



УТВЕРЖДАЮ

Директор

ФГБУН Институт математических проблем  
биологии РАН – филиал  
ИММ им. М.В. Келдыша РАН  
д.ф.-м.н., проф. В. Д. Лахно

«29» апреля 2016 г.

### ОТЗЫВ

ведущей организации на диссертационную работу

Занегиной Ольги Николаевны

на тему

“Сравнительная характеристика структур ДНК-белковых комплексов”,  
представленную на соискание ученой степени кандидата биологических наук по  
специальности 03.01.09 (математическая биология, биоинформатика).

**Актуальность для науки и практики.** Актуальность темы диссертации О.Н. Занегиной определяется важностью понимания механизмов узнавания белком ДНК, для чего, в свою очередь, необходимо изучение структур ДНК-белковых комплексов. Знание закономерностей узнавания ДНК такими белками, как транскрипционные факторы, эндонуклеазы, метилтрансферазы может быть полезно как для фундаментальных исследований, так и для приложений к медицине и биотехнологии. В частности, возможность предсказания специфичности транскрипционных факторов может облегчить изучение систем регуляции бактерий.

**Новизна и значимость основных научных результатов.** Работа носит теоретический характер и проведена на основе анализа всех доступных

структур ДНК-белковых комплексов. Основные научные результаты, полученные автором, таковы:

1. Предложена классификация структур комплексов белковых доменов с ДНК, основанная на взаимодействующих структурных элементах белка и ДНК. Принцип классификации позволяет обновлять её при появлении новых структур. В работе классифицированы 1942 структуры, относящиеся к 314 ДНК-контактирующим доменам.

2. Впервые предложена классификация семейств ДНК-узнающих белковых доменов (в терминологии номенклатуры SCOP). Отнесение семейства к определенному классу (всего выделено 17 таких классов) производится на основе консервативных особенностей ДНК-белкового взаимодействия, характерных для семейства. Класс определен для всех семейств, представленных тремя и более различными доменами в комплексах с ДНК.

3. Разработан подход к описанию консервативно расположенных молекул воды в структурах гомологичных макромолекул, прежде всего в структурах ДНК-белковых комплексов, включающих гомологичные белки.

4. Предложенные классификации интегрированы в базу данных взаимодействий нуклеиновых кислот с белками (NPIDB), что упрощает её использование при анализе ДНК-белковых структур. Также в NPIDB интегрировано подробное описание семейств ДНК-белковых комплексов, включающее описание консервативных молекул воды.

5. С использованием функционала базы данных NPIDB разработан подход к структурному анализу консервативных особенностей гомологичных белков. Данный подход был применён в анализе консервативных особенностей узнавания ДНК в семействах TATA-бок узнающих белков и семейства LAGLIDADG\_1.

Результаты получены при участии автора впервые и опубликованы в четырёх научных статьях в международных журналах "Journal of Bioinformatics and Computational Biology" и "Nucleic Acids Research".

**Структура работы.** Диссертация состоит из введения, пяти глав, заключения и списка литературы. Общий объем работы составляет 138 страниц, включая 6 таблиц и 50 рисунков. Библиография содержит 137 ссылок.

Первая глава содержит подробный литературный обзор, рассматривающий физические основы ДНК-белковых взаимодействий, механизмы ДНК-белкового узнавания, созданные ранее классификации ДНК-белковых взаимодействий и баз данных структур белков, ДНК и их комплексов.

Во второй главе описываются предложенные автором классификации ДНК-белковых комплексов и их семейств. Классификация индивидуальных комплексов применена к 1942 структурам комплексов белковых доменов с ДНК, классификация семейств — к 34 семействам ДНК-узнающих белковых доменов SCOP. На основе проведенного анализа описаны закономерности ДНК-белкового взаимодействия. В частности, отмечен вклад гидрофобных контактов в формирование ДНК-белкового комплекса и приведены возможные причины вариаций способов узнавания белковым доменом ДНК в рамках одного семейства SCOP.

В третьей главе описываются результаты поиска консервативных молекул воды в семействах гомологичных белков. Исследована роль консервативных молекул воды в формировании активного центра транскетолазы, а также выявлены консервативные молекулы воды в 72 семействах ДНК-связывающих белковых доменов SCOP.

В главе 4 описан функционал базы данных комплексов нуклеиновых кислот с белками (NPIDB), в том числе содержащий разработанную классификацию ДНК-белковых взаимодействий.

Глава 5 посвящена результатам подробного анализа консервативных контактов — прямых и опосредованных водой водородных связей, а также гидрофобных кластеров — в семействах TATA-бок связывающего белка и LAGLIDADG\_1 хоминг эндонуклеаз. Примечательно, что в первом случае были выявлены консервативно контактирующие с ДНК аминокислотные остатки, а во втором — консервативно контактирующие аминокислотные позиции.

### **Практическая значимость**

Полученные результаты, в частности интеграция разработанных классификаций в базу NPIDB, могут иметь значение для понимания механизмов и закономерностей ДНК-белкового узнавания, что может способствовать прогрессу в таких областях, как медицина и фармакология (создание противобактериальных, противоопухолевых и противовоспалительных препаратов), биотехнология и сельское хозяйство (повышение производительности сельскохозяйственных животных и растений, борьба с вредителями и возбудителями заболеваний). Сказанное выше особенно важно ввиду быстрого роста объема структурной информации о ДНК-белковых комплексах.

### **Рекомендации по использованию результатов и выводов диссертации**

Разработанная классификация и подход к структурному анализу гомологичных белков имеют методическую значимость и могут быть использованы в курсах биоинформатики для студентов университетов. Кроме того, описанные в работе подходы и результаты могут быть использованы при анализе конкретных семейств ДНК-узнающих белков, прежде всего транскрипционных факторов. Результаты и выводы диссертации могут быть использованы в следующих научных учреждениях: Федеральное государственное бюджетное учреждение науки Институт математических проблем биологии Российской академии наук — филиал Федерального государственного учреждения «Федеральный исследовательский центр Институт прикладной математики им. М.В. Келдыша Российской академии наук»; Федеральное государственное бюджетное учреждение науки Институт общей генетики им. Н.И. Вавилова Российской академии наук; Федеральное государственное бюджетное учреждение науки Институт молекулярной генетики Российской академии наук; Федеральное государственное бюджетное учреждение науки Институт биологии гена Российской академии наук; Федеральное государственное бюджетное учреждение науки Институт проблем передачи информации им. А.А. Харкевича Российской академии наук;

Федеральное государственное бюджетное образовательное учреждение высшего образования «Московский государственный университет имени М.В.Ломоносова», а также в других научно-исследовательских организациях, в которых ведутся исследования по молекулярной биологии.

#### **Общие замечания.**

Текст диссертации не свободен от недостатков. Имеются неточные утверждения, например на стр. 52: «в малой бороздке бета-листы встречаются чаще, чем альфа-спирали» — это утверждение противоречит находящемуся на той же странице рисунку 10, из которого следует, что контакт альфа-спирали с малой бороздкой встретился в 19 семействах, а контакт бета-листа с малой бороздкой — в 10 семействах, то есть примерно вдвое реже, а не чаще, как утверждается в тексте. На странице 55 анонсирована закономерность, относящаяся к числу и характеру связей с сахаро-фосфатным остовом, однако дальнейший текст содержит лишь описание двух примеров, и остаётся непонятным, в чём же состоит упомянутая закономерность. Имеется ещё несколько недостатков подобного рода, однако все они не относятся к основным результатам работы и не могут быть причиной её низкой оценки.

#### **Заключение**

Диссертация представляет собой завершённую научно-исследовательскую работу на актуальную тему. Новые научные результаты, полученные диссертантом, имеют существенное значение для биологической науки и могут быть использованы специалистами в области молекулярной биологии, фармакологии, биоинженерии и биотехнологии. Выводы и рекомендации достаточно обоснованы. Работа отвечает требованиям п. 9 «Положения о порядке присуждения ученых степеней», предъявляемым к кандидатским диссертациям, а её автор заслуживает присуждения ей ученой степени кандидата биологических наук по специальности 03.01.09 — «Математическая биология, биоинформатика».

Отзыв на диссертацию и автореферат обсужден на расширенном семинаре  
лаборатории прикладной математики ИМПБ РАН — филиала ИПМ им. М.В.  
Келдыша РАН 21.04.2016 г., протокол № 5.

Научный сотрудник лаборатории прикладной математики  
ФГБУН ИМПБ РАН — филиала ИПМ им. М.В. Келдыша РАН  
к.ф.-м.н.  
28 апреля 2016 г.

*V. V. Yakovlev*

В.В. Яковлев  
тел. (4967) 318530

Подпись сотрудника лаборатории прикладной математики, к.ф.-м.н. Виктора  
Владимировича Яковлева удостоверяю

Начальник отдела кадров  
ИМПБ РАН – филиала ИПМ им. М.В. Келдыша Р



Т.А.Галушко