

Ю.А. Мазур<sup>1</sup>, Н.Ю. Опарина<sup>2</sup>

<sup>1</sup> Московский физико-технический институт (государственный университет)

<sup>2</sup> Институт молекулярной биологии им. Энгельгардта РАН

### **Молекулярное моделирование распознавания стоп-кодона фактором терминации трансляции 1-го типа eRF1**

Терминация трансляции – завершающий этап синтеза белка, при котором происходит гидролиз пептидил-тРНК и высвобождается новосинтезированный полипептид. В этом процессе принимаются участие белковые факторы, функциональные аналоги тРНК, специфические связывающиеся со стоп-кодонами. У эукариот все три стоп-кодона (UAA, UAG или UGA) в мРНК распознаются фактором eRF1 (eukaryotic peptide chain release factor 1). eRF1 состоит из трех доменов, причем N-концевой домен отвечает за специфическое взаимодействие со стоп-кодоном. Многочисленные исследования показали влияние на специфичность работы eRF1 точечных мутаций в различных областях N-концевого домена, однако ни сайт взаимодействия со стоп-кодоном, ни механизм декодирования до сих пор остаются невыясненными [1,2]. В нашей работе применены методы молекулярного моделирования для поиска участка взаимодействия 2-го и 3-го нуклеотидов стоп-кодона с доменом 1 eRF1 (N-домен белка eRF1).

Работа разделена на две части: во-первых, это первичный (грубый) поиск места связывания N-домена и лиганда с помощью программы для молекулярного докинга Autodock (версия 3) [3,4], а, во-вторых, более точный расчет энергий взаимодействия с помощью программы молекулярной динамики Amber [5]. На первом этапе происходит отбор потенциальных сайтов взаимодействия N-концевого домена eRF1 с лигандом, в качестве которого выбран РНК-дипурин (AA, AG, GA или GG). Такой лиганд был выбран на основании двух предпосылок: во-первых, стоп-кодона и триптофановый кодон представляют собой мотив U-пурин-пурин, при этом показано, что мотив UGG не может работать в качестве стоп-кодона даже в модельной системе в отсутствие триптофановой тРНК. Поэтому комплексы GG с N-концевым доменом eRF1 не должны быть сопоставимы по стабильности с комплексами динуклеотидов AA, GA и GA. Во-вторых, известные экспериментальные сведения дают основание сомневаться в одновременном распознавании первого нуклеотида U и остальных нуклеотидов стоп-

кодона, позволяя предположить двухэтапный процесс. Эти данные позволяют считать допустимой применение такой минимальной модельной системы для моделирования.

При молекулярном докинге было получено большое количество моделей комплексов, неразличимых по энергии взаимодействия (определенной по скоринг-функции Autodock). Поэтому для дальнейшей работы была написана программа, производящая отбор «потенциально специфичных» структур с помощью фильтров. Фильтры учитывали энергию комплексов после докинга, распределение водородных связей между сахаро-фосфатным остовом и нуклеотидной частью лиганда и пр. Следует отметить, что отбор структур только по этим признакам также не позволил получить характерных моделей взаимодействия N-концевого домена eRF1 с фрагментом стоп-кодона. Мы предположили, что к важным элементам распознавания стоп-кодона при терминации трансляции следует отнести окружение рибосомной РНК в А-сайте рибосомы, поскольку фактор eRF1 неспособен специфически связывать стоп-кодон в растворе. Кодон, предшествующий стоп-кодону в процессе терминации находится в Р-сайте рибосомы в связанном с тРНК состоянии. В связи с этими данными можно предполагать, что все стоп-кодоны должны взаимодействовать с одним и тем же локальным районом N-концевого домена eRF1. Отобранные комплексы, прошедшие фильтр, были сгруппированы в структурные кластеры, в зависимости от распределения в них водородных связей между белком и мРНК. Из каждого кластера были отобраны лучшие по энергиям структуры, которые далее стали начальными данными для молекулярной динамики. Поскольку нашей задачей являлся поиск высокоспецифичного сайта, взаимодействующего со стоп-кодоном, дополнительно была проведена проверка комплексов белка с AG и GA «мутациями», а именно: в каждом их исходных комплексов был заменен нуклеотид А на G и, т.о., получен неспецифический комплекс (поскольку UGG не связывается с eRF1). Была проведена молекулярная динамика с помощью SANDER (пакет программ Amber) при непрямом указании растворителя. При тех же условиях провели молекулярную динамику N-концевого домена eRF1 в отсутствии лигандов и свободных динуклеотидов AA, AG, GA и GG. На основании сравнения потенциальных энергий комплекса белок-лиганд и пары свободных белка и соответствующего лиганда в равновесной стадии молекулярной динамики была определена энергия взаимодействия. Среди комплексов белка с AG и GA были отобраны те варианты, для которых замена А на G (с получением в результате комплекса с GG) приводила к значительной дестабилизации комплекса. Для AA с самого начала был отобран только один стабильных вариант

комплекса, прошедший фильтр, поэтому дополнительный отбор не проводился. В результате для каждого из «специфических» динуклеотидов AA, GA и AG, представляющих все три известных стоп-кодона (UAA, UAG и UGA) получена одна рабочая модель взаимодействия с N-концевым доменом eRF1. При этом во всех случаях наблюдали пространственно сходное расположение лиганда на белке. При молекулярном докинге GG подобного расположения лиганда получить не удалось, что можно считать дополнительным свидетельством того, что полученная нами модель взаимодействия близка к реальной и специфична.

Полученные модели комплексов были исследованы, проанализированы аминокислотные остатки в eRF1, взаимодействующие с лигандами и изучен тип взаимодействия. Для большинства аминокислот, идентифицированных в данной работе как контактные, экспериментально показано сильное влияние на функции белка, а также показана их высокая консервативность для всех эукариот, узнающих три стоп-кодона[6,7].

#### Литература

1. *Buckingham R.H., Grentzmann G., Kisselev L.* Polypeptide Chain Release Factors // *Mol. Microbiol.*, Vol. 24, P. 449-456, 1997.
2. *Caskey C.T.*, Peptide chain termination. *Trends Biochem. Sci.* **5**:234-237, 1980.
3. *Moitessier N, Westhof E, Hanessian S.* Docking of aminoglycosides to hydrated and flexible RNA. *J Med Chem.* 2006 Feb 9;49(3):1023-33
4. *Morris, G. M., Goodsell, D. S., Halliday, R.S., Huey, R., Hart, W. E., Belew, R. K. and Olson, A. J.* (1998), Automated Docking Using a Lamarckian Genetic Algorithm and Empirical Binding Free Energy Function *J. Computational Chemistry*, **19**: 1639-1662.
5. *D.A. Case, T.E. Cheatham, III, T. Darden, H. Gohlke, R. Luo, K.M. Merz, Jr., A. Onufriev, C. Simmerling, B. Wang and R. Woods.* The Amber biomolecular simulation programs. *J. Computat. Chem.* **26**, 1668-1688 (2005).
6. *Bertram G, Bell HA, Ritchie DW, Fullerton G, Stansfield I.* Terminating eukaryote translation: domain 1 of release factor eRF1 functions in stop codon recognition. *RNA.* 2000 Sep;6(9):1236-47.
7. *Kim OT, Yura K, Go N, Harumoto T.* Newly sequenced eRF1s from ciliates: the diversity of stop codon usage and the molecular surfaces that are important for stop codon interactions. *Gene.* 2005 Feb 14;346:277-86. Epub 2005 Jan 4. Erratum in: *Gene.* 2005 Jun 6;352:137.