



УДК 577.143.5

© 1990 г.

А. А. Бухаров, В. Л. Колосов, А. С. Золотарев

ФОТОСИСТЕМА II РЖИ. НУКЛЕОТИДНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ
ГЕНА *psbC*, КОДИРУЮЩЕГО 43-кДа
ХЛОРОФИЛЛ(а)-СВЯЗЫВАЮЩИЙ БЕЛОК

Институт биоорганической химии им. М. М. Шемькина АН СССР, Москва

Установлена структура участка хлоропластной ДНК ржи, содержащего ген *psbC*, который кодирует полипептидную цепь 43-кДа хлорофилл(а)-связывающей субъединицы фотосистемы II (белок СРа-2). На расстоянии 140 п. о. за стоп-кодonom гена *psbC* на противоположной цепи ДНК расположена последовательность гена *trnS*(UGA), кодирующая тРНК^{Ser}, а 5'-концевая часть гена *psbC*, как и у других видов растений, на 50 п. о. перекрывается с 3'-концевым участком гена *psbD*, который кодирует 32-кДа белок реакционного центра ФС II — D2.

Анализ аминокислотной последовательности продукта гена *psbC* обнаруживает ряд общих черт со структурой продукта гена *psbB* [1] — белка СРа-1, отражающих, очевидно, сходство функций, выполняемых этими белками, а также значительную консервативность этой последовательности в различных группах кислородвыделяющих фотосинтетиков.

Комплекс ядра фотосистемы II состоит из 7 полипептидов гидрофобной природы размером от 4 до 47 кДа [2, 3], с которыми ассоциировано около 50 молекул хлорофилла *a* [4] и ряд других простетических групп. При этом с двумя наиболее крупными субъединицами, называемыми хлорофилл-связывающими белками СРа-1 (размером 47 кДа) и СРа-2 (43 кДа), связано 20—21 и 26 молекул хлорофилла *a* соответственно [5], а также приблизительно по 3 молекулы β-каротина.

Такая насыщенность этих полипептидов молекулами пигментов обусловлена, по-видимому, тем, что их основная функция, выполняемая в ФС II, — участие наряду с Chl*a/b*-связывающими полипептидами светособирающего комплекса в поглощении квантов света и передаче энергии на пигмент реакционного центра P_{680} .

В результате ряда работ [6, 7], посвященных определению структуры хлоропластных генов *psbB* и *psbC*, кодирующих белки СРа-1 и СРа-2 соответственно, выяснилось, что оба полипептида обладают сходным аминокислотным составом и имеют общую схему строения. Основной чертой их структуры является наличие 6 участков выраженной гидрофобности, существующих, вероятно, в виде трансмембранных α-спиралей [8]. При этом доминирующим элементом структуры обоих белков является гидрофильный сегмент, разделяющий предполагаемые трансмембранные тяжи V и VI и составляющий около 40 и 30% всей длины полипептидных цепей СРа-1 и СРа-2 соответственно.

В опытах с моноклональными антителами и кросс-сшивающими реагентами [9, 10] показано, что у белка СРа-1 этот сегмент обращен в люмен и взаимодействует с 33-кДа Mn-стабилизирующим белком кислородвыделяющего комплекса. Белок СРа-2 также обнаруживает если не контакт, то по крайней мере пространственную близость к гидрофильным белкам кислородвыделяющего комплекса (их удаление из препарата ФС II резко повышает чувствительность белка СРа-2 к трипсиновому протеолизу [10]). Таким образом, экспонированные в люмен участки двух наиболее

Сокращения: ФС — фотосистема, Chl — хлорофилл, хлДНК — хлоропластная ДНК.

объемных субъединиц ФС II образуют, вероятно, необходимую поверхность для ассоциации полипептидов кислородвыделяющего комплекса с ядром.

Кроме того, предполагается, что Chl a -связывающие субъединицы обеспечивают правильную сборку и стабильность всего комплекса ядра ФС II [11]. Так, генно-инженерными методами показано, что модификация генов *psbB* [12] или *psbC* [11], нарушающая синтез нативных полипептидов CPa-1 и CPa-2, ведет к полной потере активности ФС II в мутантных клетках. Вместе с тем избирательное удаление из уже собранного комплекса ФС II, осуществленное недавно для белка CPa-2 [5], не приводит к утрате или уменьшению способности препарата к светозависимому переносу электронов между центрами Z и Q $_A$.

Таким образом, принципиальный вопрос о функциональной роли Chl a -связывающих белков в ядре ФС II все еще не вполне ясен. Остаются открытыми и вопросы, связанные с деталями структуры этих белков. В частности, неизвестно, с какими аминокислотными остатками полипептидных цепей связываются молекулы хлорофилла.

В настоящей работе, являющейся продолжением наших исследований по установлению структур полипептидов, образующих комплекс ФС II ржи [1, 13—16], мы описываем определение нуклеотидной последовательности гена *psbC* и обсуждаем выведенную из нее первичную структуру белка CPa-2 в сравнении с соответствующими структурами других организмов, а также структуру белка CPa-1.

Как и в случае с геном *psbB* [1], задача получения фрагмента хлДНК ржи, содержащего необходимую последовательность, была решена методом направленного клонирования фрагментов, гибридизующихся с синтетическими олигонуклеотидами, структура которых соответствовала участкам опубликованной последовательности генов *psbC* и *psbD* шпината [17] (ген *psbD* располагается в хлДНК всех изученных видов растений непосредственно перед *psbC*. Два гена при этом имеют общий участок длиной 50 п. о.).

При выработке стратегии клонирования использовали тот факт, что в кодирующей области гена *psbC* шпината находится единичный участок, узнаваемый рестриктазой *Hind*III (нуклеотиды 1931—1936 опубликованной последовательности [17]). Полагая, что в связи со значительной консервативностью структур хлДНК подобная ситуация характерна и для *psbC* гена ржи, его клонировали в виде двух *Eco*RI — *Hind*III-фрагментов, первый из которых — E5H1 (1,8 т. п. о.) — гибридизовался с зондом D, синтезированным на основании структуры участка гена *psbD* шпината (нуклеотиды 436—480 [17]), а второй — E5H2 (1,3 т. п. о.) — с зондом С на участок гена *psbC*, расположенный ниже *Hind*III-сайта (нуклеотиды 2034—2076 [17]).

При этом вначале в смеси рестриктных фрагментов хлДНК, обработанной *Eco*RI, методом блот-гибридизации был идентифицирован фрагмент E5 (3,1 т. п. о.) как гибридизующийся с зондами D и С и содержащий, видимо, генный кластер *psbD* — *psbC*. Вырезанный (в виде участка геля из легкоплавкой агарозы) фрагмент разрезали рестриктазой *Hind*III, предварительно освободившись от агарозы с помощью обработки расплава фенолом, и лигировали с вектором, разрезанным рестриктазами *Eco*RI и *Hind*III.

На стадии секвенирования генов выяснилось, что в гене *psbC* ржи кроме сохранившегося на том же, что и в структуре шпината, месте *Hind*III-сайта появился еще один — на 118 п. о. ближе к 5'-концу. Таким образом, для того чтобы определить структуру всей кодирующей области гена *psbC*, было предпринято клонирование еще и *Bam*HI-фрагмента B15' (3,25 т. п. о.), включающего в себя участок гена между двумя *Hind*III-сайтами.

Его клонирование осуществляли, лигируя вырезанный из геля фрагмент с соответствующим вектором.

Полученные рекомбинантные плазмиды pTSecE5H1, pTSecE5H2, pTSecB15' использовали при определении структуры гена *psbC*. При этом вставки, вырезанные из соответствующей плазмиды, элюировали

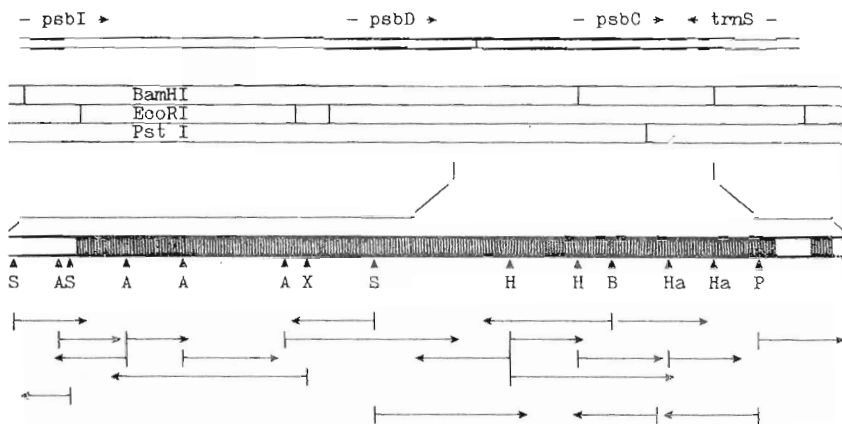


Рис. 1. Схема расщепления фрагмента хлДНК ржи, содержащего ген *psbC* и примыкающие к нему районы, различными рестриктазами и установления нуклеотидной последовательности его участка, включающего в себя гены *psbC* и *trnS*. А — *AluI*, В — *BamHI*, Вg — *BglII*, Е — *EcoRI*, Н — *HindIII*, Ha — *HaeIII*, S — *Sau3A*, X — *XbaI*

из геля, разрежали рестриктазами, узнающими четырехнуклеотидные последовательности, и клонировали в фаговые векторы серии M13mp. Стратегия определения нуклеотидной последовательности схематически представлена на рис. 1.

В показанной на рис. 2 области Е5-фрагмента хлДНК ржи располагается одна протяженная открытая рамка считывания (ОРС473), соответствующая гену *psbC*, причем первые 50 п. о. этой рамки и последовательность перед ее АТG-кодоном представляет собой 3'-концевой участок гена *psbD*, а за ее стоп-кодоном на противоположной цепи ДНК обнаруживается последовательность, гомологичная гену *trnS^{UGA}* других видов растений [18, 19].³

Кроме потенциального УСР (участок связывания рибосом), предшествующего АТG-кодону гена *psbC*, за этим кодоном расположена еще одна последовательность, комплементарная 3'-концу 16S рРНК (рис. 2), а через 9 нуклеотидов после нее находится — в той же рамке считывания, что и АТG,— GTG-кодон, который, как показано для некоторых генов *E. coli*, может служить инициатором трансляции [20]. Поскольку показано [21], что выделяемый из тилакоидов продукт гена *psbC* — СРА-2-белок — начинается с Thr⁴⁶ (результат процессинга), можно предположить, что первым транслируемым (как Met) кодоном гена *psbC* является GTG и в ходе посттрансляционной модификации удаляется не 14-членный пептид, а дипептид Met-Glu.

Между *psbC*- и *trnS*-генами на расстоянии 40 п. о. от стоп-кодона гена *psbC* располагается пара инвертированных повторов, выполняющая, вероятно, функцию терминатора транскрипции. Гомологичная структура обнаруживается и на хлДНК шпината, однако в этом случае она удалена от гена *psbC* на 140 п. о. Интересно, что при почти полном совпадении последовательностей, образующих 11-членные сегменты повтора в хлДНК двух видов (хлДНК ржи содержит единичную замену в первом сегменте, в результате которой при образовании структуры вида «петля — шпилька» пара G·C, имеющая место в случае шпината, заменяется парой G·T), гептануклеотидные последовательности, образующие «петлю», находятся в них в противоположных ориентациях (рис. 3).

Сравнение нуклеотидной последовательности гена *trnS* с соответствующей последовательностью шпината [18] обнаруживает, что все отличия двух структур локализованы в районе, кодирующем так называемую добавочную петлю тРНК, которая у ржи на 5 нуклеотидов короче. Последовательности гена *trnS* ржи и других однодольных растений — кукурузы и ячменя — идентичны [19].

TGAGTGCATTTGGCGTAGTTGGCTTGGCTCTGAACCTACGTGCCTATGACTTTGTTTCCC 60
 AGGAAATCCGTCGAGCGGAAGATCCTGAAATTTGAGACTTTCTACACCAAATAATCTTT 120
 TAAACGAGGATATTCGTGCGTGGATGGCACTCAGGATCAGCCCTATGAAATCTTATA 180
 TCCCTGAGGAGGTTCTACCACGTGGAACCTCTTTAATGGAACCTTTCGTTTATAGCTGGT 240
 CGTGACCAAAGAAACCACCGGCTTTGCTTGGTGGGCTGGGAATGCCAGACTTATCAATTTG 300
 TCCGGTAAACTACTTGGAGCTCATGTAGCCCATGCCGGAATAATCGTATTTCTGGGCCGG 360
 GCAATGAACCTATTTGAAGTGGCCCATTTCTGACCAGAAAAGCCCATGTATGAACAAGGG 420
 TTGATTTTAACTTCCACACTTAGCTACTCTAGGTGGGGAGTAGGGCCAGGGGGGAAGTT 480
 CTAGATACTTTTCCATACTTTGTATCTGGGCTACTTCACCTAATTTCCCTCCGAGTCTTA 540
 GGCTTCGGTGGCATTATACAGCCCTTCTGGGACCCGAGACTCTTGAGGAATCGTTTCCA 600
 TTCTTTGGTTATGTCTGGAAGATCGAAAATAAAATGACTACAATTTTGGGTATTCACCTA 660
 ATTTTGTAAAGTCTAGCTGGCTTTTCTTCTAGTACTCAAGGCTCTTTATTTTGGCGGTGA 720
 TATGATACCTGGGCCCTGGGGGGGAGATGTAAGAAAAATTAACCAATTTGACCCCTAGT 780
 CCCAGTGTATATTTGCTTATTTACTAAAATCTCCTTTTGGTGGAGAAGGGTGCATTGTT 840
 AGTGTAGATGATTTAGAAGATATAATGGTGGACATGTAAGTTGGGTTTATTTGTGTA 900
 TTTGGCGGAATTTGGCATATTTTAAACCAAACCCCTCGCATGGGCTCGCCGTCATTGTA 960
 TGGTCTGGAAGAAGCTTACTTGTCTTATAGTTTACCTGCTTATCTGTCTTTGGTTTTATC 1020
 GCTGTGTGTTTTGTATGGTTCAATAATACAGCTTATTCGAGTGAAGTTTTATGGACCCACC 1080
 GGGCCAGAAGCTTCTCAAGCTCAAGCATTTACTTTTCTAGTTAGAGACCAGCCCTCTGG 1140
 GCTAATGTGGATCCGCTCAAGGACCCACAGGTTTAGGTAATAATCTAAATGCGTTCCCA 1200
 ACTGGGGAAGTTATCTTTGGAGGGGAAACTATGCGTTTTTGGGACCTTCCTGCTCCATGG 1260
 TTAGAACCTCTAAGGGGCCCAACCGTTTTGGACTTGAGTGGTTGAAAAAGACATACAA 1320
 CCTTGGCAAAGAACGACGCTCAGCAGAATATATGACCCACGCTCCTTTAGGCTCTTAAAT 1380
 TCCGTGGGTGGCTAGCTACCGAGATCAATGCAGTTAATTAATGCTCTCTCTAGAAAGTTGG 1440
 TTATCTACTTCTCATTTTGTCTAGGATCTCTCCCTTTTGGGCCATTTATGGCATGCA 1500
 GGAAGAGCCCGAGCTGCTGCAGCAGGTTTTGAAAAGGGAATCGATCGTGATTTGGAACCT 1560
 GTTCTTTACATGAACCCCTCTTAACTAAGATTTCTTATTTATACCTGTTCTACTTTTTTT 1620
 CTGTTCTGGCTCCGTTATTTCCATCTAGCCGACCCATTCATTCCTTTTTTAAGAAAGAAAG 1680
 TAACCGACAGAAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1740
 GGAGAGAGAGGGGATTCGGAACCCCTCGATAGTTCCCTAGAACTATACCGGTTTTCAGACCG 1800
 CAGCTATCAACCACTCAGCCATCTGTCCACAGCCTAATCTTATTTTACTCC 1852

Рис. 2. Нуклеотидная последовательность фрагмента хлДНК ржи, содержащего гены *psbC* (166—1584), *trnS* (1828—1740), а также 3'-концевой участок гена *psbD*. Жирным прифтом выделены инициирующие и стоп-кодоны, а также антикодон гена *trnS*. Подчеркнуты потенциальные участки связывания рибосом и последовательность гена *trnS*. Встречными стрелками отмечены инвертированные повторы

На рис. 4 показана аминокислотная последовательность продукта гена *psbC* ржи в сравнении с соответствующими последовательностями других организмов. Поскольку вопрос о первичном продукте трансляции гена не решен, мы приводим структуру продукта, соответствующего максимально длинной рамке считывания. Однако дальнейшее обсуждение деталей строения белка СРа-2 относится к его зрелой полипептидной цепи длиной 459 а. о.

17	G R L L A V H I M H T A L V S G W A G S M A L Y E L A	CPa-1
	: : : : :	
	G K L L G A H V A H A G L I V F W A G A M N L F E V A	CPa-2
47		
446	S P R G W F T F G H A T F A L L F F F G H I W H G A R	CPa-1
	: : : : : : : : :	
	S P R S W L S T S H F V L G F F L F V G H L W H A G R	CPa-2
421		

Рис. 5. Сравнение структур двух участков полипептидных цепей CPa-1 и CPa-2 белков ржи, включающих в себя предполагаемые трансмембранные сегменты I и VI (отмечены линиями). Штрихами соединены идентичные остатки, двоеточием — консервативные замены. Цифрами указаны позиции аминокислотных остатков в соответствующих полипептидных цепях

Экспериментальная часть

хлДНК выделяли как описано ранее [1]. Рестрикционный анализ вели согласно работе [23], используя электрофорез в 0,5% агарозном геле. Двухнаправленный блот-перенос и гибридизацию нитроцеллюлозных блот-реплик с олигонуклеотидными зондами проводили как описано [24].

Плазмидную ДНК выделяли по методам [25]. Получение компетентных клеток и гибридизационный отбор необходимых рекомбинантных клонов осуществляли по методикам, описанным в работе [26]. Лигирование фрагментов ДНК в расплаве легкоплавкой агарозы проводили согласно сообщению [27].

Векторами для клонирования служили плаزمиды рTZ19R и репликативная форма ДНК фагов M13mp18, M13mp19, обработанные соответствующими рестриктазами. Дефосфорилирование концов векторов не проводилось.

Переклонирование фрагментов ДНК в фаговые векторы M13mp18, M13mp19 и наработку однонитевой матрицы осуществляли по методикам [28, 29]. Крупные фрагменты клонировали, вырезая их из геля и лигируя с вектором в расплаве, мелкие клонировали в виде соответствующей рестрикционной смеси с последующим гибридизационным анализом полученных рекомбинантных клонов и Т-скринингом [28]. В клонировании и при наработке однонитевой ДНК-матрицы использовали клетки JM109 (Rec⁻) [30].

Секвенирование вели по методу Сэнгера [31] с модификациями [32]. Олигонуклеотидные зонды синтезировали фосфорамидитным методом на синтезаторе Applied Biosystems.

Авторы выражают благодарность Н. С. Быстрову за синтез олигонуклеотидных зондов.

СПИСОК ЛИТЕРАТУРЫ

1. Бухаров А. А., Колосов В. Л., Золотарев А. С., Абдулаев Н. Г. // Биоорганич. химия. 1989. V. 15. № 7. P. 927—939.
2. Ikeuchi M., Inoue Y. // FEBS Lett. 1988. V. 241. № 1/2. P. 99—104.
3. Webber A. N., Packman L., Chapman D. J., Barber J., Gray J. C. // FEBS Lett. 1989. V. 242. № 2. P. 259—262.
4. Zuber H. // Photochem. Photobiol. 1985. V. 42. № 6. P. 821—844.
5. Yamaguchi N., Takahashi Y., Satoh K. // Plant Cell. Physiol. 1988. V. 29. № 1. P. 123—131.
6. Morris J., Herrmann R. G. // Nucl. Acid Res. 1984. V. 12. № 6. P. 2837—2850.
7. Alt J., Morris J., Westhoff P., Herrmann G. // Curr. Genet. 1984. V. 8. № 8. P. 597—606.

8. *Chisholm D., Williams J. G. K.* // Plant Mol. Biol. 1988. V. 10. № 4. P. 293—301.
9. *Bricker T. M., Frankel L. K.* // Arch. Biochem. and Biophys. 1987. V. 256. № 1. P. 295—301.
10. *Isogai Y., Yamamoto Y., Nishimura M.* // FEBS Lett. 1985. V. 187. № 2. P. 240—244.
11. *Dzelzkalns V. A., Bogorad L.* // EMBO J. 1988. V. 7. № 2. P. 333—338.
12. *Vermaas W. F. J., Williams J. G. K., Rutherford A. W., Mathis P., Arntzen C. J.* // Proc. Nat. Acad. Sci. USA. 1986. V. 83. № 24. P. 9474—9477.
13. *Bukharov A. A., Kolosov V. L., Zolotarev A. S.* // Nucl. Acids Res. 1988. V. 16. № 17. P. 8737.
14. *Bukharov A. A., Kolosov V. L., Klezovich O. N., Zolotarev A. S.* // Nucl. Acids Res. 1989. V. 17. № 2. P. 798.
15. *Kolosov V. L., Bukharov A. A., Zolotarev A. S.* // Nucl. Acids Res. 1989. V. 17. № 4. P. 1759.
16. *Kolosov V. L., Klezovich O. N., Zolotarev A. S.* // Nucl. Acids Res. 1989. V. 17. № 4. P. 1760.
17. *Alt J., Morris J., Westhoff P., Herrmann G.* // Curr. Genet. 1984. V. 8. № 8. P. 597—606.
18. *Holschuh K., Bottomley W., Whitfield P. R.* // Nucl. Acids Res. 1984. V. 12. № 23. P. 8819—8834.
19. *Rasmussen O. F., Bookjans G., Stummann B. M., Henningsen K. W.* // Plant Mol. Biol. 1984. V. 3. P. 191—199.
20. *Steinmuller K., Ley A. C., Steinmetz A. A., Sayre R. T., Bogorad L.* // Mol. Gen. Genet. 1988. V. 216. № 1. P. 60—69.
21. *Michel H., Hunt D. F., Shabanowitz J., Bennet J.* // J. Biol. Chem. 1988. V. 263. № 3. P. 1123—1130.
22. *Reverdatto S. V., Andreeva A. V., Buryakova A. A., Chakhmakhcheva O. G., Efimov V. A.* // Nucl. Acids Res. 1989. V. 17. № 10. P. 3996.
23. *Fuchs R., Blakesley R.* // Meth. Enzymol. 1983. V. 100. P. 3—38.
24. *Meinkoth J., Wahl G.* // Anal. Biochem. 1984. V. 138. № 2. P. 267—284.
25. *Monstein H.-J., Geijer T.* // Biochem. Inter. 1986. V. 12. № 6. P. 889—896.
26. *Маниатис Т., Фрич Э., Сэмбрук Дж.* Методы генетической инженерии: молекулярное клонирование. М.: Мир, 1984.
27. *Murray J. A. H.* // Nucl. Acids Res. 1986. V. 14. № 24. P. 10118.
28. *Messing J.* // Meth. Enzymol. 1983. V. 101. P. 20—78.
29. *Eperon I. C.* // Anal. Biochem. 1986. V. 156. № 2. P. 406—412.
30. *Yanisch-Perron C., Vieira J., Messing J.* // Gene. 1985. V. 33. № 1. P. 103—119.
31. *Sanger F., Nicklen S., Coulson A. R.* // Proc. Nat. Acad. Sci. USA. 1977. V. 74. № 12. P. 5463—5467.
32. *McGraw III R. A.* // Anal. Biochem. 1984. V. 143. № 2. P. 298—303.

Поступила в редакцию
31.X.1989

A. A. BUKHAROV, V. L. KOLOSOV, A. S. ZOLOTAREV

**RYE PHOTOSYSTEM II: NUCLEOTIDE SEQUENCE OF THE *psbC* GENE
ENCODING 43-kDa CHLOROPHYLL(*a*)-BINDING PROTEIN**

*M. M. Shemyakin Institute of Bioorganic Chemistry, Academy of Sciences
of the USSR, Moscow*

The structure of the rye chloroplast DNA, which contains *psbC* gene coding for 43-kDa chlorophyll(*a*)-binding subunit of photosystem II, is determined. The sequence of *trnS* (UGA) gene encoding tRNA Ser is located at a distance of 140 bp downstream from the stop codon of *psbC* gene on the opposite DNA strand. The 5'-terminal part of *psbC* gene, like in other plants, overlaps by 50 bp the 3'-terminal region of *psbD* gene coding for D2 protein of photosystem II. The amino acid sequence of the *psbC* gene product reveals common features with the structure of the *psbB* gene product (CPa-1 protein). The structural similarity of these two proteins seems to reflect their similar functions.