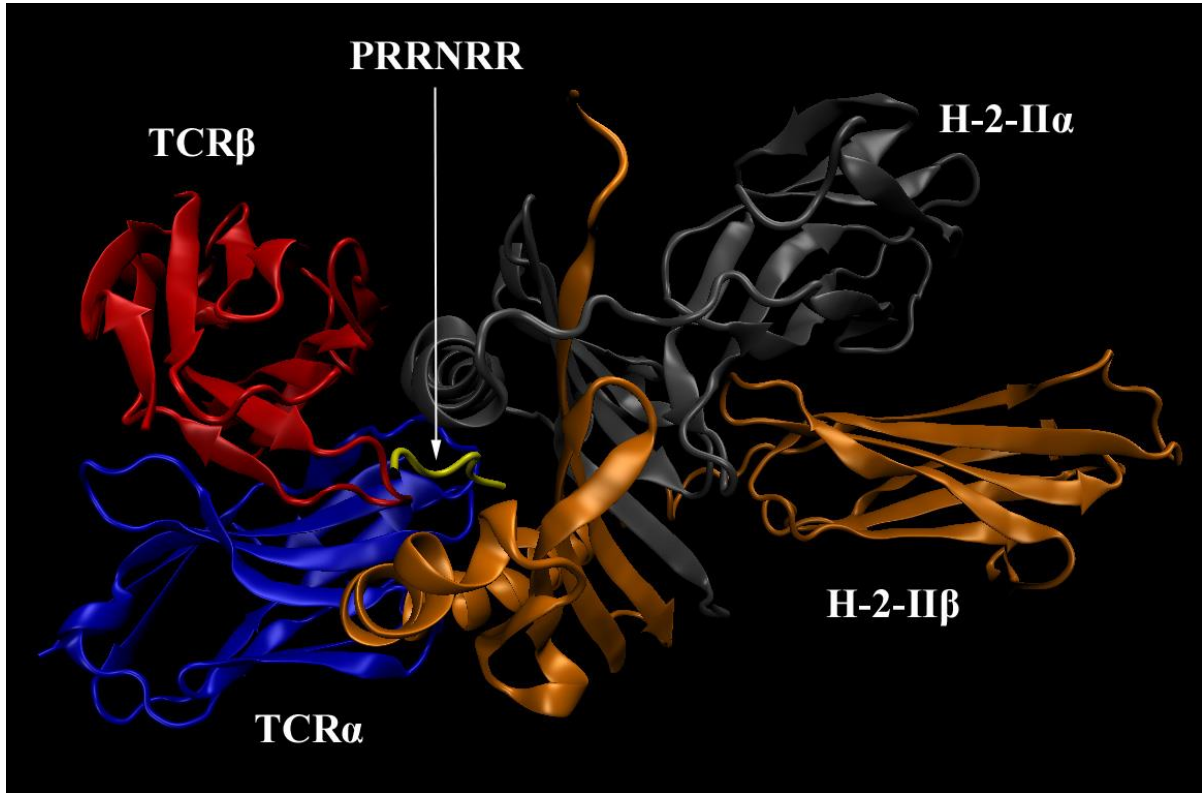


Рис. 2. Структура полученного с использованием метода Тагучи пептида H-PRRNRR-OH в составе комплекса. H-2-II α , H-2-II β – α - и β -цепи главного комплекса гистосовместимости II класса; TCR α , TCR β – α - и β -цепи T-клеточного рецептора.

Таким образом, предложенный подход показал свою применимость для решения сложной задачи оптимизации структуры пептида путем анализа его комплекса с соответствующими рецепторами. Применение метода Тагучи позволяет уменьшить число экспериментов, требуемых для определения параметров структуры лиганда в составе белкового комплекса (25 против 15625, в нашем случае).

Иммунокомпетентность пептида можно количественно охарактеризована *in vitro* в результате теста на пролиферативную активность (стимуляцию) культуры T-клеток, из животных, предварительно сенсibilизированных гомологичным антигеном. В такой системе происходит активация CD4⁺ T-клеток, с использованием двух сигналов, один из которых опосредован взаимодействием пептида в составе комплекса белков главного комплекса гистосовместимости и T-клеточного рецептора [16], что полностью соответствует архитектуре комплекса 2Z31, использованного в настоящей работе для моделирования структуры лиганда.

В практическом плане полученный результат может быть использован в работах по оптимизации и поиску биологически активных молекул, в том числе для синтеза потенциальных лигандов с целью последующего подтверждения и анализа их биологической активности в системах *in vitro* и *in vivo* [17, 18].

Преимущество предложенного подхода для оптимизации структуры пептидов проявляется в возможности использования его в биологических экспериментах, когда критическим является необходимость сокращения, как самого числа этих экспериментов, так и числа синтезируемых соединений для анализа.

ЗАКЛЮЧЕНИЕ

Показана применимость метода Тагучи для оптимизации структуры пептидов с целью разработки новых молекул лигандов. В результате оптимизации структуры пептидного лиганда получена новая аминокислотная последовательность H-PRRNRR-OH. Данная последовательность обеспечивает меньшую MM энергию по сравнению с

природным фрагментом основного белка миелина H-SQYRPS-OH в комплексе с T-клеточным рецептором и главным комплексом гистосовместимости.

Работа поддержана грантом Министерства образования и науки Российской Федерации, Соглашение №14.607.21.0133 (RFMEFI60715X0133).

СПИСОК ЛИТЕРАТУРЫ

1. Verma J., Khedkar V.M., Coutinho E.C. 3D-QSAR in drug design – a review. *Curr. Top. Med. Chem.* 2010. V. 10. P. 95–115.
2. Taguchi G. *System of experimental design: engineering methods to optimize quality and minimize costs*. V. 1. UNIPUB/Kraus International Publications, 1987.
3. Abbasian M.A., Moallem M., Fahimi B. Double stator switched reluctance machines (DSSRM): fundamentals and magnetic force analysis. *IEEE Trans. on Energy Con.* 2010. V. 25. P. 589–597.
4. Chapman P.L., Sudhoff S.D. Design and precise realization of optimized current waveform for switched reluctance drive. *IEEE Trans. Power Electronics.* 2002. V. 17. P. 76–83.
5. Feiz J., Finch J.W., Metwally H.M.B. A novel switched reluctance motor with multiple teeth per stator pole and comparison of such motors. *Electric Power System Research.* 1995. V. 34. P. 197–203.
6. Cobb B.D., Clarkson J.M. A simple procedure for optimising the polymerase chain reaction (PCR) using modified Taguchi methods. *Nucleic Acids Res.* 1994. V. 22. P. 3801–3805.
7. Burch G.J., Ferguson C.H., Cartwright G., Kwong F.Y. Application of Taguchi experimental design to the optimisation of a baculovirus expression system. *Biochem. Soc. Trans.* 1995. V. 23. P. 107S.
8. Jeney C., Dobay O., Lengyel A., Adam E., Nasz I. Taguchi optimisation of ELISA procedures. *J. Immunol. Methods.* 1999. V. 223. P. 137–146.
9. Vijayasathy S., Datta S. Method: U.S. Patent. № 20080200649.
10. Rogov S.I., Nekrasov A.N. A numerical measure of amino acid residues similarity based on the analysis of their surroundings in natural protein sequences. *Protein Engineering.* 2001. V. 14. P. 459–463.
11. Guex N., Peitsch M.C. SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis.* 1997. V. 18. P. 2714–2723.
12. Данилкович А.В., Соколов Е. В., Тихонов Д.А., Шадрин Т. Е., Удовиченко И.П. Молекулярная динамика комплексов самоорганизующихся ионных пептидов (RADA)₄. *Мат. биол. и биоинф.* 2011. Т. 6. С. 92–101. doi: [10.17537/2011.6.92](https://doi.org/10.17537/2011.6.92)
13. Danilkovich A.V., Sobolev E.V., Tikhonov D.A., Udovichenko I.P., Lipkin V.M. Distinctive H-(RLDL)₄-OH peptide complexes potentiate nanostructure self-assembling in water. *Dokl. Biochem. Biophys.* 2012. V. 443. P. 96–99.
14. van Gunsteren W.F., Billeter S.R., Eising A.A., Hünenberger P.H., Krüger P., Mark A.E., Scott W.R.P., Tironi I.G. *Biomolecular Simulation: The GROMOS96 Manual and User Guide*. Switzerland, Zürich: Vdf Hochschulverlag AG an der ETH Zürich, 1996. 1042 p.
15. Martenson R.E. Myelin basic protein speciation. In: *Experimental Allergic Encephalomyelitis: A Useful Model for Multiple Sclerosis: Progress in Clinical and Biological Research*. V. 146. Eds. E.C. Alvord Jr., M.W. Kies, A.J. Suckling. New York: Alan R. Liss, Inc, 1983.
16. Stepaniak J.A., Gould K.E., Sun D., Swanborg R.H. A comparative study of experimental autoimmune encephalomyelitis in Lewis and DA rats. *J. Immunol.* 1995. V. 155. P. 2762–2769.
17. Burns F.R., Li X.B., Shen N., Offner H., Chou Y.K., Vandenbark A.A., Heber-Katz E. Both rat and mouse T cell receptors specific for the encephalitogenic determinant of

myelin basic protein use similar $V\alpha$ and $V\beta$ chain genes even though the major histocompatibility complex and encephalitogenic determinants being recognized are different. *J. Exp. Med.* 1989. V. 169. P.27–39.

18. Lenschow D.J., Walunas T.L., Bluestone J.A. CD28/B7 system of T cell costimulation. *Annu. Rev. Immunol.* 1996. V. 14. P. 233–258.

Рукопись поступила в редакцию 01.12.2016.

Дата опубликования 08.12.2016.