

Суперкомпьютерное моделирование структуры и спектров фитохромов*

И.В. Поляков^{1,2}, В.А. Миронов^{1,2}, А.А. Московский^{1,2}, Б.Л. Григоренко^{1,2},
А.В. Немухин^{1,2}

Химический факультет МГУ имени М.В. Ломоносова¹, Институт биохимической физики им. М.Н. Эмануэля РАН²

В настоящий момент активно развивается использование генетически кодируемых маркеров для исследований *in vivo*. При рациональном дизайне новых мутантов важным критерием является поглощение и испускание таких биомаркеров в дальнекрасной и ближней ИК области (район 700-900нм) с высоким квантовым выходом (яркостью), что довольно проблематично обеспечить в случае использования мутантов из семейства зеленого флуоресцентного белка. В настоящий момент рассматривается использование модифицированных фитохромов[1], поскольку они включают в себя билиновые хромофоры, которые имеют гораздо более протяженную сопряженную систему связей по сравнению с хромофорами зеленых и красных флуоресцентных белков (рис. 1).

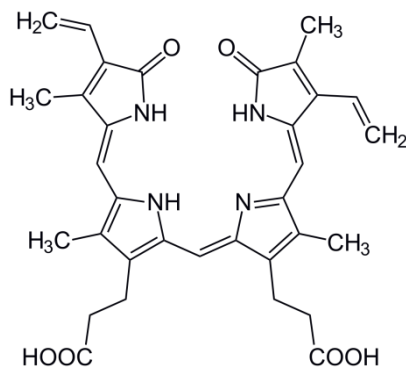


Рис. 1. Структурная формула биливердина

Большой проблемой является то, что протонирование таких хромофоров в растворе и белковой матрице может быть различным (у нейтральной формы протонированы 3 азота- рис. 1, у катионной все четыре), что может значительно сказываться как на спектре поглощения, так и на яркости сигнала.

Временно-зависимый метод теории функционала плотности (TDDFT) является наиболее широко используемым методом расчета энергий вертикальных переходов больших органических хромофоров [2], который при этом хорошо распараллеливается и масштабируется на современных вычислительных системах. Однако, ошибки TDDFT зачастую очень высоки для билиновых систем [3], его использование даже для расчетного качественного скрининга мутаций находится под большим вопросом.

Проведенное нами тестирование молекул билиновых хромофоров в газовой фазе показало, что ошибки составляют 0.1-0.5 эВ для нейтральной модели и 0.2-0.8 эВ для катиона. Большой проблемой является то, что гибридные функционалы (B3LYP) и meta-GGA функционалы (M06L) дают хорошее описание для нейтральной формы (ошибка менее 0.1 эВ), но для катиона они завышают энергию перехода на 0.4 эВ. При этом GGA функционалы (BLYP) завышают энергию перехода для нейтральной формы на 0.5 эВ, но для катиона подходят лучше (ошибка менее 0.2 эВ). По сути, невозможно сделать однозначный выбор в пользу какого-то конкретного функционала плотности (или даже группы), что затрудняет использование данного метода для скрининга новых перспективных мутантов. По нашему опыту наиболее надежным подходом к расчету спектров органических хромофоров является метод расширенной квазивывож-

* Работа выполнена в рамках проекта РНФ 17-13-01051.

денной многоконfigurационной теории возмущений (ХМCQDPT) [4]. В частности, для катиона модифицированного биливердина 15Za [3] в водном растворе он дал 1.85 эВ по результатам данной работы (эксперимент 1.83 эВ), тогда как из литературы известно, что родственный метод CASPT2 воспроизводит энергию гораздо хуже – 1.95 эВ [3].

При рассмотрении модели хромофора в белковой матрице ошибка метода TDDFT ведет себя гораздо лучше и качественно воспроизводит сдвиг поглощения при изменении протонированного состояния хромофора. Например, LC-BLYP дает 549 и 532 нм для катионной и нейтральной форм хромофора соответственно, тогда как ХМCQDPT- 668 и 660 нм. Эти результаты показывают, что лишь на основании расчета энергий вертикальных переходов невозможно определить какая форма хромофора присутствует в том или ином белке, ведь разница в значениях (0.02 эВ) меньше, чем ошибка метода относительно экспериментального результата (0.1 эВ). Однако, КМ/ММ расчеты все же дают однозначный результат, так как катионная форма в рассматриваемом нами мутанте оказалась однозначно стабильнее, чем нейтральная – на 3.3 ккал/моль, то есть в 252 раза при температуре 300К,

С вычислительной точки зрения наши задачи делятся на 3 части - молекулярная динамика для подготовки трехмерной полноатомной модели, квантовые или комбинированные КМ/ММ расчеты геометрических конфигураций, расчет энергий вертикальных переходов (спектров). Молекулярная динамика действительно хорошо масштабируется и быстро считается при использовании узлов с ускорителями NVIDIA, так как код популярных пакетов NAMD и GROMACS был своевременно переписан и очень хорошо оптимизирован их разработчиками. Квантовые или комбинированные расчеты (КМ/ММ) в настоящий момент основаны на очень старом и неэффективном коде, поэтому практически не задействуют GPU ускорители в рамках доступных прикладных пакетов, хотя и способны хорошо масштабироваться на десятки и сотни узлов, тысячи обычных процессорных ядер. Расчеты спектров в основном используют лишь ресурсы одного мощного SMP узла с большой памятью и быстрым хранилищем данных. Отдельной проблемой является очень ограниченная поддержка векторизации в прикладных пакетах квантовой химии, тогда как современные x86-64 процессоры имеют поддержку AVX2 и даже AVX512.

Работа выполнена на суперкомпьютерах «Ломоносов», «Ломоносов-2» НИВЦ МГУ и «МВС-10» МСЦ РАН. Использовались программные пакеты NWChem, GROMACS, GAMESS (US) и Firefly (PC GAMESS).

Литература

1. Chernov K., Redchuk T., Omelina E., Verkhusha V. Near-Infrared Fluorescent Proteins, Biosensors, and Optogenetic tools Engineered from Phytochromes // *Chem. Rev.* 2017. 117, 9. P. 6423-6446. DOI: 10.1021/acs.chemrev.6b00700
2. Jacquemin D., Wathelet V., Perpète E., Adamo C. Extensive TD-DFT Benchmark: Singlet-Excited States of Organic Molecules // *J. Chem. Theory Comput.* 2009. 5, 9. P. 2420-2435. DOI: 10.1021/ct900298e
3. Falkolf O., Durbeej Bo. Red-light absorption and fluorescence of phytochrome chromophores: A comparative theoretical study // *Chem. Phys.* 213. 425. P. 19-28. DOI: 10.1016/j.chemphys.2013.07.018
4. Granovsky A. Extended multi-configuration quasi-degenerate perturbation theory: the new approach to multi-state multi-reference perturbation theory// *J. Chem. Phys.* 2011. 134, 21. P. 214113. DOI: 10.1063/1.3596699