



## БЕТА-СПИРАЛЬНЫЙ ДОМЕН БАКТЕРИОФАГА T4 УПРАВЛЯЕТ УКЛАДКОЙ ФРАГМЕНТА ДЛИННЫХ ХВОСТОВЫХ ФИБРИЛЛ В СОСТАВЕ ХИМЕРНОГО БЕЛКА

© 2010 г. Р. Н. Чупров-Неточин, Н. М. Файзуллина, Н. Н. Сыкилинда, М. Н. Симакова, В. В. Месянжинов, К. А. Мирошников\*

Учреждение Российской академии наук Институт биоорганической химии им. акад. М.М. Шемякина и Ю.А. Овчинникова РАН, 117997 ГСП, Москва, В-437, ул. Миклухо-Маклая, 16/10

Поступила в редакцию 15.05.2009 г. Принята к печати 22.05.2009 г.

Ключевая стадия заражения клетки *Escherichia coli* бактериофагом T4 — связывание с поверхностью клетки-хозяина — определяется специфичностью белков длинных хвостовых фибрилл фага, в частности gp37. В природных условиях сборка и олигомеризация этого белка требует участия по крайней мере двух дополнительных белковых факторов — gp57A и gp38, что сильно затрудняет получение рекомбинантной формы gp37. Для преодоления этого препятствия была применена одна из современных белково-инженерных стратегий — создание химерного белка, имеющего в своем составе “белок-носитель”, направляющий правильную сборку целевого белка. В этом качестве был использован тримерный β-спиральный домен другого белка фага T4 — gp5. Было показано, что этот домен, представляющий собой жесткую тримерную полипептидную призму, обладает свойствами перспективного белка-носителя. К C-концу домена gp5 в единой рамке считывания был присоединен фрагмент белка gp37, имеющий в своем составе пять пентапептидных повторов Gly-X-His-X-His, определяющих связывание с рецепторами на поверхности бактериальной клетки. Полученный химерный белок образует тример с нативной конформацией gp37 и проявляет биологическую активность.

*Ключевые слова:* бактериофаг T4, хвостовые фибриллы, фолдинг белка, белковая инженерия.

### ВВЕДЕНИЕ

Бактериофаг T4, инфицирующий *Escherichia coli*, представляет собой сложную наномашину. На заключительной стадии инфекции сборка новообразованных компонентов вирусной частицы управляется рядом белковых ассистирующих факторов (шаперонов). Большинство этих белков кодируется геномом T4 [1].

Длинные хвостовые фибриллы (ДХФ) бактериофага T4 играют основную роль на ранних стадиях инфекции. Они связываются с рецепторами на поверхности бактерии и инициируют необратимое присоединение частицы фага и инъекцию ДНК в клетку [2]. Каждая из шести ДХФ представляет собой протяженный комплекс длиной около 160 нм, который состоит из гибко соединенных дистальной и проксимальной частей примерно равной длины [1]. Проксимальная часть ДХФ образована тримером белка gp34, а дистальная часть включает тримеры gp36 и gp37 [3]. Белок gp37 (190 кДа) содержит “молекулярную иглу”, в состав которой входят пять

гомологичных повторов Gly-X-His-X-His, непосредственно ответственных за контакт с рецепторами клетки [4]. Кроме того, в последовательности gp37 содержатся тандемные повторы длиной 35–40 а.о. каждый [3] и ряд более коротких “криптоповторов” [5, 6]. Подобная периодичность наблюдается и в других белках ДХФ [6].

В силу периодичности последовательности полипептида его укладка в биологически активную пространственную конформацию (фолдинг) и олигомеризация ДХФ, в частности gp37, — термодинамически достаточно сложный процесс. В процессе инфекции фолдинг направляется белками-помощниками, кодируемыми генами 57A и 38. Белок gp38 требуется для олигомеризации gp37, а gp57A несет более общую ассистирующую функцию, участвуя в сборке gp34, gp37 и белка коротких фибрилл gp12 [7, 8]. Факт участия gp38 и gp57A в сборке ДХФ экспериментально подтвержден путем амбер-мутаций в соответствующих генах, однако точные детали механизма и термодинамики процесса и участвующих в нем аминокислотных остатков до сих пор неизвестны. Таким образом, получение рекомбинантных форм белковых компонентов ДХФ и их характеристика чрезвычайно затруднены. Для преодоления “шаперонного фактора” в правильном сворачивании рекомбинантных фибриллярных белков ис-

Сокращения: ДХФ — длинные хвостовые фибриллы; КД — круговой дихроизм; gp — продукт гена (gene product); PMSF — фенилметилсульфонилфторид.

\* Автор для связи (тел./факс: (495) 336-60-22; эл. почта: kmi@ibch.ru).

пользуются различные стратегии: расширение набора методов рефолдинга белка из тел включения, подбор штаммов *E. coli* и условий экспрессии, улучшающих фолдинг, делеционный мутагенез, коэкспрессия с шаперонами, и разные белково-инженерные подходы.

Одним из современных методов преодоления образования тел включения и агрегации рекомбинантных белков, синтезируемых *E. coli*, является технология создания химерных белков (fusion technology) [9]. В основе метода лежит конструирование плазмиды, экспрессирующей в непрерывной рамке считывания ген, состоящий из двух частей. Одна часть химерного гена кодирует целевой белок, а другая – белок-носитель, который способствует укладке целевого белка, увеличивает его растворимость и повышает уровень экспрессии. В роли носителей в литературе были описаны ряд белков [10]. Основные свойства белка, используемого в качестве носителя при создании химеры, включают высокую растворимость, одинаковую с целевым белком степень олигомеризации и наличие избыточной энергии, пригодной для приведения полипептидной цепи целевого белка в необходимую упорядоченную конформацию. В составе частицы бактериофага T4 имеются два белка, имеющих домены, подходящие для роли белка-носителя. Один из них – С-концевой домен (фолдон) белка (gpwac) был детально изучен [11], и была проверена возможность его использования в белковой инженерии [12]. Другим потенциальным белком-носителем может быть тримерный β-спиральный С-концевой домен белка gp5 [13]. В данной работе мы представляем стратегию присоединения фрагмента белка ДХФ gp37 к этому домену.

## РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Для гетерологической продукции рекомбинантных белков в научных, медицинских и промышленных целях первым выбором обычно служит энтеробактерия *E. coli*. Благодаря высокому выходу продукции, невысокой стоимости и хорошо отлаженным технологиям экспрессии в системе *E. coli* было получено множество рекомбинантных белков [9, 14]. Для проявления биологической активности рекомбинантные белки должны принять свою нативную трехмерную конформацию (фолд). Фолдинг белка – спонтанный процесс, движимый разницей в энергии Гиббса между нативным и развернутым состояниями белка. Фолдинг белка может быть очень эффективным процессом, и некоторые белки могут принимать нативную конформацию в течение миллисекунд, то есть быстрее, чем происходит биосинтез на рибосоме. Обычно, впрочем, фолдинг происходит намного медленнее в связи с тем, что включает в себя медленные этапы, в частности изомеризацию пролина. В цитозоле *E. coli* рекомбинантные белки, даже бактериальной природы, нередко не могут достаточно быстро принять нужный

фолд и образуют тела включения или нерастворимые агрегаты [14].

Для предотвращения агрегации используются различные стратегии – подбор альтернативных клеточных хозяев для экспрессии, конструирование делеционных мутантов, коэкспрессию с шаперонами и инженерия химерных белков. В качестве партнера-носителя в химерном белке успешно использовались глутатион-S-трансфераза *Shistosoma japonicum*, тиоредоксин и мальтозосвязывающий белок *E. coli*, а также ряд других белков [9]. Успех применения конкретного белка-партнера не универсален для предотвращения образования тел включения других рекомбинантных белков, в особенности для олигомерных целевых белков.

Помимо полноразмерных белков, известных своей термодинамической стабильностью и легкостью фолдинга, в качестве партнера-носителя могут использоваться белковые домены, для которых в природном белке известна их роль в качестве инициатора и движителя укладки полипептидной цепи (ядра нуклеации фолдинга, или фолдоны) [15, 16]. В случае фибриллярных белков удобным белково-инженерным фолдоном служит С-концевой домен фибритина (gpwac) бактериофага T4. Фолдон представляет собой небольшую пропеллероподобную структуру, состоящую из 27 а.о., которая располагается на С-конце тримерного фибриллярного coiled-coil белка длиной ~530 Å [11]. Было показано, что для фибритина наличие фолдона является необходимым и достаточным условием для правильного фолдинга и тримеризации как *in vitro*, так и *in vivo* [17, 18].

Инициация фолдинга и тримеризации с помощью фолдона фибритина использовалась в белково-инженерных экспериментах для получения физиологически активных хвостовых фибрилл бактериофага T4 [19, 20], хвостовой иглы бактериофага P26 [21], коротких фибрилл коллагена [22, 23], гликопротеинов вируса HIV1 [24], вируса бешенства [25], белка шипа аденовируса [26, 27]. Определенно, использование фолдона фибритина в белково-инженерных экспериментах чрезвычайно перспективно [12], однако метод имеет ряд ограничений. Например, расположение фолдона на N-конце химерного белка заметно снижает его эффективность [27]. При существенных различиях в структуре целевого белка и фибритина возможны заметные искажения структуры химерного белка по сравнению с нативной [20]. Кроме того, в ряде случаев внедрение сайта протеолитического расщепления в линкер между фолдоном и целевым белком может нарушать метастабильное состояние химерного белка и привести к быстрому выпадению в осадок отщепленного целевого белка (Месянжинов, Мирошников, неопубликованные результаты). Таким образом, поиск других доменов, способных выполнять роль фолдона-носителя в белковых химерах, – достаточно актуальная задача белковой инженерии.

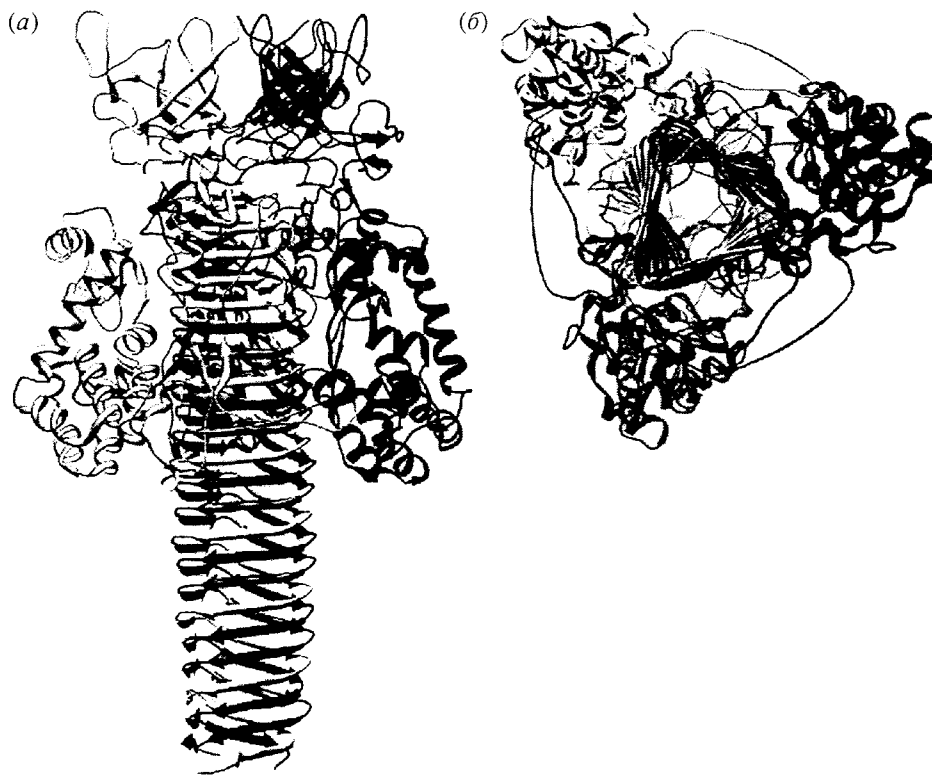


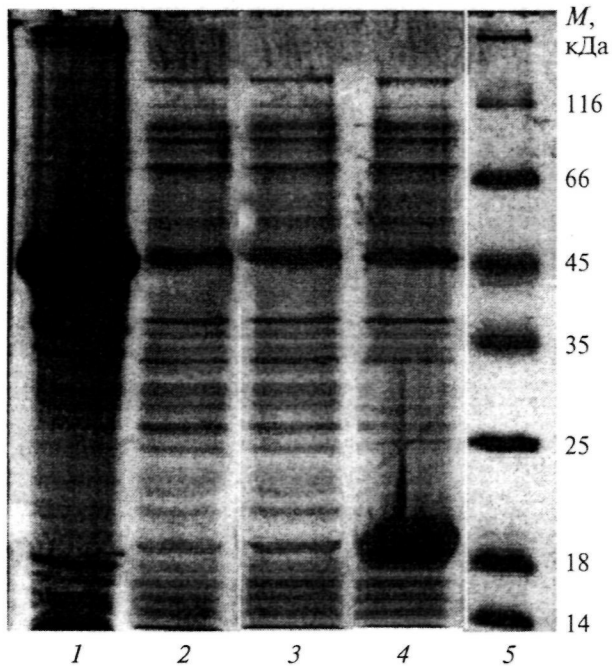
Рис. 1. Структура белка gp5 бактериофага Т4 (модифицировано из [13]).  $\alpha$ -Спирали показаны в виде спиралей,  $\beta$ -тяжи — стрелками. *a* — вид сбоку, *б* — вид сверху.

В составе базальной пластинки бактериофага Т4 имеется комплекс, участвующий в разрушении пептидогликанового слоя клеточной стенки на начальной стадии инфекции. Этот комплекс, формирующий центральную втулку базальной пластинки, состоит из белков gp5 и gp27 [28]. Структура комплекса gp5–gp27 была определена с помощью рентгеноструктурного анализа с разрешением 2.9 Å [13]. Комплекс представляет собой факелоподобную структуру высотой 190 Å, в которой тример gp27 расположен сверху. В тримере gp5 C-концевые части полипептидных цепей свернуты в жесткую спиральную иглу, верхний конец которой окружают каталитические домены, сформированные N-концевыми частями белка (рис. 1). Три цепи C-концевого домена gp5 (остатки 435–575) закручены вокруг оси третьего порядка и образуют равностороннюю призму со стороной в 110 Å. 18  $\beta$ -тяжей gp5 (по 6 в каждой цепи тримера) образуют трехзаходную  $\beta$ -спираль, в которой каждая цепь делает шесть полных оборотов. Ширина грани призмы постепенно уменьшается с 33 на N-конце до 25 Å на C-конце, что формирует структуру, подобную игле.

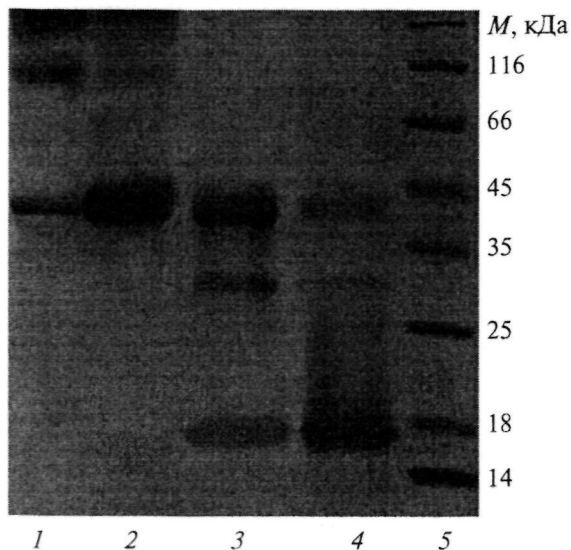
Рекомбинантная форма C-концевого домена gp5, обозначенная как 5HENS (His-tag – Enterokinase – Non-Stop, см. ниже) формирует SDS-устойчивый тример. Прочность этого домена обусловлена наличием 11 октапептидных повторов Val-X-Gly-X-X-X-X-X,

которые стабилизируют взаимное положение  $\beta$ -тяжей в призме и способствуют правильному сворачиванию как C-концевого домена, так и gp5 в целом [13, 29]. Наша гипотеза, инициировавшая настоящую работу, заключалась в том, что самособирающаяся  $\beta$ -спираль gp5 обеспечит пространственное выравнивание и тримеризацию присоединенных к ней полипептидных цепей фибриллярного белка gp37 и, таким образом, обеспечит формирование биологически активного фолда целевого белка. В качестве модельного полипептида был выбран C-концевой фрагмент gp37 (остатки 826–1026), включающий все пять гомологичных повторов Gly-X-His-X-His. Этот фрагмент был назван 37VPA по первым трем N-концевым аминокислотам последовательности. Ген, кодирующий этот полипептид, был присоединен в единой рамке считывания после гена, кодирующего  $\beta$ -спиральный участок gp5, без стоп-кодона. Конструкция 5HENS была создана таким образом, что на N-конце химерного белка синтезируется полигистидиновая метка для аффинно-хроматографической очистки.

При гетерологической экспрессии в *E. coli* химерный белок синтезировался в растворимой форме, в то время как собственно целевой белок 37VPA, клонированный под T7-промотор плазмидного вектора, образовывал нерастворимые тела включения (рис. 2). Очищенный белок gp5HENS-37VPA сохра-



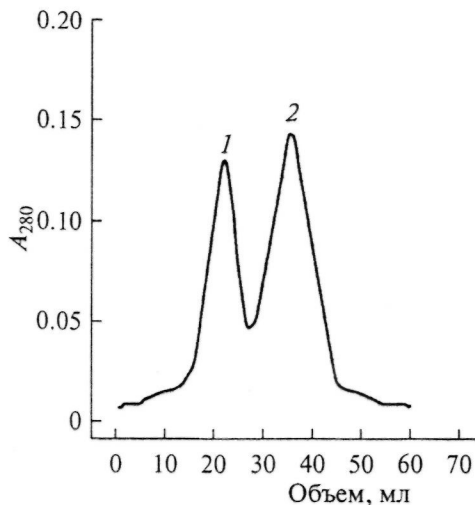
**Рис. 2.** Электрофорез в 12% SDS-ПААГ рекомбинантных белков gp5HENS-37VPA (1, 2) и gp37VPA (3, 4) после разрушения синтезирующих их клеток ультразвуком и центрифугирования: 1, 4 – растворимые фракции, 2, 3 – нерастворимые; 5 – маркеры молекулярных масс (Fermentas).



**Рис. 3.** Денатурация белков gp5HENS-37VPA (1, 2) и gp5HENS (3, 4) в присутствии SDS. Очищенные препараты в 10% SDS-ПААГ: 1, 3 – непрогретые образцы, 2, 4 – прогретые при 100°C. Дорожка 5 – маркеры молекулярных масс (Fermentas).

рактально для фибриллярных белков. С помощью инженерного подхода преодолеть этот фактор, препятствующий кристаллизации белка и определению его структуры с помощью рентгеновского анализа, не удалось.

Данные КД-спектроскопии химерного белка показывают долю  $\beta$ -тяжей, равную 53% и содержание  $\alpha$ -спирали 24% (рис. 5). Такое перераспределение элементов вторичной структуры по сравнению с практически полностью  $\beta$ -структурированным белком-носителем gp5 соответствует рассчитанному



**Рис. 4.** Разделение олигомерных форм gp5HENS-37VPA гель-фильтрацией на Toyopearl HW-55. Пик 1 – агрегированные формы, пик 2 – тримеры белка.

нял стабильность в растворе в течение нескольких недель при 4°C.

Ряд фибриллярных белков вирусного происхождения обладают характерным свойством устойчивости к SDS при электрофорезе, т.е. без нагревания компактная структура белка не полностью подвергается денатурации и белок мигрирует в геле с подвижностью, примерно соответствующей олигомерным формам [17, 18].  $\beta$ -Спиральный домен-носитель химерного белка обладает таким свойством, оно передается и всему химерному белку (рис. 3). При нанесении на калиброванную гель-фильтрационную колонку Toyopearl HW-55 очищенный препарат химерного белка gp5HENS-37VPA элюируется в виде симметричного пика с объемом элюции, соответствующим расчетной массе ~115 кДа, что подтверждает компактную тримерную форму белка в растворе (рис. 4, пик 2). Кроме того, наблюдается пик, соответствующий агрегированным формам белка, составляющим значительную долю. Следует отметить, что при увеличении концентрации gp5HENS-37VPA в растворе для проведения экспериментов по гель-фильтрации или кристаллизационному скринингу выше 3–4 мг/мл, а также при длительном хранении белок склонен агрегировать. Агрегация при повышении концентрации наблюдалась также при исследовании целых ДХФ [3] и некоторых белков, формирующих их [19, 30], то есть это свойство ха-

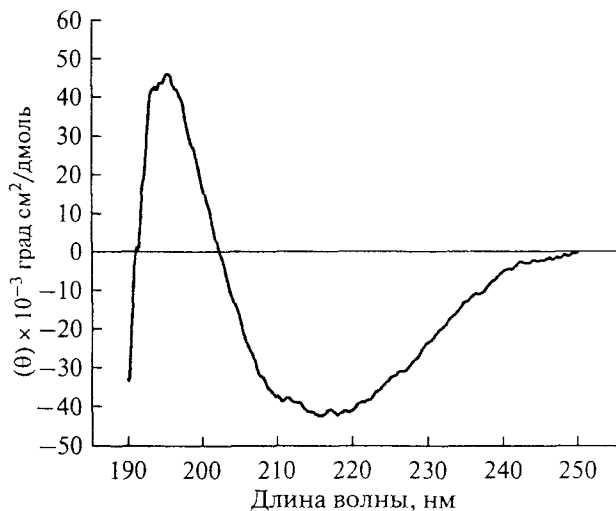


Рис. 5. КД-спектр gr5HENS-37VPA (10 мМ Na-фосфатный буфер).

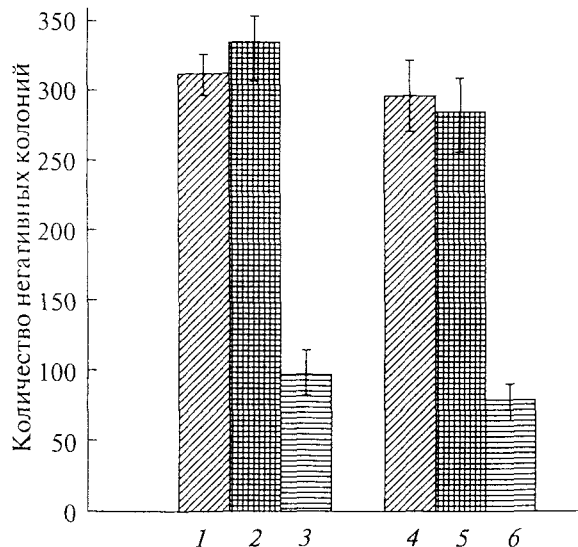


Рис. 6. Инфекционность бактериофага Т4 на клетках *E. coli* В/Е1 — без добавления белка (1, 4), в присутствии белка gr5HENS-37VPA в концентрации 0.3 мг/мл (3) и 2 мг/мл (6), в присутствии бычьего сывороточного альбумина (BSA) в качестве контроля в тех же концентрациях (2 и 5 соответственно)

при помощи программы 3D-PSSN [31] и предположенному на основе исследования целых ДХФ [3] строению С-концевой “иглы” gr37. Таким образом, можно обоснованно предположить, что присоединенный к β-спиральному белку-носителю фрагмент сохраняет свою нативную структуру.

Это предположение подтверждается измерением биологической активности химерного белка. В силу того, что контакт ДХФ фага с поверхностью клетки — обратимый процесс, а рецептор связывания ДХФ точно не определен [2], выделение комплекса и установление физико-химических параметров связывания невозможно. Однако наблюдается заметное ингибирование процесса инфекции клеток *E. coli* бактериофагом Т4 в присутствии белка gr5HENS-37VPA (рис. 6).

По совокупности проведенных экспериментов можно констатировать, что использование β-спирального домена gr5 бактериофага Т4 — перспективное направление белковой инженерии, предназначенное для получения рекомбинантных белков, которые по той или иной причине не способны принять биологически активную конформацию в организме-продуценте. В ходе работы был создан плазмидный вектор, который может использоваться для клонирования таких белков.

### ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

Процедуры ПЦР и молекулярного клонирования проводили стандартными методами [32] с учетом рекомендаций фирм-производителей. Использовали штаммы *E. coli* NovaBlue, BL21(DE3) (Novagen) и В/Е1, эндонуклеазы рестрикции, Т4-ДНК-лигазу и соответствующие буферы для реакций (Fermentas).

**Конструирование плазмиды p5HENS (His-tag — Enterokinase — Non-Stop).** Последовательность ДНК,

кодирующую С-концевой домен белка gr5 (остатки 435–575) ПЦР-амплифицировали с матрицы генома фага Т4 с использованием праймеров (5' → 3') **GGTCATATGGTTGATAATTGGTATGATATA** и **GGAAGTGGATCCGCCAATGTCAATCCTCGA** (сайты узнавания NdeI и BamHI выделены). Фрагмент обработали эндонуклеазами NdeI и BamHI и клонировали в вектор pET-19b(+) (Novagen), расщепленный теми же рестриктазами. Из получившейся плазмиды был вырезан фрагмент, ограниченный сайтами рестриктаз NcoI и BamHI, и клонирован в вектор pET-23d(+) (Novagen) по соответствующим сайтам. Конечная плазмиды несет под контролем промотора Т7 ген, кодирующий полипептид следующего непрерывного строения: 10xHis-метка для очистки хелатно-аффинной хроматографией, участок, подвергаемый действию энтерокиназы для отщепления метки, С-концевой домен белка gr5 бактериофага Т4 (140 а.о.). Стоп-кодон белка вырожден в остатки Gly-Ser, после которого с участием уникальных сайтов рестрикции вектора pET возможно клонирование генов, которые будут экспрессироваться в единой рамке считывания с белком-носителем.

**Конструирование плазмиды p5HENS-37VPA.** Последовательность ДНК, кодирующую С-концевую часть белка ДХФ gr37 бактериофага Т4 (остатки 826–1026) амплифицировали с матрицы генома фага Т4 с использованием праймеров **TCAGGATCCGTTCCCTGCTGGATTT** и **AAGCGGCCGCCTTATGCTAAACGAA** (сайты узнавания BamHI и NotI выделены). Фрагмент обработали рестриктазами BamHI и NotI и клонировали в вектор p5HENS по

тем же сайтам В качестве контроля тот же фрагмент ДНК клонировали по тем же сайтам в вектор рЕТ-23d(+) Очистку плазмид производили с помощью колонок Qiafilter Spin (Qiagen) Ствоение полученных плазмид проверяли с помощью ПЦР по фланкирующим сайтам, рестрикционным анализом и ДНК-секвенированием

**Экспрессия, выделение и очистка рекомбинантного белка gp5HENS-37VPA.** Плазмидный вектор трансформировали для экспрессии в клетки *E coli* BL21(DE3) Клетки наращивали в среде 2xYT при 37°C до  $A_{600} \sim 0.6$ , синтез белка индуцировали добавлением изопропил-1-тио- $\beta$ -D-галактозида до концентрации 0.5 mM и инкубировали в течение 4 ч при 37°C при умеренной аэрации Клетки *E coli* из 0.5 л среды осаждали центрифугированием при 3500 об/мин (Sorvall RC5, ротор GS-3) в течение 15 мин, осадок ресуспендировали в 30 мл буфера А (10 mM Трис-НСl pH 8.0, 200 mM NaCl, 1 mM PMSF) Клетки разрушали ультразвуком (Techpan MD20, 5 импульсов по 15 с), нерастворимые фрагменты клеток отделяли центрифугированием при 15000 об/мин (Sorvall RC5, ротор SA600) в течение 20 мин Фильтрованный через 0.45 мкм фильтр супернатант наносили на 5 мл колонку HighTrap Chelating (GE Amersham), заряженную хлоридом никеля Элюцию белка производили линейным градиентом концентрации имидазола (0–200 mM) в буфере А Очищенный препарат белка диализовали против буфера Б (20 mM буфера Трис-НСl pH 7.5, 50 mM NaCl, 1 mM PMSF) и концентрировали с помощью Centrprep 10 (Millipore) Концентрацию белка определяли методом Брэдфорд с помощью набора реактивов Bio-Rad

**Определение молекулярной массы белка.** Олигомерное состояние белка определяли с помощью аналитической гель-фильтрации Препараты очищенного gp5HENS-37VPA в концентрации 4 мг/мл наносили на колонку Toyopearl HW-55 (Tosoh Bioscience) 0.75 × 30 см, уравновешенную буфером Б, и элюировали со скоростью 0.2 мл/мин Предварительно колонку калибровали стандартами белков (М, кДа) альдолаза (158), бычий сывороточный альбумин (66), яичный альбумин (45), цитохром с (12) По графику зависимости логарифма молекулярной массы белка от объема элюции определяли наблюдаемую молекулярную массу белка

**КД-спектры** снимали с помощью спектрополяриметра JASCO J-500 в 10 mM Na-фосфатном буфере pH 7.5 при комнатной температуре Расчет содержания вторичной структуры проводили по методу Провенчера [33] с использованием программ CONTIN [34] или Selcon [35]

**Определение биологической активности белков (конкурентное связывание).** Культуру *E coli* штамма В/Е1 подращивали в среде 2xTY (2.5–3 ч), далее к 100 мкл аликвотам клеточной культуры добавляли 100 мкл смеси, содержащей 50 мкл фага Т4 определенного титра и 50 мкл раствора белка в concentra-

ции 0.3 и 2 мг/мл и инкубировали 1.5 мин В качестве контроля к культуре клеток добавляли 50 мкл суспензии фага нужного разведения и 50 мкл физиологического раствора Реакцию останавливали добавлением 500 мкл физ. раствора и охлаждением на льду Далее реакционную смесь смешивали с верхним агаром Hershey, который наносили на чашки Петри с нижним агаром Чашки инкубировали при 37°C в течение ночи, после чего подсчитывали количество негативных колоний, образованных фагом Т4

## БЛАГОДАРНОСТИ

Работа выполнена при финансовой поддержке грантов РФФИ № 02-04-50012-а, 05-04-49636-а и 09-04-00335-а Авторы благодарят И.А. Куделину (ИБХ РАН) за помощь в выполнении КД-спектроскопии, М.В. Фильчикова и П.А. Назарова (ИБХ РАН) за активное участие на различных стадиях этого проекта

## СПИСОК ЛИТЕРАТУРЫ

- 1 Wood W B, Eiserling F A, Crowther R A // Molecular Biology of Bacteriophage T4 / Ed Karam J Washington D C ASM Press, 1994 P 282–290
- 2 Hashemolhosseini S, Montag D, Kramer L, Henning U // J Mol Biol 1994 V 241 P 524–533
- 3 Cerritelli M E, Wall J S, Simon M N, Conway J F, Steven A C // J Mol Biol 1996 V 260 P 767–780
- 4 Tetart F, Repoula F, Monod C, Krsch H M // J Mol Biol 1996 V 258 P 726–731
- 5 Qu Y, Hyman P, Harrah T, Goldberg E // J Bacteriol 2004 V 186 P 8363–8369
- 6 Симакова М Н, Симаков Н Н // Мол биология 2005 Т 39 С 321–329
- 7 Марусич Е И, Курочкина Л П, Месянжинов В В // Биохимия 1998 Т 63 С 473–482
- 8 Matsui T, Grnnaviene B E, Goldberg B, Tsugita A, Takana N, Arisaka F // J Bacteriol 1997 V 179 P 1846–1851
- 9 Sørensen H P, Mortensen K K // J Biotechnol 2005 V 115 P 113–128
- 10 Davis G D, Elisee C, Newham D M, Harrison R G // Biotechnol Bioeng 1999 V 65 P 382–388
- 11 Tao Y, Strelkov S V, Mesyanzhinov V V, Rossmann M G // Structure 1997 V 5 P 789–798
- 12 Papanikolopoulou K, van Raaij M J, Mitraki A // Methods Mol Biol 2008 V 474 P 15–33
- 13 Kanamaru S, Leiman P G, Kostyuchenko V A, Chipman P R, Mesyanzhinov V V, Arisaka F, Rossmann M G // Nature 2002 V 415 P 553–557
- 14 Baneyx F, Mujacic M // Nat Biotechnol 2004 V 22 P 1399–1408
- 15 Panchenko A R, Luthey-Schulten Z, Cole R, Wolynes P G // J Mol Biol 1997 V 272 P 95–105
- 16 Maity H, Maity M, Krishna M M, Mayne L, Englander S W // Proc Natl Acad Sci USA 2005 V 102 P 4741–4746

- 17 Летаров А В, Лондер Ю Ю, Будько С П, Месянжинов В В // Биохимия 1999 Т 64 С 974–981
- 18 Boudko S P, Londer Y Y, Letarov A V, Sernova N V, Engel J, Mesyanzhinov V V // Eur J Biochem 2002 V 269 P 833–841
- 19 Miroshnikov K A, Marusich E I, Cerritelli M E, Cheng N, Hyde C C, Steven A C, Mesyanzhinov V V // Protein Eng 1998 V 11 P 329–332
- 20 Мирошников К А, Сернова Н В, Шнейдер М М, Месянжинов В В // Биохимия 2000 Т 65 С 1600–1606
- 21 Bhardwaj A, Walker-Kopp N, Wilkens S, Singolani G // Protein Sci 2008 V 17 P 1475–1485
- 22 Frank S, Kammerer R A, Mechling D, Schulthess T, Landwehr R, Bann J, Guo Y, Lustig A, Bachinger H P, Engel J // J Mol Biol 2001 V 308 P 1081–1089
- 23 Stetefeld J, Frank S, Jenny M, Schulthess T, Kammerer R A, Boudko S, Landwehr R, Okuyama K, Engel J // Structure 2003 V 11 P 339–346
- 24 Yang X, Lee J, Mahony E M, Kwong P D, Wyatt R, Sodroski J // J Virol 2002 V 76 P 4634–4642
- 25 Sissoeff L, Mousli M, England P, Tuffereau C // J Gen Virol 2005 V 86 P 2543–2552
- 26 Papanikolopoulou K, Forge V, Goeltz P, Mitraki A // J Biol Chem 2004 V 279 P 8991–8998
- 27 Papanikolopoulou K, Teixeira S, Belrhali H, Forsyth V T, Mitraki A, van Raay M J // J Mol Biol 2004 V 342(1) P 219–227
- 28 Nakagawa H, Arisaka F, Ishu S // J Virol 1985 V 54 P 460–466
- 29 Месянжинов В В, Лейман П Г, Костюченко В А, Курочкина Л П, Мирошников К А, Сыкилинда Н Н, Шнейдер М М // Биохимия 2004 Т 69 С 1463–1476
- 30 van Raay M J, Schoehn G, Jaquinod M, Ashman K, Burda M R, Miller S // J Biol Chem 2001 V 382 P 1049–1055
- 31 Kelley L A, MacCallum R M, Sternberg M J // J Mol Biol 2000 V 299 P 499–520
- 32 Sambrook R, Russel D W // Molecular Cloning A Laboratory Manual 3d ed Cold Spring Harbor, N Y, Cold Spring Harbor Lab Press, 2001
- 33 Provencher S W, Glockner J // Biochemistry 1981 V 20 P 33–37
- 34 Venyaminov S Y, Baikalov I A, Shen Z M, Wu C -S C, Yang J T // Anal Biochem 1993 V 214 P 17–24
- 35 Sreerama N, Woody R W // Anal Biochem 1993 V 209 P 32–44

## The $\beta$ -Helical Domain of Bacteriophage T4 Controls the Folding of the Fragment of Long Tail Fibers in a Chimeric Protein

R. N. Chuprov-Netochin, N. M. Faizullina, N. N. Sykilinda, M. N. Simakova, V. V. Mesyanzhinov, and K. A. Miroshnikov#

#Phone/fax: (495) 336-60-22; e-mail: kmi@ibch.ru

Shemyakin–Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, ul Miklukho-Maklaya 16/10, Moscow, 117997 Russia

The key stage of the infection of the *Escherichia coli* cell with bacteriophage T4, the binding to the surface of the host cell, is determined by the specificity of the long tail fiber proteins of the phage, in particular, gp37. The assembly and oligomerization of this protein under natural conditions requires the participation of at least two additional protein factors, gp57A and gp38, which strongly hinders the production of the recombinant form of gp37. To overcome this problem, a modern protein engineering strategy was used, which involves the construction of a chimeric protein containing a carrier protein that drives the correct folding of the target protein. For this purpose, the trimeric  $\beta$ -helical domain of another protein of phage T4, gp5, was used. It was shown that this domain, represented as a rigid trimeric polypeptide prism, has properties favorable for use as a protein carrier. A fragment of protein gp37 containing five pentapeptides repeats, Gly-X-His-X-His, which determine the binding to the receptors on the bacterial cell surface, was fused in a continuous reading frame to the C-terminus of the domain of gp5. The resulting chimeric protein forms a trimer that has the native conformation of gp37 and exhibits biological activity.

*Key words* bacteriophage T4, tail fibers, protein folding, protein engineering