

**Программа утверждена на заседании
Ученого Совета химического факультета
Протокол № 4 от 29 мая 2014 г.**

Декан химического факультета,
Акад. РАН, профессор



/В.В. Лунин/

Рабочая программа дисциплины (модуля)

1. Наименование дисциплины (модуля): **Биоинформатика и основы компьютерного моделирования биоструктур**

Цель: Знакомство аспирантов с основами биоинформатики и особенностями компьютерного моделирования структуры природных биополимеров и надмолекулярных ансамблей.

Задачи: Получение аспирантом навыков компьютерного моделирования биоструктур

2. Уровень высшего образования аспирантура.

3. Направление подготовки: 04.06.01 Химические науки, направленность (профиль) 03.01.04 Биохимия, 03.01.06 Биотехнология, 02.00.15 Кинетика и катализ.

4. Место дисциплины (модуля) в структуре ООП: вариативная часть ООП, блок «Дисциплины (модули)»

5. Планируемые результаты обучения по дисциплине (модулю), соотнесенные с планируемыми результатами освоения образовательной программы (компетенциями выпускников)

Формируемые компетенции (код компетенции)	Код и наименование индикатора достижения компетенции	Планируемые результаты обучения по дисциплине (модулю)
УК-1 способность к критическому ана-	УК-1.1 анализирует методологические про-	Знать: методы критического анализа и

лизу и оценке современных научных достижений, генерированию новых идей при решении исследовательских и практических задач, в том числе в междисциплинарных областях	блемы, возникающих при решении исследовательских и практических задач, альтернативные варианты их решения, проводит оценку достоинств и недостатков реализации этих вариантов	оценки современных научных достижений, а также методы генерирования новых идей при решении исследовательских и практических задач, в том числе в междисциплинарных областях.
СПК-2 способен изучать биосистемы методами биоинформатики и компьютерного моделирования	СПК-2.1 пользуется базами данных для поиска структур белков СПК-2.2 владеет методами молекулярной динамики для моделирования биосистем	Знать: современные структурные базы данных белков Знать: теоретические основы метода молекулярной динамики Уметь: проводить расчеты биосистем методами молекулярной динамики Владеть: навыками компьютерного моделирования биосистем

6. Объем дисциплины (модуля) в зачетных единицах с указанием количества академических или астрономических часов, выделенных на контактную работу обучающихся с преподавателем (по видам учебных занятий) и на самостоятельную работу обучающихся:
Объем дисциплины (модуля) составляет 3 зачетных единицы, всего 108 часов, из которых 58 часов составляет контактная работа студента с преподавателем (48 часов занятия лекционного типа, 6 часов групповых консультаций, 4 часа - мероприятия промежуточной аттестации), 50 часов составляет самостоятельная работа учащегося.

7. Входные требования для освоения дисциплины (модуля), предварительные условия.
Для того чтобы формирование данной компетенции было возможно, обучающийся должен освоить ранее дисциплины «Математический анализ», «Линейная алгебра», «Уравнения математической физики», «Химическая кинетика и катализ», «Квантовая химия», «Физическая химия», «Ферментативный катализ», «Химические основы биологических процессов»

8. Содержание дисциплины (модуля), структурированное по темам.

Наименование и краткое содержание разделов и тем дисциплины (модуля),	Всего (часы)	В том числе	
		Контактная работа (работа во взаимодействии с преподавателем), часы	Самостоятельная работа обучающегося, часы

Форма промежуточной аттестации по дисциплине (модулю)		из них					Всего	из них		
		Занятия лекционного типа	Занятия семинарского типа	Групповые консультации	Индивидуальные консультации	Учебные занятия, направленные на проведение текущего контроля успеваемости, промежуточной аттестации		Выполнение домашних заданий	Подготовка рефератов.п.	Всего
Раздел 1	16	10					10			6
Раздел 2	16	10					10			6
Раздел 3	10	6					6			4
Раздел 4	10	6					6			4
Раздел 5	10	6					6			4
Раздел 6	16	10					10			6
Промежуточная аттестация, Зачет	30			6		4	10			20
Итого	108	48					58			50

Содержание разделов:

Раздел 1. Гомологическое моделирование белковых структур – актуальность и примеры модельных белковых структур. Банки дан-

ных белковых структур и последовательностей PDB и UNIPROT и содержащейся в этих банках информация. Экспериментальные методы расшифровки белковых структур - рентгеноструктурный анализ, метод многомерного ядерного магнитного резонанса, низкотемпературная электронная микроскопия. Структурно-консервативные участки в структуре белков и методы их определения.

Раздел 2. Математические методы построения выравниваний белковых последовательностей (матрицы стоимости замены аминокислотных остатков, метод динамического программирования для нахождения оптимального парного выравнивания белковых последовательностей, проверка значимости парного выравнивания) и алгоритмы поиска гомологов по базам данных белковых последовательностей на примере методов FASTA и BLAST. Стратегия и детали гомологического моделирования третичной структуры белков. Регуляризации полученных модельных структур белков - молекулярная механика и динамика. Встраивание протетических групп и моделирование четвертичной структуры белковой молекулы. Методы оценки правдоподобности построенной белковой структуры и сферы применимости модельной структуры.

Раздел 3. Вторичная и третичная структура нуклеиновых кислот. Методы моделирования третичной структуры нуклеиновых кислот. Методы молекулярной механики и динамики, понятие силового поля. Физический смысл членов составляющих потенциальную энергию молекулы. Способы расчета параметров для молекулярной динамики. Расчет частичных зарядов по схемам esp, CNLPG, рекомендуемые квантово-механические методы. Молекулярная динамика, интегратор. Алгоритмы контроля температуры и давления. Этапы подготовки системы к молекулярной динамике.

Раздел 4. Расчет удаленных электростатических и Ван-Дер-Ваальсовых взаимодействий. Учет поляризации растворителя (PME, reaction-field shift). Информация, получаемая из анализа траекторий, и методы обработки траекторий. Ограничения метода молекулярной динамики. Самосборка бислоя. Управляемая симуляция молекулярной динамики. Полноатомные симуляции молекулярной динамики как способ предсказания структур небольших НК. Метод обмена репликами. Основные проблемы. Загрубленное (Coarse grained) описание НК и применение к моделированию структур больших НК, как в свободном состоянии, так и в связанном с другими биополимерами.

Раздел 5. Современные методы классификации последовательностей и пространственных структур белков (структурные и последовательные домены белков, базы данных PFAM, SCOP, CATH) и подходы к построению множественных выравниваний белковых последовательностей с использованием метода скрытых Марковских моделей. *de novo* предсказание структуры белковых молекул. Развитие методов предсказания структуры белков по их первичной последовательности.

Раздел 6. Молекулярные машины – природные и искусственные. Использование молекулярных машин на основе биомоторов (F₀F₁-ATP synthase, kinesin, dynein, myosin, RNA polymerase) – связь функции и механики. Структурная нанотехнология на основе молекул ДНК. Методы построения наноструктур из молекул ДНК (самосборка).

9. Образовательные технологии.

Занятия проводятся как с помощью традиционных образовательных технологий, так и с применением современных компьютерных программ.

10. Оценочные материалы для проверки результатов обучения по дисциплине (модулю)

Шкала оценивания знаний, умений и навыков является единой для всех дисциплин (приведена в таблице ниже)

ШКАЛА И КРИТЕРИИ ОЦЕНИВАНИЯ РЕЗУЛЬТАТА ОБУЧЕНИЯ по дисциплине (модулю)				
Оценка \ Результат	2	3	4	5
Знания	Отсутствие знаний	Фрагментарные знания	Общие, но не структурированные знания	Сформированные систематические знания
Умения	Отсутствие умений	В целом успешное, но не систематическое умение	В целом успешное, но содержащее отдельные пробелы умение (допускает неточности не принципиального характера)	Успешное и систематическое умение
Навыки (владения)	Отсутствие навыков	Наличие отдельных навыков	В целом, сформированные навыки, но не в активной форме	Сформированные навыки, применяемые при решении задач

Типовые задания или иные материалы, необходимые для оценки результатов обучения, методические материалы, определяющие процедуры оценивания приведены в разделе Фонды оценочных средств.

РЕЗУЛЬТАТ ОБУЧЕНИЯ по дисциплине (модулю)	ФОРМА ОЦЕНИВАНИЯ
<p>Знать: методы критического анализа и оценки современных научных достижений, а также методы генерирования новых идей при решении исследовательских и практических задач, в том числе в междисциплинарных областях</p> <p>Знать: современные структурные базы данных белков</p> <p>Знать: теоретические основы метода молекулярной динамики</p>	<p>мероприятия текущего контроля успеваемости, устный опрос на зачете</p>

Уметь: проводить расчеты биосистем методами молекулярной динамики	ПКЗ, устный опрос на зачете
Владеть: навыками компьютерного моделирования биосистем	ПКЗ, мероприятия текущего контроля успеваемости

11. Учебно-методические материалы для самостоятельной работы по дисциплине (модулю): презентации к лекционным занятиям. Аспирантам предоставляется программа курса, план занятий и задания для самостоятельной работы, презентации к лекционным занятиям.

12. Ресурсное обеспечение:

- Перечень основной и вспомогательной учебной литературы ко всему курсу

Основная литература

1. А Леск Введение в биоинформатику. Пер. с англ. – М.: БИНОМ. Лаборатория знаний, 2009. - 318 с.
2. Х.Д.Хельтге, В. Зиппль, Д. Роньян, Г. Фолькерс Молекулярное моделирование (теория и практика) М. : БИНОМ. Лаборатория знаний, 2010.- 318 с:
3. М.В. Волькенштейн Биофизика. Москва, «Наука», 1981, с.575.
4. С.Д. Варфоломеев Химическая Энзимология. Москва, «Академия», 2005, с. 472.
5. Кантор Ч., Шиммел П. Биофизическая химия. В 3х томах. «Мир», 1984 г, т.2.
6. Страйер Л. Биохимия. В 3х томах, «Мир», 1984 г.
7. Atlas of Protein Sequence and Structure, Suppl 3, 1978, M. O. Dayhoff, ed. National Biomedical Research Foundation, 1979.
8. Koonin E.V., Galperin M.Y. "Sequence - Evolution - Function Computational Approaches in Comparative Genomics". Доступна по адресу <http://www.ncbi.nlm.nih.gov/pubmed/21089240>
9. *Geoffrey J. Barton Protein Sequence Alignment and Database Scanning.* Доступна по адресу http://www.compbio.dundee.ac.uk/papers/rev93_1/rev93_1.html
10. «2Can Support Portal – Bioinformatics» . Справочное руководство по биоинформатике, подготовленное сотрудниками Европейского института биоинформатике. Доступно по адресу <http://www.ebi.ac.uk/2can/home.html>
11. <http://www.ncbi.nlm.nih.gov/BLAST/tutorial/Altschul-1.html>

12. Thompson J.D., Higgins D.G., Gibson T.J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 1994 Nov 11; **22**(22):4673-80.
13. Schwede T., Kopp J., Guex N., and Peitsch M.C. SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Research*, 2003, **31**, 3381-3385.
14. Rachel Karchin. Hidden Markov Models and Protein Sequence Analysis. Доступно по адресу <http://www.cse.ucsc.edu/research/compbio/ismb99.handouts/KK185FP.html>
15. Eddy S. Серия статей в секции "Primer" журнала *Nature Biotechnology*, 2004, **22** (№ 7-10). Доступны по адресу http://www.liacs.nl/~hoogeboom/mcb/nature_primer.html
16. Baker D., Sali A. Protein Structure Prediction and Structural Genomics. *Science*, 2001, **294**, 93-96.
17. Schueler-Furman O., Wang C., Bradley Ph., Misura K., Backer D. Progress in modeling of protein structures and interactions. *Science*, 2005, **310**, 638-642.
18. Computer Simulation of Biomolecular Systems: Theoretical and experimental by Wilfred F. van Gunsteren, Paul K. Weiner, Tony Wilkinson - Science — 1997
19. Computational studies of RNA and DNA By Jiří Šponer, Filip Lankaš, Springer, 2006
20. GROMACS 4.0 manual.
21. DeLano, W.L. The PyMOL Molecular Graphics System (2002) on World Wide Web <http://www.pymol.org>
22. Vienna RNA Package, *Nucleic Acids Res.* 31: 3429-3431 (2003)
23. E. Vanquenef, G. Marquant, J. C. Delepine, P. Cieplak & F.-Y. Dupradeau. R.E.D. Server: a web service designed to derive RESP and ESP charges of high quality and reproducibility and to generate force field libraries for new molecules and new molecular fragments. Université de Picardie Jules Verne - Burnham Institute for Medical Research (2009).
24. Gregory R. Bowman, Xuhui Huang, Yuan Yao, Jian Sun, Gunnar Carlsson, Leonidas J. Guibas, and Vijay S. Pande, *J. Am. Chem. Soc.*, 130 (30), 9676–9678, 2008

- Описание материально-технической базы.

Лекционные занятия проводятся в специально оборудованной аудитории (к.202 кафедры химической энзимологии). Вспомогательный материал в виде презентаций доступен студентам

13. Язык преподавания – русский

14. Преподаватели:

кафедра химической энзимологии химического факультета МГУ имени М.В.Ломоносова
к.ф.-м.-н., ст.н.с Упоров Игорь Владимирович, e-mail: @.ru, телефоны (495)-939-

Фонды оценочных средств, необходимые для оценки результатов обучения

Образцы оценочных средств для текущего контроля усвоения материала и промежуточной аттестации - зачета. На зачете проверяется достижение промежуточных индикаторов компетенций, перечисленных в п.5.

1. Материалы к промежуточной аттестации (ПКЗ к зачету)

Вариант №1.

1. В чём отличие последовательностей, депонированных в БД белковых последовательностей SWISS-PROT, от последовательностей в БД TrEMBL:
 - а) Последовательности белков из ограниченного круга организмов
 - б) Курируемые последовательности белков
 - с) Последовательности белковых фрагментов
2. Сколько генов ферментов содержится в геноме человека (выберите наиболее близкое значение):
 - а) 1000
 - б) 2500
 - с) 27000
3. Какой из экспериментальных методов обладает возможностью самого точного разрешения пространственной структуры биополимеров:
 - а) Метод рентгено-структурного анализа
 - б) Спектроскопия гигантского комбинационного рассеяния
 - с) Метод низкотемпературной электронной микроскопии
4. Какая информация не содержится в БД белковых структур:
 - а) Молекулярная масса молекулы
 - б) Химическая формула простетических групп
 - с) Информация о четвертичной структуре белка
5. Выберите правильное определение белков-гомологов
 - а) Белки-гомологи – это белки демонстрирующие сходный процентный аминокислотный состав первичной последовательности

- b) Белки-гомологи – это белки, обладающие сходными физико-химическими свойствами (молекулярная масса и размер, значения pI, сходная подвижность в электрофорезе, пр.)
 - c) Белки-гомологи – это белки со схожей аминокислотной последовательностью из организмов, имеющих общего предшественника
6. Какая идея лежит в основе в основе алгоритма поиска схожих пространственных участков полипептидной цепи в белках:
- a) Поиск участков со схожей локальной конформацией полипептидной цепи
 - b) Поиск участков со схожей конформацией боковых остатков аминокислотных остатков
 - c) Поиск участков со схожей локальной конформацией полипептидной цепи и с одинаковым пространственным расположением этих участков в белковой глобуле

7. Вычислите значение стоимости следующего выравнивания с использованием единичной матрицы замен:

RKYLESNIGHTLVVTSDDKG-PDSVF
ANWLKDGHELIITSDKEGE-TSEL

- a) 5
 - b) 10
 - c) 24
8. Какой из аминокислотных остатков являются наиболее консервативными с эволюционной точки зрения (вероятность мутации минимальна):
- a) C
 - b) W
 - c) S
9. Какой математический метод используется при построении оптимального выравнивания двух белковых последовательностей:
- a) Метод наименьших квадратов
 - b) Метод динамического программирования
 - c) Метод суммирования числовых последовательностей
10. Какое утверждение свидетельствует о достоверности парного выравнивания, построенного с помощью алгоритма BLAST:
- a) Длина найденных гомологичныхх участков > 10 остатков
 - b) $10 < \text{Bit score} < 100$
 - c) $E < 0.01$
11. На основе последовательности какой молекулы строится дендрограмма в проекте «Tree of life»
- a) 15S sRNA
 - b) 16S sRNA
 - c) Нуклеиновой кислоты, входящей в 30S субъединицу рибосомы
12. Перечислите требования предъявляемые к модельной структуре:

13. Основной идеей применения метода молекулярной механики и динамики к биополимерам является:
- a) Применение к биополимерам уравнений Максвелла
 - b) Применение к биополимерам квазиклассического приближения квантовой механики
 - c) Применение к молекулам классической механики
14. Каким потенциалом описывается потенциальная энергия химической связи между атомами в классических силовых полях
- a) Потенциал Морзе
 - b) Гауссовским потенциалом
 - c) Гармоническим осциллятором
 - d) Кулоновским потенциалом
15. Перечислите все составляющие потенциальной энергии молекулы в методе МД:
16. Учёт экранирования электростатического взаимодействия производится путём:
- a) Введения двух значений диэлектрической проницаемости для растворителя и внутрибелкового пространства
 - b) Применением потенциала 6-12 для описания электростатического взаимодействия
 - c) Выбором диэлектрической проницаемости пропорциональной расстоянию между двумя зарядами
17. Какое значение шага интегрирования используется в МД расчётах биополимеров и чем оно определяется:
- a) 10 пс, конформационными перестройками боковых остатков аминокислотных остатков
 - b) 5 пс, колебаниями длин связей между тяжёлыми атомами
 - c) 1 фс, периодом колебаний C-H связей
18. Какова точность совпадения структурных доменов в белках (БД SCOP) с доменами из БД белковых семейств (PFAM):
- a) Точно совпадают
 - b) 5-10 остатков
 - c) 30-50 остатков

Вариант №2.

19. Выберите число наиболее близкое к количеству депонированных белковых последовательностей в БД TrEMBL в настоящее время:
- a) 10000
 - b) 1000000
 - c) 10000000
20. Выберите правильную последовательность аминокислот, упорядоченную в порядке естественной встречаемости (по возрасту):

- a) ACDEF
- b) PKLSC
- c) WCHAL

21. Какое количество фолдов белков известно к настоящему времени (выберите наиболее близкое число):

- a) 2100
- b) 5700
- c) 50000

22. Какова величина химического сдвига метильной группы в спектрах ЯМР (выберите правильный диапазон):

- a) 2,2 – 2,5 ppm
- b) 8,0 – 8,5 ppm
- c) 6,3 – 6,6 ppm

23. Какая из характеристик белка является более консервативной с эволюционной точки зрения:

- a) Аминокислотная последовательность
- b) Функциональное предназначение
- c) Пространственная укладка полипептидной цепи (фолдинг)

24. Каково значение среднеквадратичного отклонения атомов полипептидной цепи структурно-консервативных пространственных участков белков при их наложении друг на друга (выберите правильный диапазон):

- a) 0,1-1,5 Å
- b) 2-3Å
- c) Более 3Å

25. Какая из перечисленных эволюционных матриц замен аминокислотных остатков описывает гомологии между эволюционно удалёнными белковыми последовательностями:

- a) PAM100
- b) BLOSUM62
- c) PAM256

26. Положительные элементы эволюционной матрицы стоимости замен аминокислотных остатков свидетельствуют о том, что

- a) Вероятность мутации в процессе эволюции больше чем вероятность случайной мутации
- b) Эти аминокислоты имеют схожие физико-химические свойства боковых остатков
- c) Эти два аминокислотных остатка наиболее часто встречаются в последовательностях белков, депонированных в БД

27. Какие из перечисленных утверждений правильны:

- a) Метод построения выравнивания белковых последовательностей FASTA позволяет построить глобальное выравнивание последовательностей

- b) Метод построения выравнивания белковых последовательностей BLAST позволяет выявить только по одному локальному гомологичному участку в каждом из белков
 - c) Алгоритмы парного выравнивания белковых последовательностей, основанные на основе методе динамического программирования, позволяют находить более оптимальное выравнивание чем методы FASTA и BLAST
28. Какие из ограничений применимы к методу построения множественных выравниваний ClustalW?
- a) Количество аминокислотных последовательностей не должно превышать 30 и их суммарная длина не превосходит 10000 остатков
 - b) Аминокислотные последовательности должны демонстрировать парную гомологию со степень идентичности не менее 25%
 - c) Компьютер должен быть оснащён не менее чем 1 Gb оперативной памяти
29. Какова последовательность присвоения координат атомов в процессе моделирования по гомологиям
- a) Присваиваются координаты атомам полипептидной цепи структурно-консервативных участков, петель и боковых цепей всех остатков
 - b) Присваиваются координаты атомам полипептидной цепи петель, боковых цепей всех остатков, полипептидной цепи структурно-консервативных остатков
 - c) Вычисляются координаты атомов боковых цепей аминокислотных остатков в развёрнутой конформации, вычисляются координаты атомов полипептидной цепи петель, присваиваются координаты атомам полипептидной цепи консервативных остатков
30. Выберите минимальное значение степени идентичности между белковыми последовательностями, которое позволяет применять метод построения пространственной структуры белков по гомологиям:
- a) 80%
 - b) 17%
 - c) 40%
31. Понятие силового поля включает в себя:
- a) Деление атомов по типам
 - b) Формулу вычисления потенциальной энергии молекулы и параметры этой функции
 - c) Алгоритмы определения типов атомов и присвоения соответствующих параметров членам в функции потенциальной энергии
 - d) Все перечисленные выше элементы
32. Молекулярная механика/динамика неприменимы для описания:
- a) Равновесных структур белков и нуклеиновых кислот
 - b) Латерального движения белков в биологических мембранах
 - c) Переноса протонов через биологические мембраны

33. Выберите правильный диапазон значения силовой константы K_2 одинарной химической связи
- a) 350-490 (kcal/(molÅ²))
 - b) 540-8700 (kcal/(molÅ²))
 - c) 55-87 (kcal/(molÅ²))
34. Выберите правильную последовательность методов минимизации при типичной минимизации потенциальной энергии молекулы
- a) Метод наискорейшего спуска → метод сопряжённых градиентов → метод Ньютона-Рафсона
 - b) Метод сопряжённых градиентов → Метод наискорейшего спуска
 - c) Метод наискорейшего спуска → метод Ньютона-Рафсона
35. Какой из алгоритмов построения множественного выравнивания белковых последовательностей является наиболее чувствительным в настоящее время:
- a) ClustalW
 - b) PSI-BLAST
 - c) Методы, основанные на формализме скрытых марковских моделей

Методические материалы для проведения процедур оценивания результатов обучения

Зачет проводится по билетам. В ходе сдачи зачета проверяется, в первую очередь, формирование «знаниевой» компоненты компетенций, перечисленных в п.5, а также сформированность перечисленных в п.5 умений. Уровень знаний аспиранта по каждому вопросу оценивается на «отлично», «хорошо», «удовлетворительно», «неудовлетворительно». В случае, если на все вопросы был дан ответ, оцененный не ниже чем «удовлетворительно», аспирант получает общую оценку «зачтено». Ведомость приема зачета подписывается членами комиссии.