

УДК 577.15

## Моделирование свойств и реакций биомолекулярных систем комбинированными методами квантовой и молекулярной механики

А. В. Немухин, Б. Л. Григоренко, Е. М. Епифановский, А. А. Московский

*АЛЕКСАНДР ВЛАДИМИРОВИЧ НЕМУХИН — профессор, доктор химических наук, заведующий лабораторией химической кибернетики Химического факультета МГУ им. М.В. Ломоносова, заведующий лабораторией Института биохимической физики им. Н.М. Эмануэля РАН. Область научных интересов: молекулярное моделирование, квантовая химия.*

*БЕЛЛА ЛЮДВИГОВНА ГРИГОРЕНКО — доктор физико-математических наук, старший научный сотрудник Химического факультета МГУ им. М.В. Ломоносова. Область научных интересов: молекулярное моделирование, квантовая химия.*

*ЕВГЕНИЙ МИХАЙЛОВИЧ ЕПИФАНОВСКИЙ — аспирант Химического факультета МГУ им. М.В. Ломоносова. Область научных интересов: молекулярное моделирование, вычислительные методы квантовой химии.*

*АЛЕКСАНДР АЛЕКСАНДРОВИЧ МОСКОВСКИЙ — кандидат химических наук старший научный сотрудник Химического факультета МГУ им. М.В. Ломоносова и Института биохимической физики им. Н.М. Эмануэля РАН. Область научных интересов: молекулярное моделирование, информационные технологии.*

*119992, Москва, Ленинские горы, 1/3, Химический факультет МГУ им. М.В. Ломоносова, тел. (495)939-10-96, E-mail anet@icc.chem.msu.ru*

### Введение

Молекулярное моделирование становится все более распространенным инструментом для изучения химических и биохимических реакций. Применение компьютерных технологий может оказать заметную поддержку экспериментальным работам, позволяя существенно снизить временные и материальные затраты, в частности, на всестороннее исследование перспективных лекарственных препаратов.

Моделирование важнейших биохимических процессов — реакций ферментативного катализа всегда привлекало внимание теоретиков, однако его приложение к биомолекулярным системам было практически невозможно по следующим причинам. Методы молекулярной механики, активно применяемые, например, при моделировании пространственных структур биомолекул или в молекулярном докинге, не могут характеризовать разрыв и образование химических связей, поскольку параметры силовых полей вокруг изучаемых реакционных систем подбираются по свойствам реперных систем вблизи равновесных конфигураций реагирующих молекул.

Что касается методов квантовой механики, которые позволяют описывать изменения в электронной структуре молекул-участников химической реакции, то их применение ограничено размерами молекулярной системы и вряд ли в ближайшее время следует ожидать надежных результатов полностью квантового анализа систем, включающих более 100 атомов. Ведь даже в этом случае необходимо учитывать влияние достаточно протяженного молекулярного окружения на выделенный небольшой фрагмент системы, анализируемый методами квантовой химии. Надежды на континуальные модели конденсированной среды, рассматривающие, например, белковое окружение как

однородный диэлектрик, оправдываются только в случае получения качественных заключений, но не количественных результатов.

Определенный прорыв в развитии моделирования ферментативного катализа на молекулярном уровне связан с внедрением комбинированных методов квантовой и молекулярной механики (метод КМ/ММ). Согласно основной идее этого подхода [1] часть белковой системы, в которой происходят перераспределения химических связей, включается в квантовую подсистему, и энергии и силы в ней рассчитываются по уравнениям квантовой механики в различных приближениях квантовой химии. Большинство атомов, окружающих эту выделенную центральную часть, относится к молекулярно-механической подсистеме. Предполагается, что приближения КМ/ММ должны учитывать конформационные изменения в белковой матрице, сопровождающие химическую реакцию, и в то же время учитывать влияние конформационных изменений в большой системе на энергетический рельеф пути реакции.

При подобном механистическом подходе к моделированию химических реакций основная роль отводится расчету поверхности потенциальной энергии (ППЭ) изучаемых реакционных систем. В приближениях КМ/ММ энергия каждой точки на ППЭ складывается из энергии квантовой части в поле ММ-подсистемы и молекулярно-механической энергии. Исследование механизма реакции проводится путем анализа стационарных точек на ППЭ, т.е. геометрических конфигураций локальных минимумов, отвечающих реагентам, продуктам и возможным интермедиятам, а также седловых точек, через которые происходят переходы между минимумами. Особое внимание уделяется расчетам энергетических барьеров на пути от реагентов к продуктам. Переход к эксперименталь-

но значимым величинам, таким как константа скорости реакций, осуществляется, как правило, в рамках теории переходных состояний. Наши работы относятся именно к этому направлению, позволяющему рассчитывать точки на ППЭ для больших молекулярных систем, определять геометрические конфигурации стационарных точек и анализировать химические превращения и конформационные изменения вдоль контуров минимальной энергии на ППЭ.

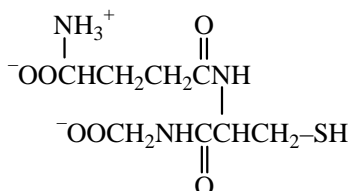
### Методика КМ/ММ

Известно несколько компьютерных реализаций метода КМ/ММ, различающихся способом учета взаимодействия КМ- и ММ-подсистем. Наша оригинальная разработка базируется на теории эффективных фрагментов [2], позволяющей достаточно точно подсчитывать вклады жестких фрагментов ММ-подсистемы в квантовый гамильтониан КМ-части и оценивать обратное влияние квантовой подсистемы на ММ-часть. Под эффективным фрагментом понимается группа атомов с фиксированными внутренними координатами, которая может перемещаться в пространстве как целое и взаимодействовать с другими эффективными фрагментами или с квантовой подсистемой.

Наиболее понятный пример приложения нашего подхода — описание процесса взаимодействия молекул растворителя, в частности воды, как эффективных фрагментов (ММ-подсистема), с молекулами растворенного вещества, которые рассматриваются на квантовом уровне (КМ-подсистема). При таком моделировании свойств раствора достигается большая экономия вычислительных усилий.

В исходном варианте метод КМ/ММ, основанный на концепции эффективных фрагментов, применялся для моделирования свойств ферментативных систем весьма ограниченно [3, 4] именно из-за необходимости поддерживать фиксированными внутренние координаты фрагментов, моделирующих молекулярные группы белковой матрицы. В рамках нашего метода [5-8] пептидные цепи белковых систем разбиваются на малые эффективные фрагменты, что обеспечивает достаточную гибкость окружения активной части фермента, и описание фрагмент-фрагментных взаимодействий (т.е. взаимодействий в ММ-подсистеме) и проводится моделирование фрагмент-фрагментных взаимодействий с помощью известных силовых полей.

Пример разбиения молекулярной системы на квантовую и молекулярно-механическую части иллюстрирует рис. 1. Здесь представлена одна из конформаций молекулы (радикала) глутатиона



К КМ-подсистеме отнесена реакционная сульфгидрильная группа SH. Ближайший «буферный» фрагмент CH<sub>3</sub> принадлежит к КМ- и ММ-подсистемам. Остальная часть молекулы — совокупность малых эффективных фрагментов (CO<sub>2</sub><sup>-</sup>, CH, NH<sub>3</sub><sup>+</sup>, CH<sub>2</sub>, CONH), которые взаимодействуют с квантовой частью

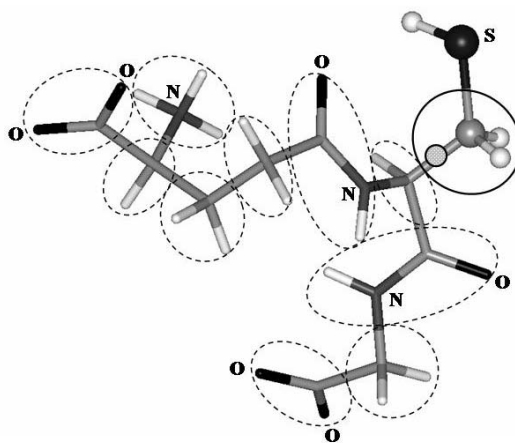


Рис. 1. Разбиение молекулы глутатиона на квантовую часть — группу SH (представлено шарами и стержнями) и ММ-совокупность эффективных фрагментов CO<sub>2</sub><sup>-</sup>, CH, NH<sub>3</sub><sup>+</sup>, CH<sub>2</sub>, CONH (стержни).

Каждый эффективный фрагмент окружен пунктирной линией. «Буферный» фрагмент CH<sub>3</sub> окружен сплошной линией, положение в нем связующего атома водорода отмечено кружком

через вклады в гамильтониан, но между собой взаимодействуют посредством молекулярно-механических силовых полей.

В рамках подобной структурной модели можно моделировать, например, реакции тушения свободных радикалов глутатионом. В таких модельных реакциях [9, 10] к квантовой части структуры глутатиона добавляется второй реагент, в частности, гидроксиметильный радикал CH<sub>2</sub>OH, а к молекулярно-механической части (эффективным фрагментам) — молекулы воды как растворителя. На рис.2 изображен рассчитанный методом КМ/ММ энергетический профиль реакции. В качестве координаты реакции по оси x удобно использовать разность расстояний от передаваемого атома водорода до атома серы R(SH) и до атома углерода R(CH). На вставке под графиком слева внизу изображена конфигурация реагентов с исходным расстоянием от углерода гидроксиметильного радикала CH<sub>2</sub>OH до серы глутатиона 3,76 Å. Иллюстрация геометрической конфигурации системы на вершине потенциального барьера показывает, что передаваемый атом водорода находится на середине пути. Вставка справа относится к концу реакции: образовались метанол CH<sub>3</sub>OH и радикал глутатиона. Важно отметить, что при оценке энергии активации (16,5 ккал/моль) были учтены вклады в энергию от изменяющихся в ходе реакции конформаций трипептида глутатиона и от молекул растворителя — воды.

Другой пример разбиения белковой структуры на подсистемы показан на рис. 3, на котором изображена молекула модельного субстрата для реакции гидролиза пептидной связи под действием трипсина. В рамках данной модели молекулярные группы ближайшего окружения пептидной связи CO—NH (представлены шарами и стержнями) описываются уравнениями квантовой химии. Остальная часть молекулярной системы (представленная линиями) отнесена к ММ-подсистеме, она рассматривается как совокупность

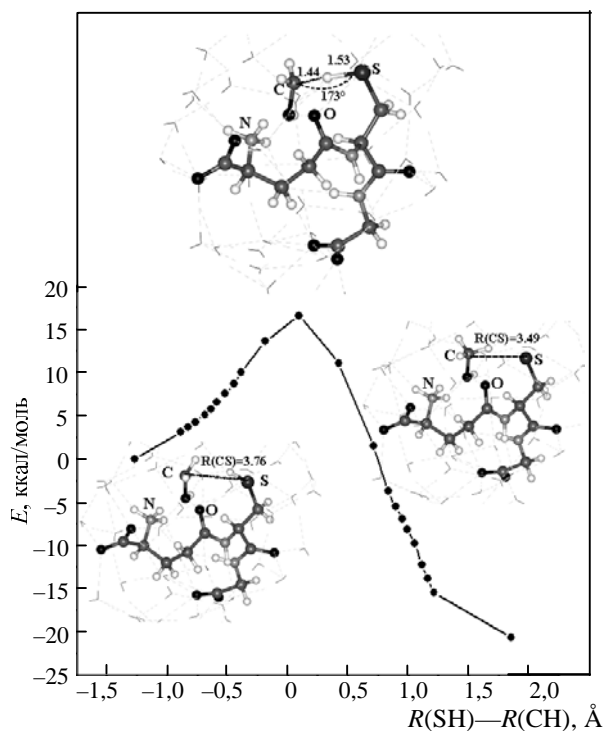


Рис. 2. Рассчитанный энергетический профиль реакции передачи атома водорода от группы SH глутатиона к гидроксиметильному радикалу  $\text{CH}_2\text{OH}$

эффективных фрагментов. На рис. 3 показаны положения связующих (или замыкающих) атомов водорода (выделены заштрихованными кругами), которые располагаются вдоль ковалентных связей, через которые проходит КМ/ММ-граница [6, 7].

#### Примеры исследования биомолекулярных систем методом КМ/ММ

##### Каталитический цикл трипсина и перемещения протонов вдоль водородных связей

Одним из первых объектов практического приложения новой методики был каталитический цикл реакции гидролиза пептидных связей сериновыми протеазами [11]. Ферменты этой группы содержат каталитическую триаду из аминокислотных остатков серина Ser195, гистидина His57 и аспарагиновой кислоты Asp102 (порядковые номера аминокислотных остатков химотрипсина). Существенное значение для действия этих ферментов имеют также группы NH остатков глицина Gly193 и Ser195, способствующие образованию так называемой оксианионной дыры, обеспечивающей стабилизацию тетраэдрических интермедиатов и на стадии ацилирования фермента при расщеплении пептидной связи в субстрате, и на стадии деацилирования. На пути реакции гидролиза молекулярная система проходит через фермент-субстратный комплекс (ES), первый тетраэдрический интермедиат (INT1), фермент-ацильный комплекс (EA), второй тетраэдрический интермедиат (INT2) [11]. Каждой структуре

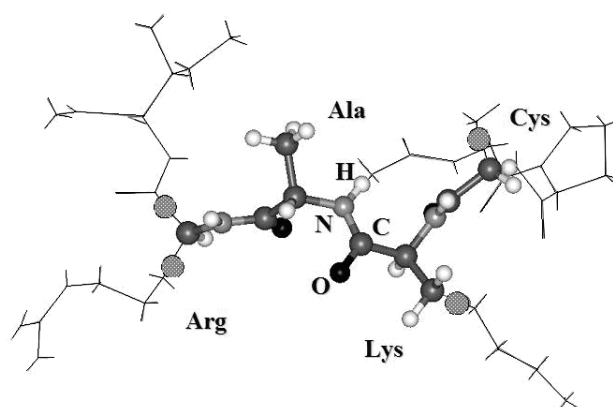


Рис. 3. Разбиение молекулы субстрата для трипсина на квантовую часть (представлено шарами и стержнями) и ММ-совокупность эффективных фрагментов (линии)

отвечает область локального минимума на потенциальной поверхности.

При моделировании каталитического цикла трипсина [12] в качестве начальных координат атомов был взят комплекс трипсин-ингибитор ВРТИ (код 2РТС из банка данных белковых структур [13]). Из молекулы ингибитора был выделен участок, моделирующий субстрат (см. рис. 3). В квантовую область взаимодействующей системы субстрат—фермент были включены фрагменты каталитической триады фермента и центральная часть модельного субстрата около разрываемой пептидной связи. Молекулярно-механическая подсистема включала примерно 750 атомов, расположенных в пределах 18 Å от кислородного центра серина каталитической триады. Расчеты подсистемы в квантовой области при локализации точек минимумов на потенциальной поверхности были проведены неэмпирическим методом самосогласованного поля Хартри—Фока с базисом 6-31G. Параметры молекулярно-механической части были выбраны по библиотеке AMBER [14]. Энергии в найденных стационарных точках ППЭ пересчитывались в квантовохимических приближениях более высокого уровня, т.е. по теории возмущений с расширенными базисами.

На рис. 4 показана суперпозиция рассчитанного методом КМ/ММ участка фермент-субстратного комплекса и экспериментальной структуры фермента с ингибитором 2РТС. Видно, что ключевые для реакционного центра расстояния воспроизводятся методом КМ/ММ достаточно хорошо.

Для выбранной системы были найдены равновесные геометрические конфигурации как точки минимумов на ППЭ, отвечающие характерным структурам на пути всего каталитического цикла (ES, INT, EA). Показано, что результаты моделирования методом КМ/ММ полностью соответствуют химическим представлениям о механизме ферментативного катализа сериновыми протеазами.

В качестве иллюстрации на рис. 5 и 6 изображены структуры, относящиеся к лимитирующей стадии реакции, т.е. переходу от фермент-субстратного комплекса (ES) к первому тетраэдрическому интермедиату

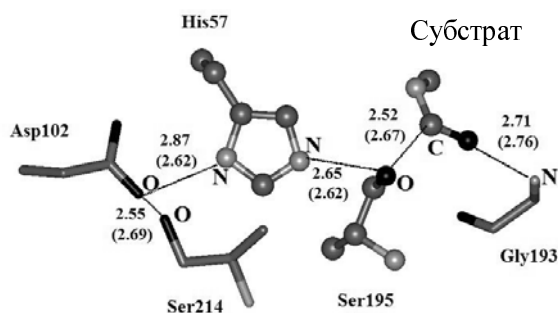


Рис. 4. Суперпозиция рассчитанного методом КМ/ММ участка фермент(трипсин)-субстратного комплекса и экспериментальной структуры комплекса трипсина с ингибитором 2PTC.

Расстояния между «тяжелыми» атомами указаны в ангстремах, в скобках — экспериментальные значения для кристалла комплекса трипсина

(INT1). Показаны химические (верхние части рисунков) и молекулярные (нижние части) структуры. На данной стадии происходит перенос протона от остатка серина Ser195 к гистидину His57, сопровождающийся образованием ковалентной связи O—C комплекса серин-субстрат.

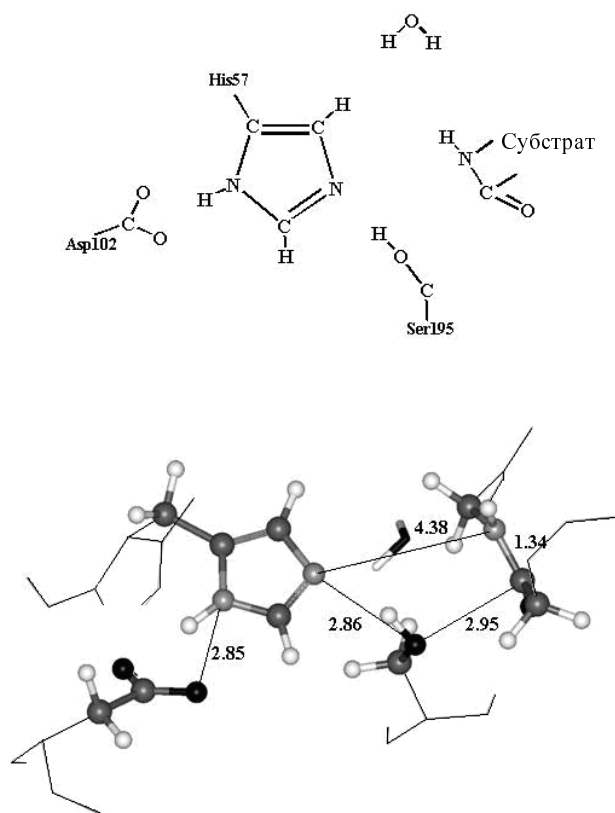


Рис. 5. Рассчитанная конфигурация фермент(трипсин)-субстратного комплекса (ES).

Межъядерные расстояния указаны в ангстремах

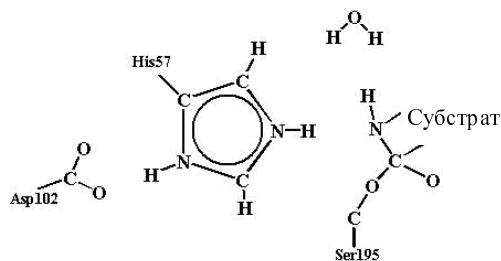


Рис. 6. Рассчитанная конфигурация первого тетраэдрического интермедиата (INT1) в реакции гидролиза субстрата трипсином.

Межъядерные расстояния указаны в ангстремах

Вычисленные в рамках метода КМ/ММ активационные барьеры не превышают 10 ккал/моль, что находится в разумном согласии с экспериментальными данными [11]. Важно отметить решающую роль белковой матрицы — по расчетам энергетического профиля стадии ацилирования, выполненным для активного центра сериновых протеаз без учета окружающих пептидных молекулярных групп методами квантовой химии, оценка энергетических барьеров составила более 25 ккал/моль [15].

Одна из причин низких энергий активации реакций ферментативного катализа связана с тем, что перемещения протонов вдоль водородных связей в белковой структуре происходят с малыми затратами энергии. В нашей работе [16] было выполнено моделирование процесса передачи протонов по ориентированным сетям водородных связей в белковом окружении. На рис. 7 показана одна из таких систем, описывающая участки канала мембранного белка М2. Передача протона осуществляется путем последовательного перемещения молекул воды в полости канала, и очевидным препятствием на этом пути до аминокислотного остатка аспарагиновой кислоты является наличие остатка гистидина.

Результаты моделирования методом КМ/ММ [16] показали, что рельеф поверхности потенциальной энергии для перемещения протона вдоль ориентированных цепей молекул воды, прикрепленных к молекулярным стенкам белка, оказывается настолько пологим, что реориентация молекул воды позволяет избежать энергетических затрат на перегруппировку химических и водородных связей. Единственный барьер на пути протона до концевой группы Asp возникает из-за необходимости разрывать связь  $N_{\epsilon}-H$  в His, расчетные оценки этой величины приводят к значениям, не превышающим 5 ккал/моль.

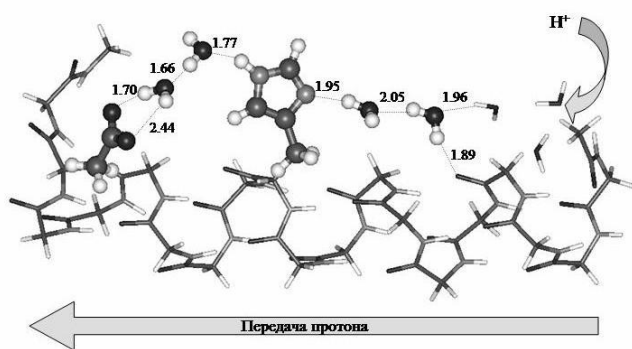


Рис. 7. Молекулярная система, участвующая в передаче протона в полости канала мембранного белка M2.

Межъядерные расстояния указаны в ангстремах. Шарами и стержнями изображены молекулярные группы, отнесенные к квантовой подсистеме: четыре молекулы воды и боковые цепи гистидина и аспарагиновой кислоты. Уменьшенными стержнями показаны молекулярные группы, отнесенные к молекулярно-механической подсистеме — три молекулы воды и полиглициновая спираль, моделирующая стенку мембранного белка.

#### Гидролиз гуанозинтрифосфата белковым комплексом RAS-GAP

Важнейшая ферментативная реакция гидролиза гуанозинтрифосфата (*GTP*), приводящая к гуанозиндифосфату (*GDP*) и неорганическому фосфату (*Pi*)



является лимитирующей стадией всего цикла связывание/гидролиз гуанозинтрифосфата, ответственного за передачу сигналов на рост клеток. В результате реакции (1) *GTP/GDP*-связывающие белки (*G*-белки) переходят из активной формы в неактивную. Замена *GDP* на *GTP* восстанавливает активную форму белков.

Эффективно гидролиз проходит в том случае, когда *G*-белок, содержащий *GTP*, образует комплекс с активирующим белком *GAP*. Сбой в работе этой молекулярной машины, в частности, замедление или прекращение реакции (1) при определенных мутациях одного из *G*-белков человека, p21<sup>ras</sup> (далее для краткости *RAS*), приводит к развитию раковых опухолей.

Экспериментальному и теоретическому изучению реакции гидролиза гуанозинтрифосфата посвящено огромное число публикаций, и, тем не менее, четких и непротиворечивых ответов на вопросы о механизме реакции (1) в различных средах и о роли ключевых аминокислотных остатков в *G*-белках не дано. Довольно распространенной точкой зрения является гипотеза ассоциативного механизма реакции гидролиза (1) по

схеме субстрат-ассистирующего катализа [17], согласно которой субстрат, т.е. молекула *GTP*, а не какая-либо молекулярная группа фермента (*G*-белка) играет роль общего основания. Однако все попытки рассчитать энергетический профиль реакции, протекающей по этому механизму, в рамках различных молекулярных моделей приводят к нереально высоким активационным барьерам — более 40 ккал/моль.

В наших работах [18, 19] методом КМ/ММ было выполнено моделирование реакции гидролиза гуанозинтрифосфата для случая с реальным белковым окружением.

Ранее, в 1997 г., по данным рентгеноструктурного анализа была определена структура комплекса *RAS-GAP* (p21<sup>ras</sup>-p120<sup>GAP</sup>) с аналогом *GTP* — системой *GDP+AlF<sub>3</sub>* [20] (эта работа является фактически единственным источником информации о строении активного центра фермента *RAS* в комплексе с ускорителем *GAP*). Обычно считается, что структура 1WQ1 из банка данных белковых структур [13], относящаяся к системе *RAS-GAP-GDP+AlF<sub>3</sub>*, соответствует переходному состоянию биосистемы в реакции гидролиза (1), поскольку плоская группа *AlF<sub>3</sub>* имитирует предполагаемую конфигурацию  $\gamma$ -фосфатной группы *GTP* на вершине энергетического барьера. Представление о структуре 1WQ1 дает рис. 8.

На основе атомных координат структуры 1WQ1 была построена молекулярная модель для расчетов методом КМ/ММ, включающая 43 атома в КМ-подсистеме и 1622 атома в ММ-части. Общий размер выбранной модельной системы, в которую входят *GTP*, гидролитическая вода, ближайшие к реакционной зоне аминокислотные остатки белков *RAS* и *GAP* и молекулы воды, составляет примерно 40 Å.

В квантовую подсистему включены фосфатные группы *GTP*, молекула воды, катион магния (который кроме фосфатных групп *GTP* координирует также отнесенные к молекулярно-механической подсистеме

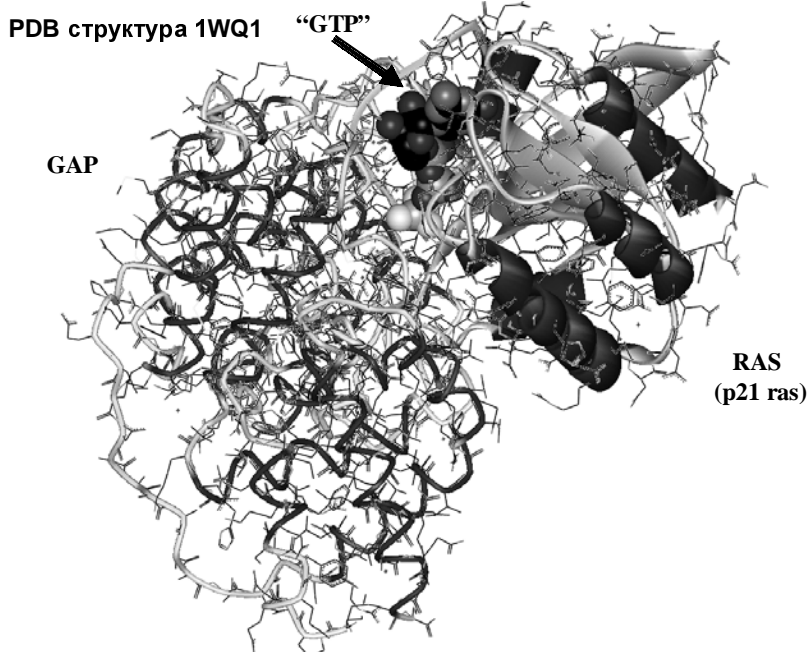


Рис. 8. Структура белкового комплекса 1WQ1

боковые цепи аминокислотных остатков Thr35 и Ser17 от RAS и две молекулы воды), фрагмент глутамина Gln61 от RAS, фрагмент Arg789 от GAP («аргининовый палец»). Именно эти аминокислотные остатки считаются ключевыми при функционировании всего ферментативного комплекса [5]. В процедуру оптимизации геометрических параметров по минимуму полной энергии включены все 43 атома квантовой части и большинство атомов ММ-части. В расчетах КМ-части использовали неэмпирический метод Хартри—Фока с базисом 6-31G, расчет в ММ-части проводили по параметрам силового поля AMBER.

На рис. 9 сопоставлены расстояния между некоторыми «тяжелыми» атомами, полученные для равновесной геометрической конфигурации модельной системы, и экспериментальные расстояния, полученные для структуры 1WQ1 при разрешении 2,5 Å. Следует отметить, что в кристалле вместо GTP находится аналог (GDP+AlF<sub>3</sub>), причем AlF<sub>3</sub> располагается на месте γ-фосфатной группы GTP.

Наиболее важные результаты сопоставления геометрических параметров следующие. Для подавляющего большинства координат молекулярных групп внутри белковой системы получено достаточно хорошее согласие рассчитанных и экспериментальных параметров (с учетом погрешности разрешения). В частности, расстояния между карбонильным кислородом Arg789 и азотом Gln61 практически совпадают. Этот параметр весьма важен, поскольку одна из функций остатка Arg789 от GAP состоит в том, чтобы «подтянуть» остаток Gln61 от RAS достаточно близко к молекуле гидролитической воды.

Замена чужеродного фрагмента AlF<sub>3</sub> на PO<sub>3</sub><sup>-</sup>, т.е. реставрация молекулы GTP при моделировании, приводит к ряду изменений в геометрической конфигурации модельного белкового комплекса. Во-первых, Arg789 становится достаточно близко к мостиковым атомам кислорода O(P<sub>β</sub>) и O(P<sub>α</sub>): расстояния N—O сокращаются на 0,4–0,5 Å. Во-вторых, остатки Thr35 и Gln61 подтягиваются к кислороду гидролитической воды: расстояния O—O сокращаются на 0,5 Å.

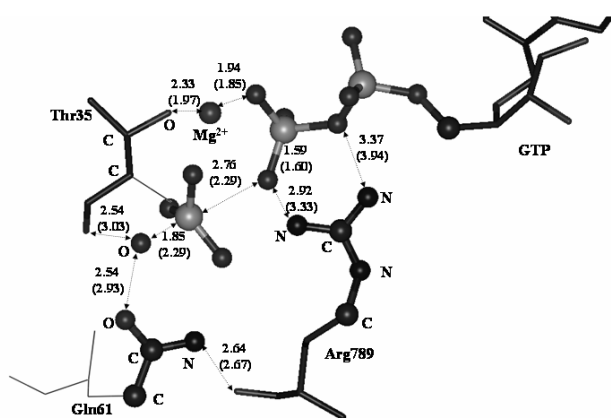


Рис. 9. Сопоставление расстояний (Å) между некоторыми «тяжелыми» атомами в равновесной геометрической конфигурации модельной системы гуанозинтрифосфат-RAS-GAP (значения без скобок) и в кристалле 1WQ1 (значения в скобках)

Наиболее важным является заключение, что расстояние P<sub>γ</sub>—O(P<sub>β</sub>) (2,76 Å) отвечает практически разорванной связи P—O и что γ-фосфатная группа PO<sub>3</sub><sup>-</sup> располагается близко от молекулы гидролитической воды. Это означает, что разделение GTP на GDP + PO<sub>3</sub><sup>-</sup> внутри белкового комплекса происходит под действием окружающих молекулярных групп белков до начала собственно реакции гидролиза, поскольку молекула гидролитической воды в рассматриваемой конфигурации остается практически в неизменном виде. Решающими взаимодействиями внутри сеток водородных связей являются:

- взаимодействие Arg789 (GAP)—Gln61 (RAS), «подтягивающее» Gln61 к реакционному центру,
- взаимодействия Thr35 и Gln61 с молекулой воды, ориентирующие ее соответствующим образом по отношению к γ-фосфатной группе,
- взаимодействие подходящим образом ориентированной молекулы воды с GTP, что приводит к стереохимическому выворачиванию «зонтика» PO<sub>3</sub>.

Для проверки высказанных утверждений был проведен численный эксперимент с заменой боковой цепи Arg на остаток лейцина Leu («мутация» в GAP). Оптимизация геометрических параметров в подобной модифицированной системе не приводит к разъединению GTP на GDP и PO<sub>3</sub><sup>-</sup>: равновесное расстояние P<sub>γ</sub>—O(P<sub>β</sub>) составляет 1,93 Å. Эти результаты были получены по расчетам методом КМ/ММ. Рассчитанный рельеф поверхности потенциальной энергии показывает, что процесс разъединения GTP происходит без энергетического барьера. Дополнительные расчеты в рамках полностью квантового подхода с использованием приближений Хартри—Фока и теории функционала электронной плотности для 150-атомного кластера, включающего атомы реакционного центра, показали, что для разрыва связи P<sub>γ</sub>—O(P<sub>β</sub>) может потребоваться небольшой активационный барьер [18].

Рис. 10 иллюстрирует механизм второй, заключительной стадии реакции гидролиза гуанозинтрифосфата белковым комплексом RAS-GAP. Здесь представлена та же конфигурация, что и на рис. 9, но с изображением некоторых атомов водорода. Дальнейший ход химических превращений может проходить по двум

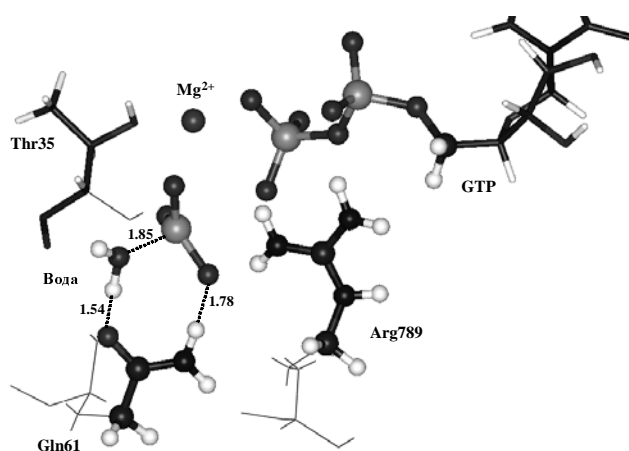
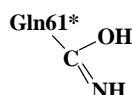


Рис. 10. Схема, иллюстрирующая механизм второй стадии реакции гидролиза гуанозинтрифосфата белковым комплексом RAS-GAP

возможным каналам. Можно принять, что реакция будет протекать по схеме субстрат-ассистирующего катализа [17], согласно которой протон молекулы воды присоединяется к кислороду  $\gamma$ -фосфатной группы, а гидроксил — к атому фосфора  $P_\gamma$ . Для подобного механизма был выполнен расчет полного энергетического профиля пути реакции методом КМ/ММ и получены структуры реагентов ( $GDP \dots PO_3^- + H_2O$ ) и продуктов ( $GDP + H_2PO_4^-$ ). Однако этот путь реакции невозможен вследствие нереально высокого активационного барьера (более 30 ккал/моль).

Можно предложить и другой путь реакции. На рис. 10 видно, что молекула воды ориентирована водородными связями с карбонильными группами остатков Gln61 и Thr35 в позиции, удобной для перегруппировки внутри 8-членного цикла. А именно, при двойном переносе протонов вдоль направлений  $(Gln)N-H \dots O$  (длина водородной связи 1,78 Å) и  $(Gln)O \dots H-O$  (длина водородной связи 1,54 Å) и присоединении образующегося гидроксила от воды к  $P_\gamma$  может образоваться неорганический фосфат  $H_2PO_4^-$ . Расчеты соответствующего энергетического профиля методом КМ/ММ дают оценку активационного барьера на этом пути менее 10 ккал/моль, что вполне разумно для биохимических процессов.

При подобной перегруппировке глутамин остается в таутомерной форме



что не может вызывать серьезных возражений — исходная форма достаточно легко может быть восстановлена, если вблизи окажется молекула воды после освобождения продуктов.

В литературе удалось найти упоминание о возможности таутомерного механизма реакции гидролиза связей в G-белках, высказанное как гипотеза [21]. Но эта гипотеза позднее была отвергнута в теоретической работе [22] на основании полуколичественных оценок энергетических затрат. Действительно, в изолированном состоянии энергия таутомерной формы Gln61\* должна быть выше (примерно на 18 ккал/моль). Однако в данном молекулярном окружении необходимо принимать во внимание наличие рядом с остатком глутамин аниона  $PO_3^-$ , что существенно меняет энергетику перегруппировки.

Таким образом, результаты моделирования дают основания утверждать, что реакция гидролиза гуанозинтрифосфата в белковом комплексе RAS-GAP происходит по схеме, представленной на рис. 11.

Роль ключевых остатков Gln61(RAS) и Arg789(GAP) полностью ясна. Понятно, что мутации по этим позициям должны приводить к возрастанию активационного барьера реакции или полностью менять механизм катализа.

#### Вычислительные аспекты

Компьютерная программа, реализующая данный вариант метода КМ/ММ, т.е. позволяющая рассчитывать энергии и силы в больших биомолекулярных системах, представляет собой комбинацию пакета программ квантовой химии GAMESS [23] (точнее ее версии PC GAMESS, ориентированной на компьюте-

ры архитектуры Intel, созданной сотрудником лаборатории химической кибернетики А.А. Грановским, URL <http://lcc.chem.msu.ru/gran/gamess/index.html>) и пакета программ молекулярного моделирования TINKER (J. Ponder, URL <http://dasher.wustl.edu/tinker>).

Первый вариант программы трудно назвать «дружественным для пользователя» — каждый раз необходимо вводить достаточно большой объем начальных данных, включая обычные параметры квантовохимических задач (информацию о базисных функциях, параметры, управляющие сходимостью численных методов решения соответствующих матричных уравнений, параметры метода учета эффектов электронной корреляции и т.д.), информацию об эффективных фрагментах, информацию о молекулярно-механических силовых полях. Мы постоянно совершенствуем методики, программы и интерфейсы так, чтобы применение метода КМ/ММ к конкретным задачам стало более доступным для широкого круга пользователей.

Наиболее трудоемкая часть работы связана с оптимизацией геометрических параметров молекулярной системы, т.е. с нахождением для каждого значения координат реакции координат всех частиц по минимуму полной энергии, включающей и квантовую, и молекулярно-механическую составляющие. Сове-

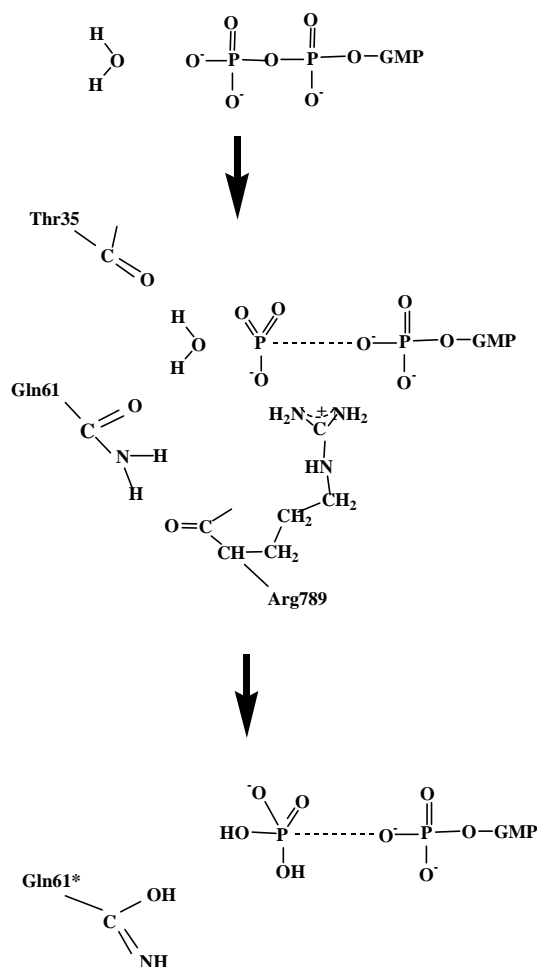


Рис. 11. Схема, иллюстрирующая механизм реакции гидролиза гуанозинтрифосфата комплексом RAS-GAP

менные алгоритмы оптимизации требуют расчетов энергии и градиентов энергии для систем с числом степеней свободы порядка тысяч. Каждый такой расчет энергии и градиента проводится с характерным числом базисных функций, составляющим порядка нескольких сотен. Понятно, что при решении подобных задач необходимо ориентироваться на использование принципов параллельных вычислений, когда одна задача обрабатывается совокупностью компьютерных процессоров. Как правило, применяются вычислительные кластеры, состоящие из комплексов машин с числом узлов от 4 до 32. Весьма перспективна одна из современных тенденций в развитии компьютерных технологий на основе идеи распределенных вычислений, т.е. использование для решения одной задачи пространственно разделенных, но связанных электронными сетями вычислительных кластеров, например кластера на Химическом факультете МГУ и кластера в Институте биохимической физики РАН. Применение подобных вычислительных приемов требует создания специальных программных средств, что также составляет предмет наших усилий [24].

В свою очередь достижения компьютерных технологий позволяют ставить более амбициозные задачи для молекулярного моделирования. Поскольку быстрые вычисления градиентов энергии (сил) для сложных систем, состоящих из квантовых и молекулярно-механических частей, становятся практически возможными, то реальными становятся задачи расчетов молекулярно-динамических траекторий с квантовыми силами в КМ-подсистеме. Реализация подобной методики позволяет преодолеть ряд ограничений стратегии КМ/ММ: при моделировании можно ввести важнейший параметр — температуру, в частности, перейти к расчетам энергии Гиббса; можно добиться значительно более эффективного сканирования обширных участков потенциальной поверхности, а не ограничиваться только областями локальных минимумов.

Нами предприняты первые шаги в направлении создания комбинированного метода квантовой механики и молекулярной динамики (КМ/МД). В частности, разработана программа, реализующая метод молекулярной динамики применительно к твердым телам, поскольку программа КМ/ММ позволяет вычислять силы, действующие на эффективные фрагменты как на твердые тела. Для обеспечения параллельных вычислений применяется система автоматического динамического распараллеливания OpenTS [24]. Первые приложения метода КМ/МД относились к расчетам динамики шара (10 Å) из молекул воды (154 молекулы), а также динамики состоящего из 36 аминокислотных остатков белка виллина, который часто используется для моделирования сворачивания пептидных цепей (URL <http://folding.stanford.edu/villin/>).

### Заключение

Итак, новый комбинированный метод квантовой и молекулярной механики (КМ/ММ) позволяет моделировать свойства и химические реакции в любых белковых системах. Наибольший интерес представляет построение энергетических профилей ферментативных реакций, что дает возможность исследовать меха-

низмы реакций ферментативного катализа на уровне элементарных стадий. Что касается структурных задач, которые технически более простых, но не менее важных, с помощью методики КМ/ММ можно, в частности, достаточно точно определить положения атомов водорода в белках, исследованных рентгеноструктурными методами. В качестве перспективного направления представляется создание программного обеспечения для смешанных квантово-классических молекулярно-динамических расчетов.

При написании статьи были использованы работы, поддержанные Российским фондом фундаментальных исследований (проект № 05-07-90146) и Российской академией наук (программа 10 Отделения химии и наук о материалах).

### ЛИТЕРАТУРА

1. Warshel A., Levitt M. J. Mol. Biol., 1976, v. 103, p. 227.
2. Gordon M.S., Freitag M.A., Bandyopadhyay P. e. a. J. Phys. Chem. A., 2001, v. 105, p. 293.
3. Wladkowski B.D., Krauss M., Stevens W.J. J. Amer. Chem. Soc., 1995, v. 117, p. 10537.
4. Kairys V., Jensen J.H. J. Phys. Chem. A., 2000, v. 104, p. 6656.
5. Nemukhin A.V., Grigorenko B.L., Bochenkova A.V. e. a. J. Molec. Struct. (THEOCHEM), 2002, v. 581, p. 167.
6. Grigorenko B.L., Nemukhin A.V., Topol I.A., Burt S.K. J. Phys. Chem. A, 2002, v. 106, p. 10663.
7. Nemukhin A.V., Grigorenko B.L., Topol I.A., Burt S.K. J. Comp. Chem., 2003, v. 24, p. 1410.
8. Nemukhin A.V., Grigorenko B.L., Bochenkova A.V. e. a. Struct. Chem., 2004, v. 15, p. 3.
9. Nemukhin A.V., Grigorenko B.L., Topol I.A., Burt S.K. Phys. Chem. Chem. Phys., 2003, v. 6, p. 1031.
10. Немухин А.В., Григоренко Б.Л., Князева М.А. Ж. физ. химии, 2004, т. 78, с. 1028.
11. Hedstrom L. Chem. Rev., 2002, v. 102, p. 4501.
12. Nemukhin A.V., Grigorenko B.L., Rogov A.V. e. a. Theor. Chem. Acc., 2004, v. 111, p. 36.
13. Berman H.M., Westbrook J., Feng Z. e. a. Nucl. Acid. Res., 2000, v. 28, p. 235.
14. Cornell W.D., Cieplak P., Bayly C.I. e. a. J. Amer. Chem. Soc., 1995, v. 117, p. 5179.
15. Nemukhin A.V., Topol I.A., Burt S.K. Int. J. Quant. Chem., 2002, v. 88, p. 34.
16. Nemukhin A.V., Topol I.A., Grigorenko B.L., Burt S.K. J. Phys. Chem. B., 2002, v. 107, p. 2958.
17. Kosloff M., Selinger Z. Trends Biochem. Sci., 2001, v. 26, p. 161.
18. Topol I.A., Cachau R.E., Nemukhin A.V. e. a. Biochim. Biophys. Acta, 2004, v. 1700, p. 125.
19. Grigorenko B.L., Nemukhin A.V., Topol I.A. e. a. Proteins: Stuct. Func. Bioinf., 2005, v. 60, p. 495.
20. Scheffzek K., Ahmadian M.R., Kabsch W. e. a. Science, 1997, v. 277, p. 333.
21. Sondek J., Lambright D.G., Noel J.P. e. a. Nature, 1994, v. 372, p. 276.
22. Schweins T., Warshel A. Biochem., 1996, v. 35, p. 14232.
23. Schmidt M.W., Baldrige K.K., Boatz J.A. e. a. J. Comput. Chem., 1993, v. 14, p. 1347.
24. Abramov S.M., Adamovich A.I., Nemukhin A.V., Moskovsky A.A. e. a. Proceedings Conf. «Supercomputing Systems and Applications», 2004, Minsk, Belarus.