



УДК 547.962.02

ПЕРВИЧНАЯ СТРУКТУРА ФАКТОРА ЭЛОНГАЦИИ G  
ИЗ *ESCHERICHIA COLI*IX. ИССЛЕДОВАНИЕ СТРУКТУРЫ ПЕПТИДОВ  
БРОМЦИАНОВОГО РАСЩЕПЛЕНИЯ G-ФАКТОРА, ВЫДЕЛЕННЫХ  
НА ТИОЛ-АКТИВИРОВАННОЙ СЕФАРОСЕ, И ПРОДУКТОВ РАСЩЕПЛЕНИЯ  
G-ФАКТОРА ПО СВЯЗЯМ Asp-Pro.  
ПОЛНАЯ ПЕРВИЧНАЯ СТРУКТУРА\**Овчинников Ю. А., Алахов Ю. В., Бундулис Ю. П.,  
Бундуле М. А., Винокуров Л. М., Мотуз Л. П., Козлов В. И.**Институт белка Академии наук СССР, г. Пушкино Московской обл.*

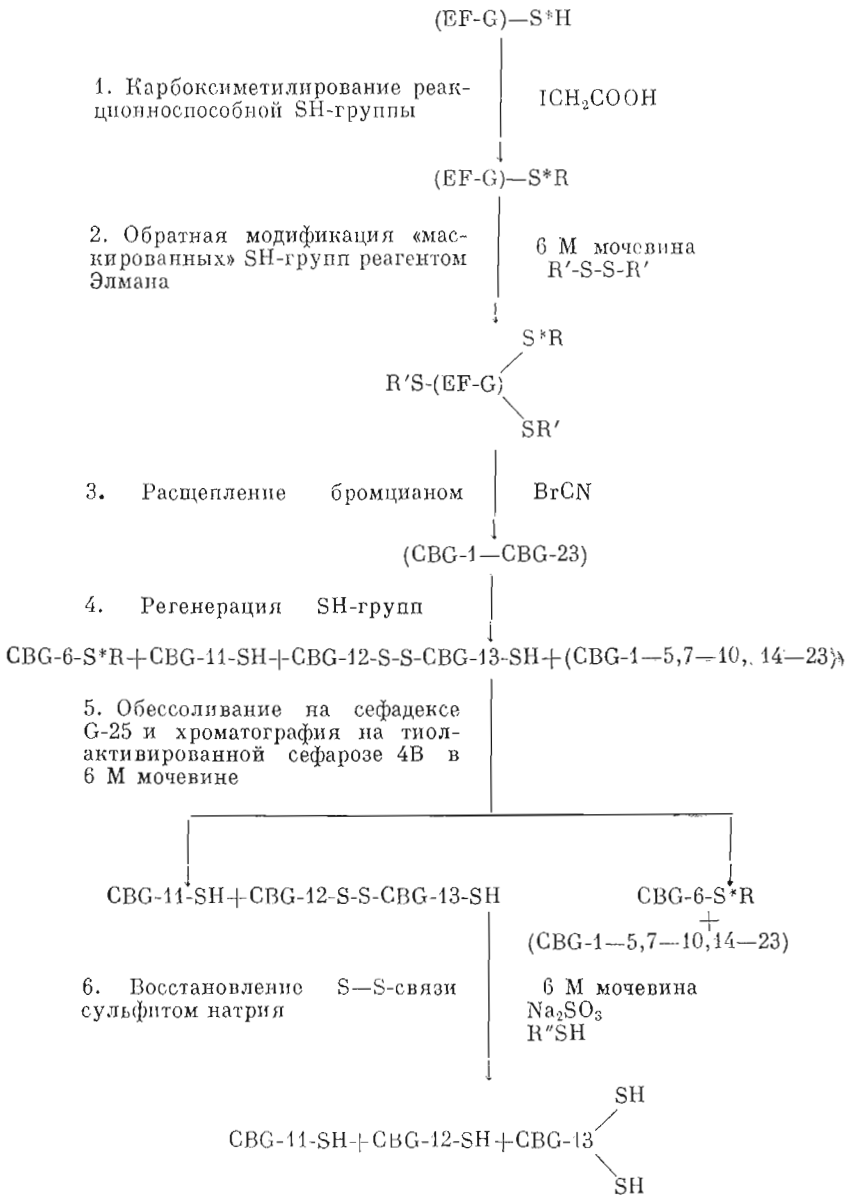
Определена аминокислотная последовательность цистеин- и цистинсодержащих пептидов, полученных в результате расщепления G-фактора бромцианом. Для их структурного анализа использовали расщепление бромциановых пептидов или их фрагментов трипсином, химотрипсином, термолизином, стафилококковой глутаминовой протеиназой и ограниченный кислотный гидролиз. Исследованы продукты расщепления G-фактора по связям Asp-Pro. Полученные данные в сумме с ранее опубликованными позволили представить полную первичную структуру фактора элонгации G. Полипептидная цепь белка состоит из 701 аминокислотного остатка и имеет точную молекулярную массу 77321,46.

Свойства сульфгидрильных групп фактора элонгации G были исследованы титрованием тиольными реагентами: 5,5'-дитиобис(2-нитробензоатом), *n*-оксимеркурибензоатом и *N*-этилмалеилмидом. Было показано, что G-фактор содержит три [2] или четыре [3] сульфгидрильные группы на молекулу белка. По данным аминокислотного анализа, где остатки цистеина определялись в виде цистеиновой кислоты, молекула белка содержит 5 [4] или 6 [5] остатков цистеина.

Показано, что один остаток цистеина особенно легко модифицируется сульфгидрильными реагентами, что приводит к ингибированию как взаимодействия G-фактора с рибосомой, так и несопряженной GTP-азной реакции [2, 4, 6, 7]. Вместе с тем известно, что гуаниловые нуклеотиды защищают эту свободную SH-группу от действия тиольных реагентов [3]. Фотохимической модификацией G-фактора аналогом GTP было показано, что район полипептидной цепи белка, включающий экспонированную SH-группу, непосредственно участвует в связывании GTP [8], хотя сама SH-группа не входит непосредственно в GTP-связывающий участок. Исследование фрагментов ограниченного трипсинолиза G-фактора в условиях, когда белок сохраняет нативную структуру, в частности фрагмента T<sub>7</sub>, позволило локализовать единственный экспонированный остаток цистеина [9, 10]. Далее было показано, что этот остаток цистеина находится в положении 113 полипептидной цепи белка [11]. При исследовании продуктов бромцианового расщепления G-фактора [12] оказалось, что еще четыре остатка цистеина входят в состав фрагмента T<sub>1</sub> (обозначение фрагментов см. в [13]). Пептиды бромцианового расщепления белка, содержащие остатки цистеина, были выделены в количествах, недостаточных для установления их полных аминокислотных последовательностей [12]. Установить места расположения остатков цистеина, два из которых выявляются только после восстановления сульфитом патрия, не удалось также после исследования продуктов гидролиза фрагмента T<sub>1</sub> глутаминовой протеиназой [14] и трипсином [1]. Для выяснения структуры цистеинсодер-

\* Сообщение VIII см. [1].

Ступенчатое выделение пептидов бромцианового расщепления G-фактора, содержащих «маскированные» SH-группы и S—S-связь



жащих пептидов бромцианового расщепления G-фактора и локализации цистеиновых остатков, образующих S—S-связь, нами была разработана методика выделения этих пептидов на тиол-активированной сефарозе 4В (схема 1).

Поскольку структура участка полипептидной цепи белка, включающего экспонированную SH-группу, была уже известна, эта SH-группа была блокирована иодуксусной кислотой. Перед расщеплением G-фактора бромцианом он был обработан реагентом Элмана в 6 М мочеvine для предотвращения окисления свободных SH-групп. После расщепления бромцианом продукты были обработаны β-меркаптоэтанолом для удаления защитных групп с остатков цистеина и после обессоливания хроматографировались на колонке с тиол-активированной сефарозой 4В. Связывание с носителем контролировали по увеличению поглощения при 343 нм. Оказалось, что с тиол-сефарозой связалось три пептида: CBG-11, CBG-12 и

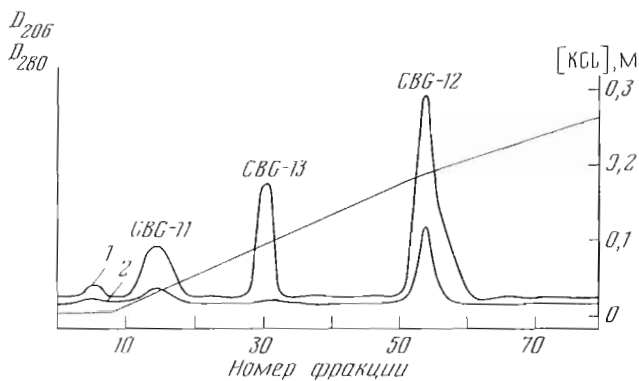


Рис. 1. Хроматография смеси пептидов CBG-11 – CBG-13 бромцианового расщепления G-фактора на колонке (1,0×10 см) с QAE-сефадексом А-25, уравновешенным буфером 20 мМ трис-НСl (рН 8,0), содержащим 6 М мочевины, в градиенте концентрации КСl. Скорость элюирования 10 мл/ч, объем фракций 10 мл. 1 –  $D_{206}$ , 2 –  $D_{280}$

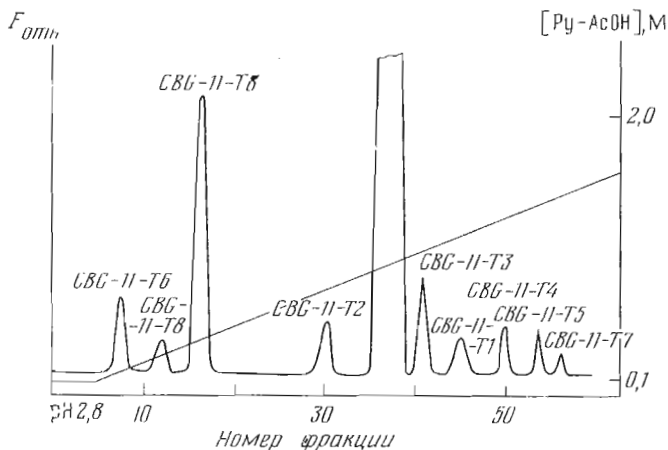
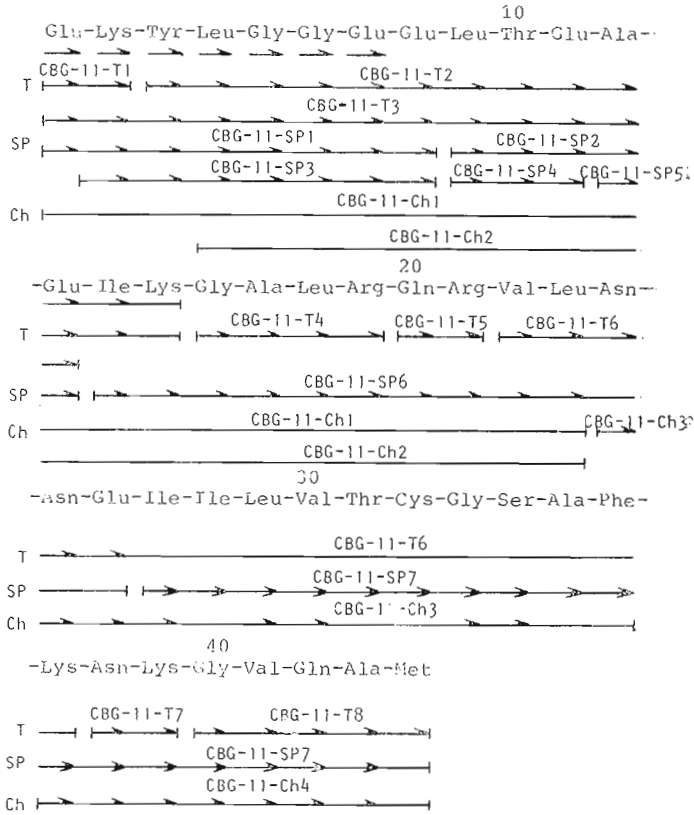


Рис. 2. Высокоэффективная хроматография смеси пептидов, полученной после триптического гидролиза пептида CBG-11 на колонке (0,6×40 см) с Chromobeads (тип Р) в градиенте концентрации пиридин-ацетата (0,1–2,0 М) и рН (2,8–5,5). Скорость элюирования 20 мл/ч, объем фракций 1 мл

CBG-13 (обозначения пептидов бромцианового расщепления G-фактора см. в работе [12]). Два пептида были ковалентно связаны S–S-связью и после восстановления S–S-связи сульфитом натрия разделялись на QAE-сефадексе А-25 (рис. 1). Дополнительное количество этих пептидов было выделено из элюата, содержащего пептиды, не связанные с тиол-активированной сефарозой, после обработки его сульфитом натрия и повторной хроматографии на тиол-активированной сефарозе. Полученная смесь тиолсодержащих пептидов была карбоксиметилирована иодуксусной кислотой, затем восстановлена сульфитом натрия и карбоксиметилирована [ $^{14}$ C]иодуксусной кислотой. После разделения на QAE-сефадексе А-25 все три тиолсодержащих пептида бромцианового расщепления G-фактора были выделены в чистом виде. Используя такой подход, нам удалось выделить все три пептида с общим выходом около 50%.

Пептид CBG-11, по данным аминокислотного анализа, содержит 44 аминокислотных остатка. Ручным методом Эдмана (в дансильной модификации) была установлена последовательность первых восьми остатков. На первом этапе структурных исследований пептида CBG-11 нами был выбран триптический гидролиз, поскольку количество остатков лизина и аргинина позволяло получить набор пептидов, аминокислотную последовательность которых легко установить ручным методом Эдмана. После

## Исследование структуры пептида CBG-11



*Примечания.* --- здесь и далее — определение аминокислотной последовательности методом Эдмана в дансильной модификации; → — определение аминокислотной последовательности автоматическим методом Эдмана; T — пептиды триптического гидролиза; SP — пептиды гидролиза глутаминовой протеиназой; Ch — пептиды триптического гидролиза.

разделения смеси пептидов методом высокоэффективной ионообменной хроматографии (рис. 2) были выделены и охарактеризованы все 7 пептидов, полученных при триптическом гидролизе пептида CBG-11. N-Концевая аминокислотная последовательность пептида CBG-11-T6 оказалась идентичной последовательности пептида LTIV-11, выделенного из продуктов триптического расщепления фрагмента T<sub>4</sub>, модифицированного малеиновым ангидридом [1]. Единственный остаток цистеина находится в этом пептиде, но он не содержит радиоактивную метку и, следовательно, не образует S-S-связи.

Для получения перекрытий между триптическими пептидами и установления аминокислотной последовательности неизвестного участка пептида CBG-11-T6 нами был предпринят гидролиз пептида CBG-11 глутаминовой протеиназой. Смесь пептидов разделяли двумерной техникой на пластинках с тонким слоем целлюлозы. В результате были выделены пептиды CBG-11-SP1—CBG-11-SP7 (схема 2). Исследование аминокислотной последовательности пептида CBG-11-SP6 позволило объединить в одну полипептидную цепь триптические пептиды CBG-11-T3, CBG-11-T4, CBG-11-T5 и CBG-11-T6. Пептид CBG-11-SP7 был выделен с незначительным выходом, так что ручным методом Эдмана удалось установить последовательность только первых трех аминокислотных остатков.

Для выяснения структуры C-концевой части пептида CBG-11 мы исследовали продукты химотриптического гидролиза этого пептида. После

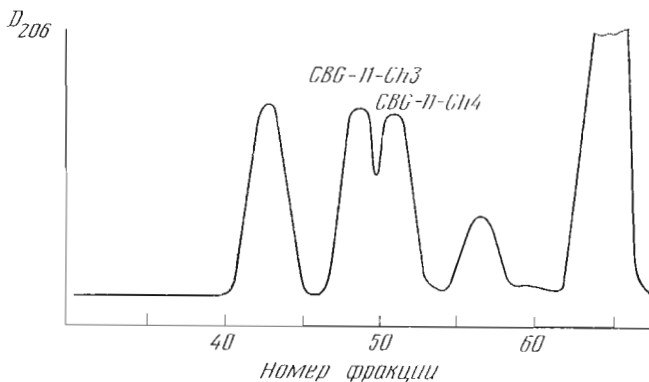


Рис. 3. Гель-фильтрация смеси пептидов химотриптического гидролиза пептида СВГ-11 на колонке (1,0×200 см) с сефадексом G-25, уравновешенным аммиачной водой, pH 9. Скорость элюирования 2 мл/ч, объем фракций 2 мл

разделения смеси пептидов методом гель-фильтрации на сефадексе G-25 (рис. 3) были выделены интересные нас пептиды СВГ-11-Сh3 и СВГ-11-Сh4 (схема 2). Полную аминокислотную последовательность пептида СВГ-11-Сh3 установить не удалось, а пептид СВГ-11-Сh4 объединял триптические пептиды СВГ-11-Т7 и СВГ-11-Т8. Для завершения работы по установлению первичной структуры пептида СВГ-11 было необходимо подтвердить перекрывание пептидов СВГ-11-Т6 и СВГ-11-Т7 (схема 2), а также положение некоторых аминокислотных остатков в С-концевой части пептида СВГ-11-Сh3. Повторное исследование структуры пептида СВГ-11, содержащего гомосерин, твердофазным автоматическим методом Эдмана без предварительного разделения смеси пептидов дало возможность установить полную аминокислотную последовательность пептида СВГ-11. Единственный остаток цистеина этого пептида находится в положении 32 (схема 2).

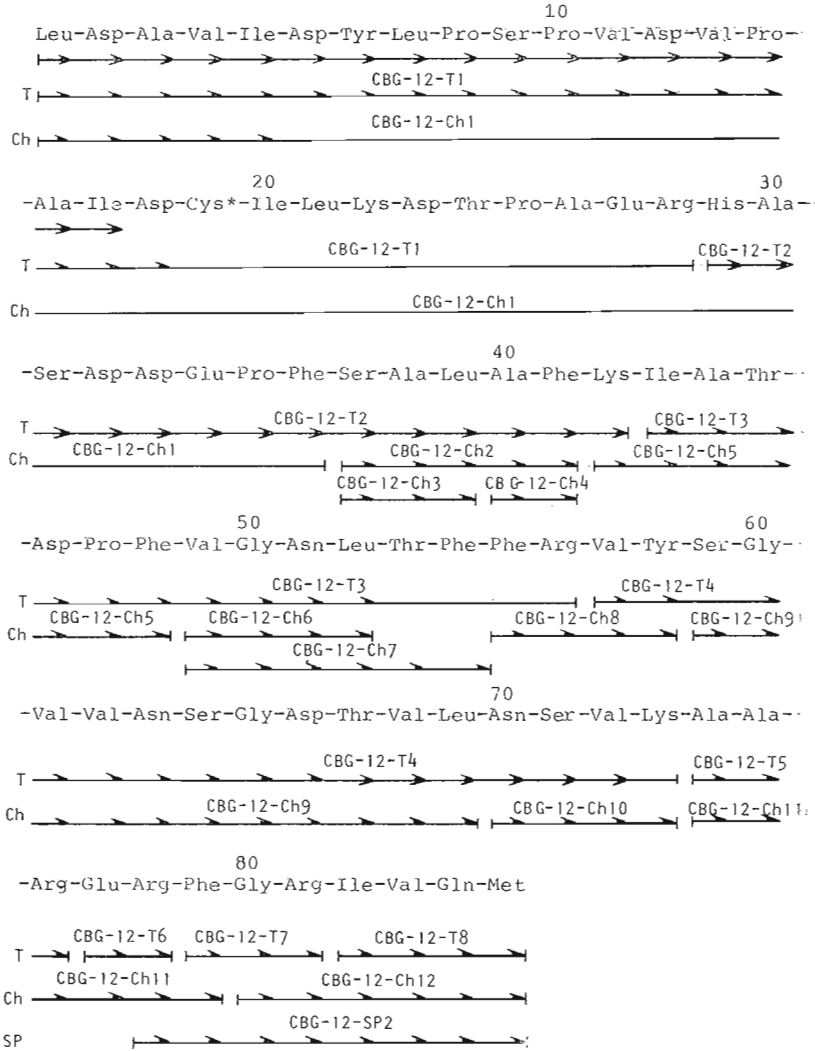
Пептид СВГ-12 является самым крупным пептидом бромцианового расщепления G-фактора (85 аминокислотных остатков). В результате разделения смеси пептидов триптического гидролиза СВГ-12 методом высокоэффективной хроматографии на колонке с ионообменной смолой Chromobeads (тип Р) были выделены и охарактеризованы пептиды СВГ-12-Т1—СВГ-12-Т8 (схема 3). Пептид СВГ-12-Т1, по данным аминокислотного анализа, содержит как остаток лизина, так и остаток аргинина. Кроме того, этот пептид содержит остаток цистеина с радиоактивной меткой. При исследовании структуры пептида СВГ-12-Т1, так же как и исходного пептида СВГ-12, удалось установить последовательность лишь 18 аминокислотных остатков и поэтому остаток цистеина, образующий S—S-связь, локализован не был. Пептид СВГ-12-Т2 оказался идентичным пептиду LTIV-12, а пептиды СВГ-12-Т4 и СВГ-12-Т5 в сумме соответствовали пептиду LTIV-14 триптического гидролиза фрагмента T<sub>1</sub>, модифицированного малеиновым ангидридом (обозначения триптических пептидов см. в работе [1]). Структура мелких пептидов СВГ-12-Т6, СВГ-12-Т7 и СВГ-12-Т8 была установлена полностью ручным методом Эдмана (в дансильной модификации).

Ферментативное расщепление пептида СВГ-12 глутаминовой протеназой привело к образованию двух пептидов. После их разделения на колонке с сефадексом G-25 оказалось, что больший из них охватывает почти всю последовательность пептида СВГ-12, а исследование структуры другого пептида (СВГ-12-SP2, схема 3) позволило объединить в полипептидную цепь триптические пептиды СВГ-12-Т6, СВГ-12-Т7 и СВГ-12-Т8.

Для выяснения аминокислотной последовательности неизвестного участка в N-концевой части пептида СВГ-12, а также получения недостающих перекрываний пептидов триптического гидролиза было использовано расщепление пептида СВГ-12 химотрипсином. В результате разделения

Схема 3

Исследование структуры пептида CBG-12



смеси пептидов методом гель-фильтрации на колонке с сефадексом G-25 (рис. 4) был получен пептид CBG-12-Ch1 и фракции, содержащие мелкие пептиды, из которых после двумерного разделения на пластинках с тонким слоем целлюлозы были выделены пептиды CBG-12-Ch2—CBG-12-Ch12. Исследование структуры этих коротких пептидов (схема 3) позволило установить аминокислотную последовательность 50 остатков C-концевой части пептида CBG-12.

Для установления полной структуры CBG-12 было необходимо определить аминокислотную последовательность пептида CBG-12-Ch1, который был подвергнут гидролизу трипсином. После предварительной гель-фильтрации смеси пептидов, полученных при триптическом гидролизе пептида CBG-12-Ch1, на колонке с сефадексом G-25 и дальнейшего двумерного разделения полученных фракций на пластинках с тонким слоем целлюлозы был получен пептид CBG-12-Ch1-T2, который не образуется при триптическом гидролизе исходного пептида CBG-12. Поскольку в состав пептида CBG-12-T1 (схема 3) входят остатки лизина и аргинина, мы предполагаем, что пептид CBG-12-Ch1-T2 является C-концевым участком пептида CBG-12-T1, в котором связь Lys-Asp не расщепляется из-за близости остатка пролина (схема 3). Для проверки этого предположения, а также

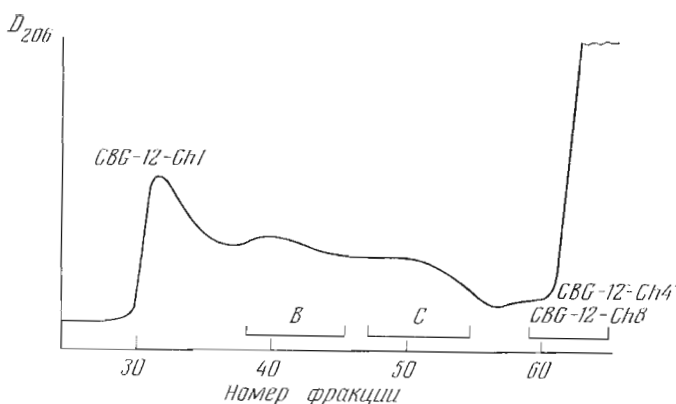
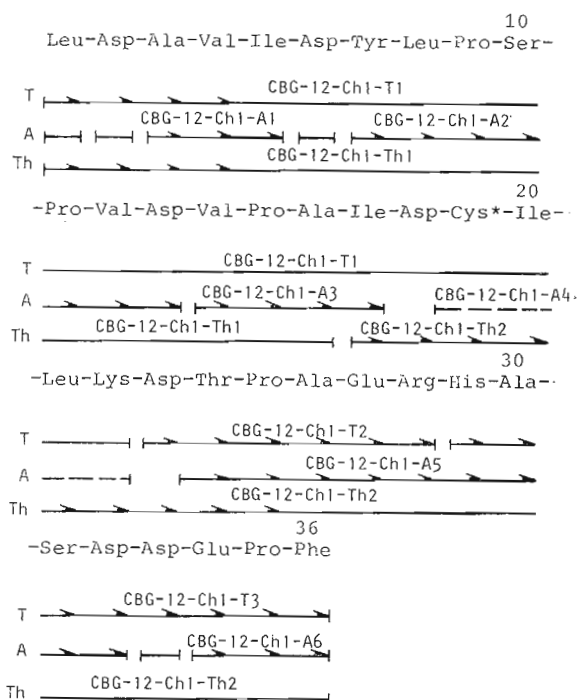


Рис. 4. Гель-фильтрация смеси пептидов химотриптического гидролиза пептида СВГ-12 на колонке (1,0×200 см) с сефадексом G-25, уравновешенным аммиачной водой, pH 9. Скорость элюирования 2 мл/ч, объем фракций 2 мл

с целью исследования труднодоступного участка пептида СВГ-12 (остатки 18–30) мы решили применить кислотный гидролиз с выщеплением остатков аспарагиновой кислоты. В результате разделения полученной смеси пептидов на пластинках с тонким слоем целлюлозы и после исследования аминокислотной последовательности выделенных чистых пептидов было получено перекрытие пептидов СВГ-12-Ch1-T2 и СВГ-12-Ch1-T3, а также подтверждена структура N-концевой части пептида СВГ-12 до положения 18 (схема 4). Пептид СВГ-12-Ch1-A4, содержащий остаток цистеина, выделить не удалось. Наличие остатка изолейцина не

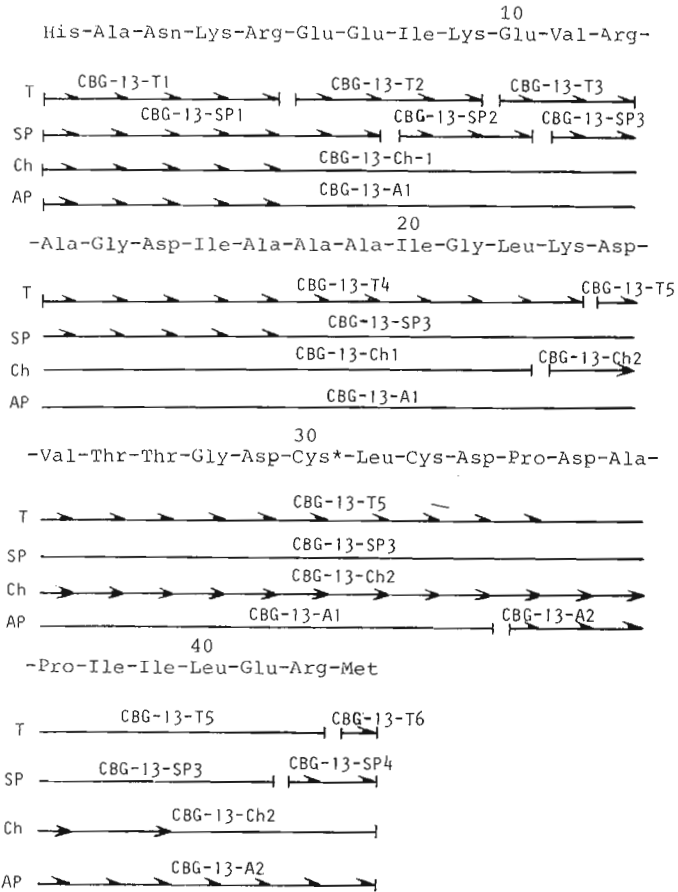
Схема 4

Исследование структуры пептида СВГ-12-Ch1



Примечания. А — пептиды, полученные кислотным гидролизом с выщеплением остатков аспарагиновой кислоты; Th — пептиды термолитического гидролиза.

## Исследование структуры пептида CBG-13



*Примечание.* AP — пептиды, полученные после расщепления CBG-13 по связям Asp-Pro.

посредственно перед неизвестной частью пептида CBG-12-Ch1 дало возможность использовать термолитический гидролиз для установления структуры цистеинсодержащего участка. Разделение термолитического гидролизата на пластинках с тонким слоем целлюлозы позволило выделить пептид CBG-12-Ch1-Th2, который содержал остаток цистеина и давал перекрывание пептидов CBG-12-Ch1-T1 и CBG-12-Ch1-T2 (схема 4). Аминокислотная последовательность этого пептида была установлена ручным методом Эдмана (в дансильной модификации) до положения 9, и тем самым была завершена работа по установлению полной структуры пептида CBG-12. Остаток цистеина, образующий S-S-связь, занимает 19-е положение в полипептидной цепи пептида CBG-12 (схема 3).

На первом этапе исследования структуры пептида CBG-13 (43 аминокислотных остатка) нами был избран триптический гидролиз. В результате деления смеси пептидов на пластинках с тонким слоем целлюлозы были выделены пептиды CBG-13-T1—CBG-13-T5 и гомосерин. Пептид CBG-13-T5 содержал радиоактивную метку; следовательно, в нем находится остаток цистеина, образующий S-S-связь. Структура пептидов CBG-13-T1—CBG-13-T4 была установлена полностью (схема 5), а пептида CBG-13-T5 — до положения 11. Оказалось, что этот пептид содержит два остатка цистеина (положения 7 и 9), первый из которых образует S-S-связь с остатком цистеина пептида CBG-12. Расположение остатков



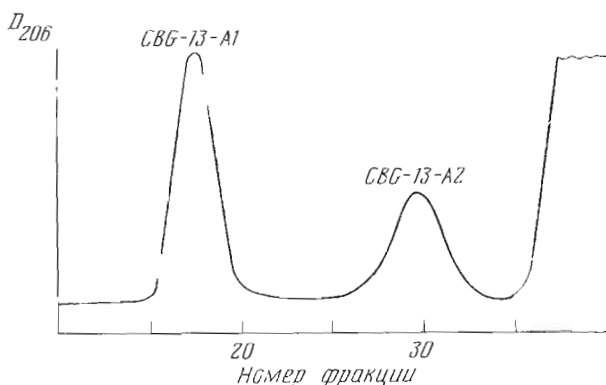


Рис. 5. Гель-фильтрация продуктов расщепления пептида СВГ-13 по связи Asp-Pro на колонке (1,0×200 см) с сефадексом G-25, уравновешенным 0,01 н. HCl. Скорость элюирования 2,8 мл/ч, объем фракций 2,8 мл

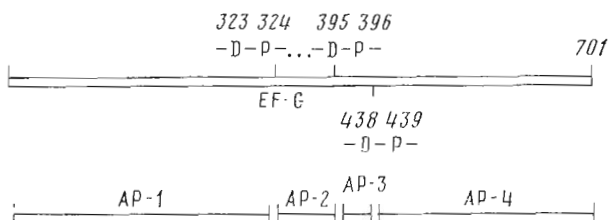


Рис. 6. Расположение фрагментов, полученных в результате расщепления G-фактора по связям Asp-Pro

цистеина в полипептидной цепи СВГ-13 и аминокислотная последовательность С-концевой части пептида СВГ-13-T5 были установлены при исследовании структуры химотриптического пептида СВГ-13-Ch2 твердофазным автоматическим методом Эдмана без предварительного разделения смеси химотриптических пептидов (схема 5).

Для получения перекрытий между пептидами триптического гидролиза пептид СВГ-13 был подвергнут расщеплению глутаминовой протеиназой. В результате двухэтапного выделения (I этап — гель-фильтрация на сефадексе G-15, II этап — двумерное разделение на пластинках с тонким слоем целлюлозы) были получены пептиды СВГ-13-SP1—СВГ-13-SP4. Исследование структуры пептидов СВГ-13-SP1, СВГ-13-SP2 и СВГ-13-SP3 позволило объединить в полипептидную цепь триптические пептиды СВГ-13-T1, СВГ-13-T2, СВГ-13-T3 и СВГ-13-T4 (схема 5). Пептид СВГ-13-SP4 оказался дипептидом Arg-Met.

Для установления полной аминокислотной последовательности пептида СВГ-13 было необходимо определить последовательность С-концевой части этого пептида. Удачное расположение связи Asp-Pro позволило применить расщепление по этой связи. Установление структуры пептида СВГ-13-A2 (схема 5) ручным методом Эдмана в дансильной модификации, после его выделения на колонке с сефадексом G-25 (рис. 5), явилось завершающей стадией исследования структуры пептида СВГ-13. Было показано, что остаток цистеина, образующий S—S-связь, находится в положении 30, а второй остаток цистеина — в положении 32 в полипептидной цепи СВГ-13.

Данные, полученные при изучении аминокислотных последовательностей пептидов бромцианового расщепления G-фактора [12], в сочетании с результатами исследования структуры пептидов, образовавшихся при расщеплении смеси фрагментов T<sub>3</sub> и T<sub>4</sub> глутаминовой протеиназой [14] и трипсином [1], и структуры фрагментов T<sub>6</sub> [11], T<sub>7</sub> [9, 10] и T<sub>5</sub> [15, 16], показали, что для завершения работ по установлению аминокислотной последовательности белка необходимо получить перекрытие пептидов

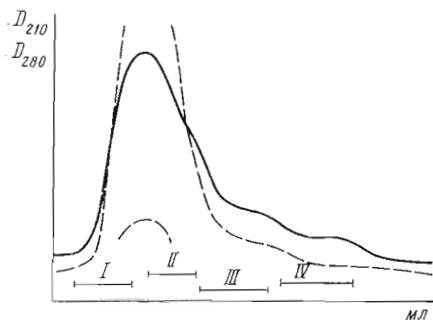


Рис. 7. Гель-фильтрация продуктов расщепления G-фактора по связям Asp-Pro на колонке (1,0×180 см) с сефадексом G-100 в буфере 20 мМ трис-HCl, pH 7,5, содержащим 6 М мочевины. Скорость элюирования 6 мл/ч

СВГ-13 и СВГ-14 бромцианового расщепления G-фактора, а также подтвердить перекрывание фрагментов T<sub>4</sub> и T<sub>5</sub> (обозначение фрагментов см. в работе [13]).

Для быстрого решения этой задачи было решено использовать расщепление G-фактора по связям Asp-Pro, так как этот метод дает возможность получить крупные фрагменты, структуру которых удобно определять автоматическим методом Эдмана. Однако для получения удовлетворительных результатов требуется значительное количество исходного материала, так как при кислотном гидролизе наблюдается неспецифическое расщепление полипептидной цепи. Недавно было предложено [17] при определении аминокислотных последовательностей автоматическим методом Эдмана использовать флуорескамин, который позволяет блокировать первичные аминогруппы, не затрагивая иминогруппу пролина. Мы решили использовать реакцию с флуорескамином для блокирования N-концевых аминогрупп продуктов неспецифического расщепления полипептидной цепи G-фактора по связям Asp-Pro. Таким образом, немодифицированными оставались только пептиды, имеющие N-концевой аминокислотный остаток пролина. После расщепления белка и обработки продуктов гидролиза флуорескамином N-концевые последовательности смеси пептидов анализировали автоматическим методом Эдмана. Было показано (рис. 6), что G-фактор содержит (как и предполагалось после исследования структуры пептидов бромцианового расщепления белка) только три связи Asp-Pro.

Наличие остатка пролина в N-концевой части пептида AP-3 позволило, используя повторную обработку флуорескамином, автоматическим методом Эдмана определить последовательность 16 аминокислотных остатков. Для исследования структуры других пептидов ограниченного кислотного гидролиза общий гидролизат G-фактора был подвергнут гель-фильтрации на колонке с сефадексом G-100 (рис. 7). Объединенные фракции I–IV анализировали автоматическим методом Эдмана. В результате анализа при использовании сравнительно малого количества G-фактора (~300 нмоль) была определена N-концевая аминокислотная последовательность пептидов AP-1 – AP-4 (таблица) и были получены важные перекрывания пептидов СВГ-13 и СВГ-14, СВГ-14 и СВГ-15, полученных при бромциановом расщеплении G-фактора, а также объединены пептиды триптического гидролиза G-концевой части пептида СВГ-12. Кроме того, было показано, что фрагменты ограниченного трипсинолиза G-фактора T<sub>4</sub> и T<sub>5</sub> [13] следуют непосредственно друг за другом. Результаты исследований, изложенные в настоящей работе, вместе с данными по первичной структуре G-фактора, полученными нами ранее (см. [1, 10–12, 14–16]), позволили завершить работу по установлению полной аминокислотной последовательности белка, полипептидная цепь которого состоит из 701 аминокислотного остатка (схема 6) и имеет M<sub>r</sub> 77321,46 [18, 19].

Аминокислотная последовательность отдельных участков белка недавно была подтверждена изучением нуклеотидной последовательности гена, кодирующего G-фактор. В частности, при изучении *Str*-оперона *E. coli* [20] была установлена нуклеотидная последовательность, соответствующая 92 аминокислотным остаткам с N-конца белка, а при выяснении нуклеотидной последовательности гена *tufA*, кодирующего фактор элонгации

Исследование N-концевой аминокислотной последовательности  
продуктов расщепления G-фактора по связям Asp-Pro

| Пептид | Аминокислотная последовательность                                                                                                   |
|--------|-------------------------------------------------------------------------------------------------------------------------------------|
| AP-1   | Ala-Arg-Thr-Thr-Pro-Ile-...<br>→ → → → → → →<br>← —————<br>T-6                                                                      |
| AP-2   | Pro-Phe-Val-Gly-Asp-Leu-Thr-Phe-Phe-Arg-Val-Tyr-<br>→ → → → → → → → → → →<br>CBG-12-T3<br>————— ←                                   |
| AP-3   | Ser-Gly-Val-...<br>→ → →<br>CBG-12-T4                                                                                               |
|        | Pro-Asp-Ala-Pro-Ile-Ile-Leu-Glu-Arg-Met-Glu-Phe-<br>→ → → → → → → → → → →<br>CBG-13<br>————— ←                                      |
|        | Pro-Glu-Pro-Val-Ile-Ser-Ile-Ala-Val-Glu-Pro-Lys-<br>→ → → → → → → → → → →<br>CBG-14<br>————— ←                                      |
|        | Thr-Lys-Ala-Asp-Gln-Glu-Lys-Met-Gly-Leu-Ala-...<br>→ → → → → → → → → →<br>CBG-14                      CBG-15                        |
| AP-4   | Pro-Ser-Phe-Arg-Val-Trp-Thr-Asp-Glu-Glu-Ser-Asn-<br>→ → → → → → → → → → →<br>T-4<br>————— ←                                         |
|        | Glu-Thr-Ile-Ile-Ala-Gly-Met-Gly-Glu-Leu-His-Leu-<br>→ → → → → → → → → → →<br>T-4<br>————— ←                                         |
|        | Asp-Ile-Ile-Val-Asp-Arg-Met-Lys-Arg-Glu-Phe-...<br>→ → → → → → → → → → →<br>T-4                                      T-5<br>————— ← |

Tu, была установлена структура участка гена, кодирующего C-концевую часть G-фактора (20 аминокислотных остатков) [21]. В обоих случаях наблюдается совпадение аминокислотных последовательностей, установленных химическим методом и выведенных на основании исследования структуры ДНК.

Аминокислотный состав G-фактора, вычисленный из аминокислотной последовательности, находится в хорошем соответствии с установленным экспериментально (с учетом поправки на точную молекулярную массу). Белок содержит пять остатков цистеина, два из которых образуют S-S-связь (положения 296 и 392 в полипептидной цепи). SH-Группа остатка цистеина в положении 113 обладает повышенной реакционной способностью, и ее модификация приводит к резкому уменьшению сродства комплекса G-фактора с ГТР к рибосоме.

Аминокислотный состав G-фактора несколько отличается от среднего состава глобулярных белков. G-Фактор содержит повышенное количество заряженных и алифатических гидрофобных аминокислотных остатков. Содержание незаряженных гидрофильных остатков несколько понижено. Количество кислых аминокислотных остатков немного превышает количество основных (15,5 и 14,5% соответственно). В этом отношении аминокислотный состав G-фактора подобен составу фактора элонгации Tu [22]. Однако фактор элонгации Tu содержит повышенное количество остатков треонина. Соотношение треонин-серин для Tu-фактора равно 2,7, для G-фактора — 1,28; среднее значение для глобулярных белков — 0,85 [23]. Фрагменты ограниченного трипсинолиза отличаются друг от друга по соотношению кислых и основных остатков. Во фрагменте T<sub>6</sub> соотношение равно 0,8, во фрагменте T<sub>7</sub> — 1, во фрагменте T<sub>4</sub> — 1,27 и во фрагменте T<sub>5</sub> — 0,74. Наибольшее количество кислых аминокислотных остатков находится во фрагменте T<sub>4</sub>, где они во многих местах группируются в кластеры, например -Glu-Tyr-Glu-Asp- (положения 202–205), -Glu-Ala-Ser-Glu-Glu-Ley-Met-Glu- (положения 227–234), -Glu-Glu-Leu-

Схема 6

Полная первичная структура фактора элонгации С

Ala-Arg-Thr-Thr-Pro-Ile-Ala-Arg-Tyr-Arg-Asn-Ile-Gly-Ile-Ser-Ala-His-Ile-Asp-Ala-10 20  
 Gly-Lys-Thr-Thr-Thr-Thr-Glu-Arg-Ile-Leu-Phe-Tyr-Thr-Gly-Val-Asn-His-Lys-Ile-Gly-40  
 Glu-Val-His-Asp-Gly-Ala-Ala-Thr-Met-Asp-Trp-Met-Glu-Gln-Glu-Gln-Glu-Arg-Gly-Ile-60  
 Thr-Ile-Thr-Ser-Ala-Ala-Thr-Thr-Ala-Phe-Trp-Ser-Gly-Met-Ala-Lys-Gln-Tyr-Glu-Pro-80  
 His-Arg-Ile-Asn<sup>2</sup>Ile-Ile-Asp-Thr-Pro-Gly-His-Val-Asp-Phe-Thr-Ile-Glu-Val-Glu-Arg-100  
 Ser-Met-Arg-Val-Leu-Asp-Gly-Ala-Val-Met-Val-Tyr-Cys-Ala-Val-Gly-Gly-Val-Gln-Pro-120  
 Gln-Ser-Glu-Thr-Val-Trp-Arg-Gln-Ala-Asn-Lys-Tyr-Lys-Val-Pro-Arg-Ile-Ala-Phe-Val-140  
 Asn-Lys-Met-Asp-Arg-Met-Gly-Ala-Asn-Phe-Leu-Lys-Val-Val-Asn-Gln-Ile-Lys-Thr-Arg-160  
 Leu-Gly-Ala-Asn-Pro-Val-Pro-Leu-Gln-Leu-Ala-Ile-Gly-Ala-Glu-Glu-His-Phe-Thr-Gly-180  
 Val-Val-Asp-Leu-Val-Lys-Met-Lys-Ala-Ile-Asn-Trp-Asn-Asp-Ala-Asp-Gln-Gly-Val-Thr-200  
 Phe-Glu-Tyr-Glu-Asp-Ile-Pro-Ala-Asp-Met-Val-Glu-Leu-Ala-Asn-Glu-Trp-His-Gln-Asn-220  
 Leu-Ile-Glu-Ser-Ala-Ala-Glu-Ala-Ser-Glu-Glu-Leu-Met-Glu-Lys-Tyr-Leu-Gly-Gly-Glu-240  
 Glu-Leu-Thr-Glu-Ala-Glu-Ile-Lys-Gly-Ala-Leu-Arg-Gln-Arg-Val-Leu-Asn-Asn-Glu-Ile-260  
 Ile-Leu-Val-Thr-Cys-Gly-Ser-Ala-Phe-Lys-Asn-Lys-Gly-Val-Gln-Ala-Met-Leu-Asp-Ala-280  
 Val-Ile-Asp-Tyr-Leu-Pro-Ser-Pro-Val-Asp-Val-Pro-Ala-Ile-Asp-Cys\*-Ile-Leu-Lys-Asp-300  
 Thr-Pro-Ala-Glu-Arg-His-Ala-Ser-Asp-Asp-Glu-Pro-Phe-Ser-Ala-Leu-Ala-Phe-Lys-Ile-320  
 Ala-Thr-Asp-Pro-Phe-Val-Gly-Asn-Leu-Thr-Phe-Phe-Arg-Val-Tyr-Ser-Gly-Val-Val-Asn-340  
 Ser-Gly-Asp-Thr-Val-Leu-Asn-Ser-Val-Lys-Ala-Ala-Arg-Glu-Arg-Phe-Gly-Arg-Ile-Val-360  
 Gln-Met-His-Ala-Asn-Lys-Arg-Glu-Glu-Ile-Lys-Glu-Val-Arg-Ala-Gly-Asp-Ile-Ala-Ala-380  
 Ala-Ile-Gly-Leu-Lys-Asp-Val-Thr-Thr-Gly-Asp-Cys\*-Leu-Cys-Asp-Pro-Asp-Ala-Pro-Ile-400  
 Ile-Leu-Glu-Arg-Met-Glu-Phe-Pro-Glu-Pro-Val-Ile-Ser-Ile-Ala-Val-Glu-Pro-Lys-Thr-420  
 Lys-Ala-Asp-Gln-Glu-Lys-Met-Gly-Leu-Ala-Leu-Gly-Arg-Leu-Ala-Lys-Glu-Asp-Pro-Ser-440  
 Phe-Arg-Val-Trp-Thr-Asp-Glu-Glu-Ser-Asn-Gln-Thr-Ile-Ile-Ala-Gly-Met-Gly-Glu-Leu-460  
 His-Leu-Asp-Ile-Ile-Val-Asp-Arg-Met-Lys-Arg-Glu-Phe-Asn-Val-Glu-Ala-Asn-Val-Gly-480  
 Lys-Pro-Gln-Val-Ala-Tyr-Arg-Glu-Thr-Ile-Arg-Gln-Lys-Val-Thr-Asp-Val-Glu-Gly-Lys-500  
 His-Ala-Lys-Gln-Ser-Gly-Gly-Arg-Gly-Gln-Tyr-Gly-His-Val-Val-Ile-Asp-Met-Tyr-Pro-520  
 Leu-Glu-Pro-Gly-Ser-Asn-Pro-Lys-Gly-Tyr-Glu-Phe-Ile-Asn-Asp-Ile-Lys-Gly-Gly-Val-540  
 Ile-Pro-Gly-Glu-Tyr-Ile-Pro-Ala-Val-Asp-Lys-Gly-Ile-Gln-Glu-Gln-Leu-Lys-Ala-Gly-560  
 Pro-Leu-Ala-Gly-Tyr-Pro-Val-Val-Asp-Met-Gly-Val-Arg-Leu-His-Phe-Gly-Ser-Tyr-Lys-580  
 Asp-Val-Asp-Ser-Ser-Glu-Leu-Ala-Phe-His-Leu-Ala-Ala-Ser-Ile-Ala-Phe-Lys-Glu-Gly-600  
 Phe-Lys-Lys-Ala-Lys-Pro-Val-Leu-Leu-Glu-Pro-Ile-Met-Lys-Val-Glu-Val-Glu-Thr-Pro-620  
 Glu-Gln-Asn-Thr-Gly-Asp-Val-Ile-Gly-Asp-Leu-Ser-Arg-Arg-Arg-Gly-Met-Leu-Lys-Gly-640  
 Gln-Gln-Ser-Glu-Val-Thr-Gly-Val-Lys-Ile-His-Ala-Gln-Val-Pro-Leu-Ser-Glu-Gln-Met-660  
 Phe-Gly-Tyr-Ala-Thr-Gln-Leu-Arg-Ser-Leu-Thr-Lys-Gly-Arg-Ala-Ser-Tyr-Thr-Met-Glu-680  
 Phe-Leu-Lys-Tyr-Asp-Glu-Ala-Pro-Ser-Asn-Val-Ala-Gln-Ala-Val-Ile-Glu-Ala-Arg-Gly-700  
 Lys

Cys\* - остатки цистеина, участвующие в образовании

S-S-связей.

-Thr-Glu-Ala-Glu- (положения 240—246), -Asp-Asp-Glu- (положения 309—311), -Glu-Glu-Ile-Lys-Glu- (положения 368—372), -Asp-Glu-Glu- (положения 446—448). Интересный кластер из кислых остатков содержится в С-концевой части фрагмента T<sub>6</sub>. Из восьми входящих во фрагмент кислых остатков четыре расположены непосредственно друг возле друга — -Asp-Trp-Met-Glu-Gln-Glu-Gln-Glu- (положения 50—57). Кроме того, основные аминокислотные остатки также образуют несколько кластеров,

например -Arg-Met-Lys-Arg- (положения 468—471), -Lys-Lys-Ala-Lys- (положения 602—605), -Arg-Arg-Arg- (положения 633—635).

Из шести остатков триптофана пять остатков расположены в N-концевой части молекулы. Остатки триптофана в положениях 51 и 71 обладают повышенной реакционной способностью и легко модифицируются под действием модифицирующих агентов [24]. Молекула белка содержит 19 остатков тирозина; 10 из них расположены в C-концевом домене. Два остатка тирозина в C-концевом домене обладают пониженным значением  $pK$  фенольного гидроксила и легко модифицируются [25]. Степень модификации как остатков триптофана, так и остатков тирозина зависит от присутствия гуаниловых нуклеотидов. Присутствие GTP или GDP блокирует модификацию остатка триптофана в положении 71 [24], а наличие нуклеотида значительно увеличивает скорость модификации остатка тирозина. С другой стороны, если G-фактор находится в составе комплекса с рибосомой и GDP, степень модификации остатков тирозина резко уменьшается [25]. Неравномерно распределены по полипептидной цепи остатки гистидина: N-концевой участок (91 аминокислотный остаток) полипептидной цепи белка содержит пять остатков гистидина, фрагменты  $T_4$  и  $T_5$  также содержат по пять остатков. Концентрирование ароматических аминокислотных остатков, в частности остатков гистидина и триптофана в N-концевой части, а остатков гистидина и тирозина — в C-концевой части полипептидной цепи белка, очевидно, связано с формированием этими частями молекулы участка взаимодействия белка с рибосомой.

### Экспериментальная часть

В работе использованы сефадексы G-15, G-25, G-100, QAE A-25 (Pharmacia, Швеция), тиол-активированная сефароза 4B (Pharmacia, Швеция), I-[ $^{14}C$ ]CH<sub>2</sub>CONH<sub>2</sub> (Amersham, Англия), катионообменная смола Chromobeads тип P (Technicon, США), стафилококковая глутаминовая протеиназа (Miles Lab., Англия), трипсин, химотрипсин (Worthington, США), бромциан (Serva, ФРГ), термолизин (Serva, ФРГ), пластинки с тонким слоем целлюлозы (Merck, ФРГ), реактивы для автоматического метода Эдмана и фенолтиогидаптоины аминокислот (Beckman, США), полиамидные пластинки (BDH, Англия),  $\beta$ -меркаптоэтанол (Serva, ФРГ). Все остальные реактивы имели квалификацию ос.ч.

*Определение N- и C-концевых аминокислот.* N-Концевые аминокислоты пептидов определяли дансильным методом с идентификацией Dns-аминокислот двумерной ТСХ на пластинках с силикагелем (5×5 см) или полиамидом (3×3 см), C-концевые — с помощью карбоксипептидаз А и В [26].

Аминокислотный состав пептидов определяли как описано в работе [12].

*Автоматический метод Эдмана.* Автоматическое определение аминокислотной последовательности пептидов бромцианового расщепления G-фактора СВГ-11 — СВГ-13 и продуктов их гидролиза проводили твердофазным автоматическим методом Эдмана на секвенаторе модели APS-240 (Rank Hilger, Англия), как описано нами ранее [12]. Анализ N-концевых последовательностей фракций, полученных при расщеплении G-фактора по связям Asp-Pro, проводили на секвенаторе модели 890 C (Beckman, США) с использованием квадрольной программы (концентрация квадрола 0,1—0,5 М). Идентификацию отщепленных фенолтиогидаптоинов аминокислот проводили методами ТСХ на пластинках Kieselgel 60 F<sub>254</sub> (Fluka, Швейцария) и высокоэффективной жидкостной хроматографией (ВЭЖХ) на колонках Ultrasphere RP-18, 5 мкм (Altex, США) или Separon RP-18, 5 мкм (ЧССР).

Аминогруппы фрагментов, полученных в результате расщепления G-фактора по связям Asp-Pro, модифицировали флуорескамином, как описано в работе [17].

*Препаративное разделение пептидов* двумерной техникой осуществляли как описано в работе [14].

Расщепление G-фактора бромцианом, а также ферментативные гидролизы выделенных пептидов CBG-11, CBG-12 и CBG-13 трипсином, химотрипсином, термолизином и глутаминовой протеиназой были проведены по методикам, описанным ранее [12]. Кислотный гидролиз пептида CBG-12-Ch-1 с выщелачиванием остатков аспарагиновой кислоты проводили как описано в работе [27].

Установление аминокислотной последовательности в выделенных пептидах ручным вариантом метода Эдмана осуществляли как в работе [28].

Модификация свободных SH-групп реагентом Эдмана (5,5'-дитиобис(2-нитробензойная кислота) и восстановление S-S-связи сульфитом натрия. К образцу G-фактора (2 мкмоль), карбоксиметилированному по экспонированной SH-группе (селективное карбоксиметилирование проводилось, как описано в [13]) в буфере 50 мМ трис-HCl, pH 8, содержащем 1 мМ EDTA и 6 М мочевины, добавили 4 мл раствора реагента Эдмана (2 мг/мл) в 10 мМ трис-HCl, pH 8. Количество выделившегося аниона нитротиибензоата определяли по приросту поглощения при 412 нм ( $\epsilon$  14 290). Избыток реагента отделяли на колонке (K 50/50, Pharmacia, Швеция) с сефадексом G-25.

Сульфитолиз проводили по методике, описанной в работе [29].

Карбоксиметилирование цистеинсодержащих пептидов бромцианового расщепления G-фактора. К смеси пептидов CBG-12 и CBG-13 (~1 мкмоль) в 50 мл буфера 0,1 М трис-HCl, pH 8,1, содержащего 1 мМ  $\beta$ -меркаптоэтанол, 1 мМ EDTA и 6 М мочевины в темной склянке, добавили 100 мг  $\text{ICH}_2\text{COONa}$  (~500 мкмоль) и 100 мКи I-[ $^{14}\text{C}$ ]CH<sub>2</sub>COOH (уд. акт. 54 мКи/ммоль); pH среды поддерживали в интервале 8,1–8,3. После инкубирования смеси в течение 30 мин при 20°С реакцию останавливали добавлением ~70 мкл  $\beta$ -меркаптоэтанола и образец обессоливали на колонке с сефадексом G-25.

Расщепление G-фактора по связям Asp-Pro. Белок (~500 нмоль G-фактора) расщепляли по методике, опубликованной ранее [30], используя 70% HCOOH. Раствор образца инкубировали при 37°С в течение 24 ч. Смесь продуктов расщепления разделяли гель-фильтрацией на колонке с сефадексом G-100.

#### ЛИТЕРАТУРА

1. Алахов Ю. Б., Бундуле М. А., Бундулис Ю. П., Винокуров Л. М., Козлов В. П. Биоорган. химия, 1983, т. 9, № 3, с. 330–342.
2. Arai N., Arai K., Nakamura S., Kaziro Y. J. Biochem., 1977, v. 82, p. 695–702.
3. Marsh R. C., Chinali G., Parmeggiani A. J. Biol. Chem., 1975, v. 250, p. 8344–8352.
4. Kaziro Y., Inoue-Yokosawa N., Kawakita M. J. Biochem., 1972, v. 72, p. 853–863.
5. Rohrbach M. S., Bodley J. W. J. Biol. Chem., 1975, v. 250, p. 6831–6836.
6. Nishizuka Y., Lipmann F. Arch. Biochem. and Biophys., 1966, v. 116, p. 344–351.
7. Kaziro Y., Inoue N., Karihi Y., Mizumoto K., Tanaka M., Kawakita M. Cold Spring Harbor Symp. Quant. Biol., 1969, v. 34, p. 385–393.
8. Girshovich A. S., Bochkareva E. S., Pozdnyakov V. A., Ovchinnikov Yu. A. FEBS Lett., 1978, v. 85, p. 283–286.
9. Alakhov Yu. B., Motuz L. P., Stengrevits O. A., Ovchinnikov Yu. A. FEBS Lett., 1978, v. 85, p. 287–290.
10. Алахов Ю. Б., Мотуз Л. П., Стенгревиц О. А., Винокуров Л. М. Биоорган. химия, 1978, т. 4, с. 1301–1313.
11. Мотуз Л. П., Бундулис Ю. П., Алахов Ю. Б. Биоорган. химия, 1979, т. 5, с. 814–826.
12. Алахов Ю. Б., Бундуле М. А., Бундулис Ю. П., Винокуров Л. М., Козлов В. П., Мотуз Л. П. Биоорган. химия, 1983, т. 9, № 3, с. 304–314.
13. Алахов Ю. Б., Мотуз Л. П., Стенгревиц О. А., Винокуров Л. М., Овчинников Ю. А. Биоорган. химия, 1977, т. 3, с. 1333–1345.
14. Алахов Ю. Б., Бундуле М. А., Бундулис Ю. П., Винокуров Л. М., Козлов В. П. Биоорган. химия, 1983, т. 9, № 3, с. 315–329.
15. Alakhov Yu. B., Dovgas N. V., Motuz L. P., Vinokurov L. M., Ovchinnikov Yu. A. FEBS Lett., 1981, v. 126, p. 183–186.
16. Алахов Ю. Б., Винокуров Л. М., Довгас Н. В., Мотуз Л. П. Биоорган. химия, 1983, т. 9, № 3, с. 293–303.
17. Bhowm A. S., Bennet J. C., Morgan P. H., Mole J. E. Anal. Biochem., 1981, v. 112, p. 158–162.
18. Алахов Ю. Б., Мотуз Л. П., Довгас Н. В., Винокуров Л. М., Бундулис Ю. П., Бундуле М. А., Козлов В. П., Овчинников Ю. А. Биоорган. химия, 1981, т. 7, с. 937–940.

19. *Ovchinnikov Yu. A., Alakhov Yu. B., Bundulis Yu. P., Bundule M. A., Dougas N. V., Kozlov V. P., Motuz L. P., Vinokurov L. M.* FEBS Lett., 1982, v. 139, p. 130-135.
20. *Post L. E., Nomura M. J.* Biot. Chem., 1980, v. 225, p. 4660-4666.
21. *Yokota T., Sugisaki H., Takunami M., Kaziro Y.* Gene, 1980, v. 12, p. 25-31.
22. *Jones M. D., Petersen T. E., Nielsen K. M., Magnusson S., Sottrup-Jensen L., Gausling K., Clark B. R. C.* Eur. J. Biochem., 1980, v. 108, p. 507-526.
23. *Dayhoff M. O., Hunt L. T.* In: Atlas of Protein Sequence and Structure /Ed. Dayhoff M. O. Washington: National Biomed. Res. Fdn., v. 5.
24. *Кашпаров И. А., Семисорнов Г. В., Алахов Ю. Б.* Биохимия, 1981, т. 46, вып. 8, с. 1488-1498.
25. *Alakhov Yu. B., Zalite I. K., Kashparov I. A.* Eur. J. Biochem., 1980, v. 105, p. 531-538.
26. *Ambler R. P.* In: Methods in Enzymology /Ed. Hirs C. H. W. New York - London: Acad. Press, 1967, v. XI, p. 156-166.
27. *Schultz J.* In: Methods in Enzymology /Ed. Hirs C. H. W. New York - London: Acad. Press, 1967, v. XI, p. 255-263.
28. *Алахов Ю. Б., Доггас Н. В., Винокуров Л. М., Вельмога И. С., Овчинников Ю. А.* Биоорг. химия, 1978, т. 4, с. 879-893.
29. *Cole R. D.* In: Methods in Enzymology /Ed. Hirs C. H. W. New York - London: Acad. Press, 1967, v. XI, p. 206-208.
30. *Landon M.* In: Methods in Enzymology /Ed. Hirs C. H. W. New York - London, 1977, v. XLVII, p. 145-149.

Поступила в редакцию  
19.VII.1982

**PRIMARY STRUCTURE OF THE ELONGATION FACTOR G FROM  
ESCHERICHIA COLI. IX. STRUCTURE OF PEPTIDES OBTAINED  
AT CYANOGEN BROMIDE CLEAVAGE OF THE G-FACTOR AND ISOLATED  
ON THIOL-ACTIVATED SEPHAROSE AS WELL AS  
OF THE PRODUCTS OF THE G-FACTOR CLEAVAGE AT Asp-Pro BONDS. THE  
COMPLETE PRIMARY STRUCTURE**

OVCHINNIKOV Yu. A., ALAKHOV Yu. B., BUNDULIS Yu. P.,  
BUNDULE M. A., VINOKUROV L. M., MOTUZ L. P., KOZLOV V. P.

*Institute of Protein Research, Academy of Sciences  
of the USSR, Pushchino*

The amino acid sequence of cysteine- and cystine-containing peptides resulting from cleavage of the G-factor by cyanogen bromide has been determined. For structure analysis cyanogen bromide peptides were further degraded using trypsin, chymotrypsin, thermolysin, staphylococcal glutamic protease, or limited acid hydrolysis. The products of the G-factor cleavage at Asp-Pro bonds were also studied. The obtained data together with those published earlier permitted to establish the complete primary structure of the elongation factor G. The polypeptide chain consists of 701 amino acid residues and has molecular mass of 77321,46.