

## ГОМОЛОГИЧНОЕ МОДЕЛИРОВАНИЕ ПРОСТРАНСТВЕННОЙ СТРУКТУРЫ GFP-ПОДОБНЫХ БЕЛКОВ

Гарковенко А. В., Пахомов А. А.

*Федеральное государственное бюджетное учреждение науки Институт биоорганической химии им. академиков М. М. Шемякина и Ю. А. Овчинникова Российской академии наук, Москва, Россия (117997, Москва, ул. Миклухо-Маклая, 16/10), e-mail: a\_garkov@mx.ibch.ru*

Основным методом определения пространственной структуры флуоресцентных белков GFP семейства является метод кристаллографии. Как альтернатива рентгеноструктурному анализу ранее был предложен метод установления структуры флуоресцентных белков, объединяющий гомологичное моделирование *in silico* для определения общей пространственной структуры белка и эксперименты *in vitro* для установления структуры хромофора, который в основном и задает спектральные свойства белка. Однако в некоторых случаях для исследований значительную ценность представляет информация об аминокислотном окружении хромофора, тогда как данные о структуре хромофора не принципиальны. В настоящей работе представлены алгоритмы моделирования пространственной структуры флуоресцентных белков с использованием программы MODELLER, которая строит модель на основе пространственных ограничений, рассчитанных с учётом выравнивания последовательностей анализируемого белка и структуры-шаблона.

Ключевые слова: гомологичное моделирование, структура белка, гетероатомы.

## HOMOLOGY MODELING OF THE SPATIAL STRUCTURE OF GFP-LIKE PROTEINS

Garkovenko A. V., Pakhomov A. A.

*Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry of the Russian Academy of Science, Moscow, Russia (117997, Moscow, Miklukho-Maklaya 16/10), e-mail: a\_garkov@mx.ibch.ru*

Crystallography is the main method for determination of the spatial structure of proteins. As an alternative, a method for the determination of fluorescent proteins spatial structure has been proposed. It combines homology modeling *in silico* for determination of overall protein structure and experiments *in vitro* for determination of the chromophore structure, which mainly defines the spectral properties of the protein. However, in some cases it is interesting to know amino acid environment of the chromophore, irrelative to the chromophore structure. In the present work algorithms for the fluorescent protein spatial structure modeling using MODELLER program are presented. This program builds the model by satisfaction of spatial restraints and accounts for sequence alignment of analyzed and template proteins.

Key words: homology modeling, protein structure, heteroatoms.

### Введение

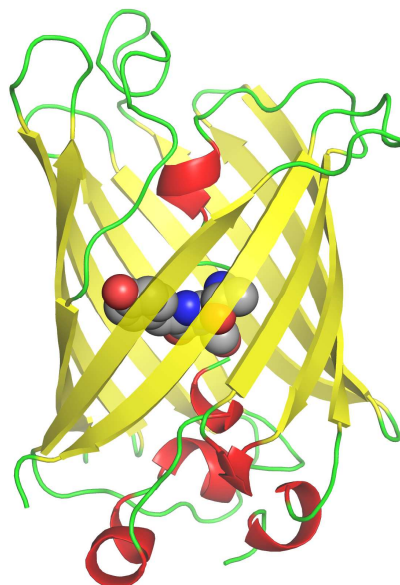
Компьютерный анализ широко применяется в научно-исследовательских работах. В частности, достижения информатики, вычислительной техники, прикладной математики и статистики используются для решения проблем, поставленных биологией, в рамках раздела естествознания биоинформатики. Одним из наиболее важных направлений современной биоинформатики является разработка алгоритмов и программ для предсказания пространственной структуры белков. Гомологичное моделирование в настоящее время наиболее успешно и часто используется для решения этой задачи [3].

В настоящей работе предполагается использовать гомологичное моделирование для предсказания структуры флуоресцентных белков. Эти белки в последнее время стали широко использоваться при исследовании живых систем. В 2008 г. была присуждена Нобелевская премия по химии «За открытие и разработку зелёного флуоресцентного белка GFP». Такая

популярность GFP (Green Fluorescent Protein, зеленый флуоресцентный белок) и родственных ему флуоресцентных белков возникла благодаря их способности к синтезу хромофора из собственных аминокислот без участия внешних кофакторов. Таким образом, цвет и физико-химические свойства определяются геном белка, что позволяет вводить флуоресцентные метки на генетическом уровне. В настоящее время разработано значительное количество генетически-кодируемых флуоресцентных зондов, биосенсоров и меток на основе флуоресцентных белков. При их разработке знание о структуре белка является одним из основополагающих.

### **Общая пространственная структура флуоресцентных белков**

Все известные флуоресцентные белки с установленной кристаллической структурой имеют упаковку  $\beta$ -бочки (рис. 1), 11  $\beta$ -листов образуют полый цилиндр с диаметром около 30 Å и длиной около 40 Å, внутри которого проходит сильно деформированная  $\alpha$ -спираль, несущая хромофор. Небольшие  $\alpha$ -спиральные участки находятся также на торцах цилиндра [5]. Хромофор, который и задает спектральные свойства белка, находится в середине и синтезируется автокаталитически из трех аминокислотных остатков самого белка (в GFP это остатки -Ser65-Tyr66-Gly67-).



*Рис. 1. Пространственная структура GFP (PDB ID: 1EMB). Хромофор внутри белка представлен в виде шаровой модели*

Гомологичное моделирование белков повсеместно применяется для предсказания их структуры. Однако оно не применимо в исходном виде для флуоресцентных белков, содержащих в своей структуре хромофор. Структура этих белков в настоящее время определяется экспериментально при помощи рентгеноструктурного анализа. В то же время

недавно было показано, что гомологичное моделирование в комбинации с масс-спектрометрическим анализом хромофорсодержащих пептидов флуоресцентных белков может быть использовано для предсказания их структуры [2]. Однако в некоторых случаях перед исследователем встает задача в оценке аминокислотного окружения хромофора, тогда как структура хромофора принципиального значения не имеет. Более того, структура хромофора относительно строго определяется цветовой группой, к которой он принадлежит. К примеру, все известные зелёные флуоресцентные белки содержат одинаковые хромофоры. Это же относится и к большинству красных белков. Всего известно менее 10 хромофоров, каждый из которых входит в определённую цветовую группу, причем часть из них может образовываться только из строго заданных хромофоробразующих аминокислотных последовательностей. Так, хромофор Kaede-типа образуется только из -His-Tyr-Gly-аминокислотной последовательности [1]. Некоторая информация о структуре хромофора может быть получена из спектральных свойств белка при денатурации в кислых и щелочных условиях [4]. Таким образом, данные о структуре хромофора могут быть с высокой вероятностью предсказаны на основании их спектральных свойств и хромофоробразующей последовательности аминокислот.

### Применение программы Modeller для флуоресцентных белков

Для моделирования структуры флуоресцентных белков наиболее рациональным представляется использование программы Modeller, фактически, являющейся стандартом для гомологичного моделирования белков. Модель строится на основе аминокислотной последовательности белка и пространственной структуры другого, гомологичного белка, причём сходство последовательностей обычно должно быть не менее 30 %. Как правило, гомология флуоресцентных белков довольно высока и составляет 30–80 %, что и позволяет применять к ним Modeller (рис. 2).

	10	20	30	40	50	60			
GFP	SKGEELFTGVVPI	LVELDGDVNGHK	FVSVSGEGEGD	ATYGKLT	TLKFICTTGKLP	VVPWPTLVTTTF.VQCF			
phiYFP	SSGALLFHGKIP	YVVEMEGNVDG	HTF SIRGKGYGD	ASVGKVDAQ	FICTTGDPVP	WSTLVTTL.AQCF			
	* * * * *	* * * * *	* * * * *	* * * * *	* * * * *	* * * * *			
	70	80	90	100	110	120	130		
GFP	SRYPDHMKRHDF	FKSAMPEGYVQ	ERTIFFKDDG	NYKTRAEVK	FEGDTLVNRI	ELKGIDFKED	GNILGH		
phiYFP	AKYGP	ELK--DFYK	SCMPEGYVQ	ERTITFE	GDGVFKTRAE	VTFENG	SVYNRVKLN	GGFQFKKDG	HVLGK
	*	*	* * * * *	* * * * *	* * * * *	* * * * *	* * * * *		
	140	150	160	170	180	190	200		
GFP	KLEYNYNSHN	VYIMADKQK	NGIKVNF	KIRHNI	ED--GSVQL	ADHYQQNT	PIGDGPVLL	PDNHYLSYQS	
phiYFP	NLEFNFTPH	CLYIWGDQ	ANHGLKSA	FKIMHEIT	GSKEDFIV	ADHTQMNT	PIGGP	VHVPEYHN	ITYHV
	* * *	* * *	* * * * *	* * * * *	* * * * *	* * * * *	* * * * *	* * *	
	210	220							
GFP	ALSKDPNE	KRDHMLLE	FVTAAGI						

phiYFP      TLKDVTDHRDNMSLVETVRAVDC  
      \*\*\*\*        \*\* \* \* \* \* \*

*Рис. 2. Выравнивание аминокислотных последовательностей белков GFP и phiYFP. Совпадающие аминокислоты обозначены звездочкой, хромофоробразующая последовательность – точкой. Степень гомологии между белками 51 %*

При выравнивании аминокислотных последовательностей моделируемого белка и шаблона в ходе моделирования возможно несколько вариантов. Во-первых, можно оставить хромофоробразующую последовательность -X-Tyr-Gly- анализируемого белка немодифицированной. Очевидно, что в этом случае в смоделированной структуре конформации аминокислот в ближайшем окружении хромофора будут отличаться от таковых в нативном белке, так как геометрия -X-Tyr-Gly- трипептида существенно модифицируется при превращении в хромофор. Тем не менее данный подход может быть использован для генерации промежуточных состояний белка, которые трудно выявить при помощи кристаллографических исследований. Этот подход может быть использован при рассмотрении механизмов формирования хромофора и моделирования исходных состояний. Во втором варианте вместо хромофоробразующей триады можно использовать -Gly-Gly-Gly- последовательность. Поскольку глицин обладает высокой торсионной свободой, данная последовательность хорошо вписывается в полость хромофора. Данный подход вообще не учитывает структуру хромофора, однако позволяет оценить относительное расположение аминокислотных остатков вокруг него. В случае, если принципиально оценить взаимодействия аминокислотного окружения с хромофором, необходимо задать его структуру. Как было сказано выше, данные о структуре хромофора можно получить на основе спектральных свойств и последовательности аминокислот, из которых он синтезируется. Однако без проведения экспериментальных исследований нельзя сказать точно, насколько верна предложенная структура. С другой стороны, можно получить несколько моделей со всеми возможными для заданных параметров структурами хромофора. В данном случае при моделировании хромофор включается в последовательность белка в виде гетерогруппы на стадии выравнивания последовательностей. При этом программа переносит хромофор в созданную модель без модификаций как жёсткое тело.

В заключение в работе предложено несколько алгоритмов для получения модели пространственной структуры флуоресцентных белков. На основании полученных моделей можно оценить аминокислотное окружение хромофора, а также возможное расположение аминокислотных остатков в ходе синтеза хромофора.

## Список литературы

- 1 Mizuno H., Mal T. P., Tong K. I., Ando R., Furuta T., Ikura M., Miyawaki A. Photo-Induced Peptide Cleavage in the Green-to-Red Conversion of a Fluorescent protein // *Mol Cell*. – 2003. – V. 12. – P.1051-1058.
- 2 Pakhomov A. A., Martynov V. I. Probing the structural determinants of yellow fluorescence of a protein from *Phialidium* sp. // *Biochem Biophys Res Co*. – 2011. – V.407, No.1. – P.230-235.
- 3 Sali A., Blundell T. L. Comparative protein modelling by satisfaction of spatial restraints // *J Mol Biol*. – 1993. – Vol. 234. – P.779-815.
- 4 Tsien R. Y. The Green Fluorescent protein // *Annu Rev Biochem*. – 1988. – V.67. – P.509-544.
- 5 Yang F., Moss L. G., Phillips G. N., Jr. The molecular structure of green fluorescent protein // *Nat Biotechnol*. – 1996. – Vol. 14, N. 10. – P. 1246-1251.

*Работа выполнена при финансовой поддержке Министерства образования и науки Российской Федерации, соглашение № 8234.*

### Рецензенты:

Шуваева Т. М., д.х.н., в.н.с., Федеральное государственное бюджетное учреждение науки Институт биоорганической химии им. академиков М. М. Шемякина и Ю. А. Овчинникова Российской академии наук, г. Москва.

Шахпаронов М. И., д.х.н., рук. группы, Федеральное государственное бюджетное учреждение науки Институт биоорганической химии им. академиков М. М. Шемякина и Ю. А. Овчинникова Российской академии наук, г. Москва.