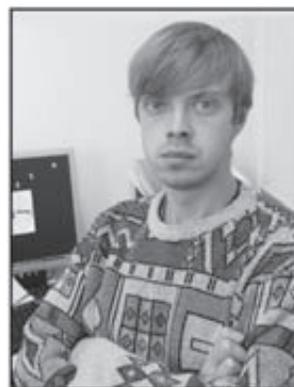




Памирский Игорь Эдуардович – к.б.н., ассистент кафедры биологической химии, Амурская государственная медицинская академия, г. Благовещенск



Голохваст Кирилл Сергеевич – к.б.н., зам. директора Института нефти и газа, Дальневосточный государственный технический университет, г. Владивосток

УДК 577. 21; 577. 3

О РОДСТВЕ ФРУСТУЛИНОВ ДИАТОМЕЙ И БЕЛКОВ ДРУГИХ ФОРМ ЖИЗНИ

ON THE RELATIONSHIP OF DIATOM FRUSTULINS AND PROTEINS OF OTHER LIFE FORMS

Данная статья является продолжением нашего исследования белков биоминерализации, обнаруженных в диатомеях, губках, магнитотаксических бактериях, а также поиска их гомологов среди белков растений, животных и других форм жизни. В частности, на уровне целых последовательностей и их отдельных участков *in silico* показана различная степень гомологии фрустулинов (*Cylindrotheca fusiformis*, *Phaeodactylum tricornutum* CCAP 1055 / 1, *Navicula pelliculosa*, *Thalassiosira pseudonana*), представляющих семейство кальций-связывающих гликопротеинов, участвующих в формировании видоспецифичного биокремнеземного панциря диатомей, с белками вирусов, бактерий, простейших многоклеточных, червей, насекомых, устриц, ланцетников, рыб, млекопитающих, в том числе человека, а также грибов и растений.

Ключевые слова: белки, фрустулины, биоминерализация

This article is a continuation of our study of biomineralization proteins found in diatoms, sponges, magnetotactic bacteria, as well as finding their homologues among proteins of plants, animals and other life forms. In particular at the level of entire sequences, and their individual sites *in silico* shown a different degree of homology of frustuls (*Cylindrotheca fusiformis*, *Phaeodactylum tricornutum* CCAP 1055 / 1, *Navicula pelliculosa*, *Thalassiosira pseudonana*), representing the family of calcium-binding glycoprotein involved in the formation of species-specific biosilica shell of diatoms with proteins of viruses, bacteria, multicellular organisms, worms, insects, oysters, amphioxus, fish, mammals, including man, as well as fungi and plants.

Key words: proteins, frustuls, biomineralization

Диатомовые водоросли имеют минеральный панцирь из аморфного диоксида кремния (SiO_2). Исследования [5, 6, 10] показали, что в формировании видоспецифичного кремниевого панциря диатомовых участвует ряд белков. К ним относятся силаффины, фрустулины, плевралины, транспортеры кремния. Силаффины, плевралины и фрустулины синтезируются в виде неактивных предшественников. Фрустулины, обнаруженные в клеточных стенках веретенообразных бактерий *S. Fusiformis*, были описаны как группа кальций-связывающих гликопротеинов, содержащая характерные кислые области (включающие около 50 аминокислотных остатков), богатые цистеином [5, 6, 8]. Иммунологические исследования показали, что фрустулины локализируются повсеместно на внешней стороне клеточной стенки диатомовых [7], хотя не исключено, что отдельные фрустулины имеют конкретные профили локализации [10].

Есть мнение, что фрустулины участвуют в образовании внешней оболочки клеточной стенки напрямую, не влияя на отложение кремния [9]. Однако установлено, что разные виды диатомей используют разные силаффины и фрустулины, которые в условиях *in vitro* определяют видоспецифические отличия осаждения кремния [11]. Также ранее было показано, что ϵ -фрустулин из диатомовой водоросли *Navicula pelliculosa*, найденный и в *Cylindrotheca fusiformis*, локализован в клеточной стенке и его синтез выражен во время строительства клеточной стенки [4].

На сегодняшний день описано 5 типов фрустулинов: α -фрустулин (первое название P75 K) (75 кДа), β -фрустулин (105 кДа), γ -фрустулин (200 кДа), δ -фрустулин (35 кДа) и ϵ -фрустулин (140 кДа) [1, 2, 5, 6, 8, 10]. Все они содержат характерные богатые цистеином области (ACR domains). Функция этого домена пока неизвестна. Для α -фрустулина выделено три изоформы, обозначаемые арабскими цифрами ($\alpha 1$, $\alpha 2$, $\alpha 3$) [3].

Исследование гомологии подобных фрустулинам белков позволит провести параллели в механизмах биоминерализации у разных видов организмов и установить филогенетические связи между ними. Наличие подобных белков у других организмов, в том числе у человека, может указывать на схожие биохимические аспекты механизмов биоминерализации, на происхождение данных белков от одной эволюционной группы, имеющих тесную связь с минералообразованием, а также более углублено понять взаимодействие живого и неживого.

Материалы и методы

Аминокислотные последовательности фрустулинов были взяты из базы данных сервера вычислительной биологии UniProt (<http://www.uniprot.org>). Исследование гомологии последовательностей белков проводилось при помощи этого же сервера (на момент обращения к серверу в базе хранилась информация о более чем 11 млн последовательностей). Для изучения фрустулинов внутри семейства использовали режим множественного выравнивания последовательностей «Align», реализованный на базе ClusalW 2.0.12 (параметры на сервере не настраиваются). Поиск белков-гомологов осуществляли при помощи инструмента «BLAST» (реализован на базе NCBI BLASTP 2.2.22 Sep-27-2009), также предусмотренного на сервере с выставлением следующих параметров: Database – UniProtKB (база данных, по кото-

рым производится поиск – UniProtKB или блоки последовательностей с 100 %, 90 % или 50 % идентичностью), Threshold – 0,0001, Matrix – Auto / BLOSUM-62 (матрица оценки вероятности для каждой позиции выравнивания), Filtering – None (фильтр регионов с низким уровнем сложности), Gapped – yes (введение пробелов при выравнивании), Hits – 250 (количество сравниваемых последовательностей). В инструкции по работе с сервером под термином «Threshold» понимается среднее значение порога статистической меры числа ожидаемых совпадений в случайных данных, обозначаемое E. При значениях E, лежащих в пределах 0,0001–0,1, данные принято считать достоверными, при значениях от 0,1 до 10, как правило, данные сомнительные, более 10 – результаты вряд ли будут иметь биологическое значение. При любом выставленном значении E все полученные результаты следует проверять вручную, т. е. они не являются истинными.

Результаты и обсуждение

На сервере UniProt размещены данные о белковых компонентах 11 фрустулинов (строение и состав углеводных компонентов не приводится). На сегодняшний день последовательности белков P75K ($\alpha 1$ -фрустулина), Alpha 2 frustulin ($\alpha 2$ -фрустулина), выделенных из *Cylindrotheca fusiformis*, Frustulin 2 и 5, выделенных из *Phaeodactylum tricorutum* CCAP 1055 / 1, Epsilon frustilin (ϵ -фрустулина), выделенных из *Navicula pelliculosa*, расшифрованы полностью. Остальные Alpha 3 frustulin ($\alpha 3$ -фрустулин) из *Cylindrotheca fusiformis*, Frustulin 3, Frustulin 5_1 и 5_2, выделенные из *Phaeodactylum tricorutum* CCAP 1055 / 1, Frustulin и неохарактеризованные белки FRU1 и FRU3 из *Thalassiosira pseudonana* – только частично.

Гомология внутри семейства фрустулинов (частично расшифрованные последовательности также подвергались сравнению) лежит в пределах 36–94 %. Ниже приводятся примеры результатов сравнения аминокислотных последовательностей с максимальной степенью гомологии из ближайших 250 гомологов для каждого фрустулина.

Белок $\alpha 1$ -фрустулин (P75K protein; ID Q39494; 442 Ам; 46,166 Да; название гена – P75K; *Cylindrotheca fusiformis*) гомологичен фрагментам белков ланцетника *Branchiostoma floridae* (C3Y2D8, C3YBS2) на 29–34 %, анемоны *Nematostella vectensis* (A7SXP8) на 40 %, дрозофил *Drosophila pseudoobscura pseudoobscura* (B5E0E8) и *Drosophila sechellia* (B4ILE7) до 51 % и 52 %, слизевика *Dictyostelium discoideum* (Q54D31) до 54 %, кукурузы *Zea mays* (B4FNV9) до 32 %, фрагментами белка клеомы *Cleome spinosa* (Q1KUM6) до 43 %.

Первичная структура $\alpha 2$ -фрустулина (Alpha 2 frustulin; ID Q39495; 438 Ам; 45,947 Да; название гена – fru alpha 2; *Cylindrotheca fusiformis*) имеет сходство с такой структурой белков и их фрагментов следующих животных и растений: анемона *Nematostella vectensis* (A7SXP8) на 39 %, слизевик *Dictyostelium discoideum* (Q54D31, Q54WQ8) на 29 % и 32 %, ланцетник *Branchiostoma floridae* (C3Y2D8, C3YBS2) на 27 % и 29 %, дрозофилы *Drosophila pseudoobscura pseudoobscura* (B5E0E8) и *Drosophila sechellia* (B4ILE7) на 34 % и 52 %, кукуруза *Zea mays* (B4FNV9) и клеома *Cleome spinosa* (Q1KUM6) на 30 %.

Фрагмент α 3-фрустулина (Alpha 3 frustulin; ID Q39496; 463 Ам; 48,532 Да; название гена – fru alpha 3; *Cylindrotheca fusiformis*) имеет гомологию с фрагментами белков анемоны *Nematostella vectensis* (A7SXP8) до 24%, нематоды *Haemonchus contortus* (Q9Y114) до 55%, африканского москита *Anopheles gambiae* (Q7PZX4) до 39%, бабочки *Spodoptera frugiperda* (Q6PST6) до 27%, дрозофилы *Drosophila pseudoobscura pseudoobscura* (Q299Y6) до 30%, жука *Tribolium castaneum* (D1MAJ8) до 28%, ланцетника *Branchiostoma floridae* (C3Y2D7) до 25%, быка *Bos Taurus* (A6QPB8) до 20%, макаки *Macaca mulatta* (B8XNP4) до 26%, пшеницы *Triticum aestivum* (P93611) и кукурузы *Zea mays* (B6TEB0) до 27%.

Аминокислотная последовательность ϵ -фрустулина (Epsilon frustilin; Q39497; 354 Ам; 37,312 Да; название гена – fru epsilon; *Navicula pelliculosa*) подобна фрагментам белков следующих животных: слизевик *Dictyostelium discoideum* (Q54M35) до 34%, анемона *Nematostella vectensis* до 34%, нематоды *Brugia malayi* (A8NMH4) и *Caenorhabditis briggsae* (A8X388) до 38% и 41%, устрица *Crassostrea nippona* (B3ITC3) до 29%, дрозофилы *Drosophila grimshawi* (B4JLR6) до 39%, рыба *Danio rerio* (Q803K3) до 41%, собака *Canis familiaris* (O18740) до 31%, мышь *Mus musculus* (P18165) до 33%, крыса *Rattus norvegicus* (B2RZ19) до 30%, хомяк *Mesocricetus auratus* (O09029) до 35%, бык *Bos taurus* (A2VE23) до 31%, человек *Homo sapiens* (C9J285, P23490, Q6FHY3) до 35%, 29% и 31%. Гомологичен белкам растений: резуховидка *Arabidopsis thaliana* (Q9LSN6) до 41%, тополь *Populus trichocarpa* (B9HL62) до 43%, клещевина *Ricinus communis* (B9T304, B9T4Y1) до 33% и 38%, томат *Solanum lycopersicum* (Q43522) до 37%, кукуруза *Zea mays* (B4FQQ2) до 41%, виноград *Vitis vinifera* (A5BG48) до 31%, рис *Oryza sativa subsp. Indica* (A2ZJC9) до 43%, сорго *Sorghum bicolor* (C5WZJ5) до 36%, фасоль *Phaseolus vulgaris* (P10495) до 34%, люцерна *Medicago truncatula* (A9YWR4) до 33%. Преобладают белки растений и высших животных.

Фрустулин 2 (Frustulin 2; ID B7G5U3; 482 Ам; 52,269 Да; название гена – fru2; *Phaeodactylum tricornutum* CCAP 1055 / 1) гомологичен фрагментам белка слизевика *Dictyostelium discoideum* (Q54WQ8) до 34%, актинии *Nematostella vectensis* (A7SXP8) до 43%, фруктовых мушек *Drosophila ananassae* (B3LZN4), *Drosophila willistoni* (B4N5R9) и *Drosophila willistoni* (B4N168) до 40%, 41% и 46%, бабочки *Trichoplusia ni* (O18510) до 28%, комара *Anopheles gambiae* (Q5TQW2) до 42%, жука *Tribolium castaneum* (D2A0A9) до 38%, ланцетника *Branchiostoma floridae* (C3Z6C6, C3YY23) до 44–45%, лягушки *Xenopus laevis* (P17437) до 26%, мыши *Mus musculus* (A2AQH4, Q5H9F3, Q5H9F3-3) до 33%, панды *Ailuropoda melanoleuca* (D2H9V1) до 32%, человека *Homo sapiens* (B7ZMM8, Q8N7U7, D2CFI5, P49918, P49918-2, C9J6V8) от 27 до 40%, кукурузы *Zea mays* (Q41805, B6TRK7) до 28% и 35%, риса *Oryza sativa subsp. Japonica* (Q5VMY9) до 32%, сорго *Sorghum bicolor* (C5Y7E8) до 28%, пырея *Lophopyrum elongatum* (Q4KXE1) до 28%, клеомы *Cleome spinosa* (Q1KUM6) до 28%, гриба *Magnaporthe grisea* (A4QSG6) до 31%.

Фрустулин 5 (Frustulin 5; ID B7S3M3; 280 Ам; 28,364 Да; название гена – fru5; *Phaeodactylum tricornutum* CCAP 1055 / 1) гомологичен фрагментам белков слизевика *Dictyostelium discoideum* (Q54WQ8) до 40%, анемоны *Nematostella vectensis* (A7T0K7) до 52%, нематод *Caenorhabditis briggsae* (A8XLQ2) и *Caenorhabditis elegans* (Q95XX2) до 35% и 42%, бабочки *Heliothis armigera* (B6CME3) до 35%, гусеницы *Mamestra configurata* (Q95U94) до 27%, дрозофил *Drosophila willistoni* (B4MHV2)

и *Drosophila virilis* (B4LHZ6) до 41–43 %, комара *Culex quinquefasciatus* (B0WB45) до 35 %, ланцетника *Branchiostoma floridae* (C3ZSQ0, C3ZAZ6) до 39–41 %, мыши *Mus musculus* (Q3TLH4, Q3TLH4-5, Q3TLH4-3, Q3TLH4-2) до 32 %, крысы *Rattus norvegicus* (P62521) до 40 %, овцы домашней *Ovis aries* (Q1WD20) до 33 %, быков *Bos taurus* (Q8WML4, Q2V0E7, Q2V0E6, B2MW78) и *Bos indicus* (B2MW75) до 30 %, панды *Ailuropoda melanoleuca* (D2I5T5) до 31 %, человека *Homo sapiens* (P98088, Q14888, Q9Y520, Q9Y520-7 и др.) до 33–40 %, огурца посевного *Cucumis sativus* (Q4VZL0) до 28 %, гриба *Aspergillus fumigatus* (Q873Y0, Q4WEP7) до 43–44 %.

Фрагмент фрустулина 5_1 (Frustulin 5_1; ID B7FXR2; 521 Ам; 59,243 Да; название гена – fru5_1; *Phaeodactylum tricornerutum* CCAP 1055 / 1) подобен белкам простейших *Leishmania braziliensis* (A4HA37) на 55 %, амёбы *Naegleria gruberi* (D2UYF5) на 50 %, слизевика *Dictyostelium discoideum* (Q8T6B7) на 52 %, трихоплакса *Trichoplax adhaerens* (B3RQW3) на 54 %, актинии *Nematostella vectensis* (A7SFE5) на 55 %, шелкопряда *Bombyx mori* (Q1HQD2) на 55 %, нематоды *Caenorhabditis elegans* (Q9XTD9) на 52 %, трематоды *Schistosoma mansoni* (C4Q844) на 53 %, клеща *Ixodes scapularis* (B7Q3I4) на 52 %, дрозофилы *Drosophila pseudoobscura pseudoobscura* (Q29IA7) на 54 %, комара *Aedes aegypti* (Q16RE9) на 53 %, ланцетника *Branchiostoma floridae* (C3Z802) на 41 %, данио *Danio rerio* (Q802V8) на 55 %, лосося *Salmo salar* (B5X0Y6, C0PU04) на 55–56 %, лягушки *Xenopus laevis* (Q7ZWW5) на 55 %, мыши *Mus musculus* (Q99LE6) на 55 %, крысы *Rattus norvegicus* (A2VD14) на 55 %, курицы *Gallus gallus* (Q5ZM62) на 55 %, быка *Bos taurus* (Q2KJA2) на 55 %, свиньи *Sus scrofa* (Q767L0) на 41 %, панды *Ailuropoda melanoleuca* (D2H1V2) на 55 %, макаки *Macaca mulatta* (Q5TM60) на 41 %, шимпанзе *Pan troglodytes* (Q7YR37) на 41 %, человека *Homo sapiens* (Q9UG63, Q75MJ1, Q75MJ0 и другие) до 55 %. Преимущественно это АТФ-связывающие транспортные белки. Также гомологичен белкам риса *Oryza sativa subsp. Japonica* (Q84ZB2) и *Oryza sativa subsp. indica* (A2Z4A5) на 57–60 %, сорго *Sorghum bicolor* (C5YMS9) на 57 %, кукурузы *Zea mays* (C0P4H6) на 57 %, претиса *Pteris vittata* (Q2TUN6) на 57 %, ели *Picea sitchensis* (A9NVC5) на 56 %, клещевины *Ricinus communis* (B9SCH6) на 57 %, тополя *Populus trichocarpa* (B9N856) на 55 %, резуховидки *Arabidopsis thaliana* (Q9FJH6) на 56 %, винограда *Vitis vinifera* (A5BPR3) на 42 %, мха *Physcomitrella patens subsp. Patens* (A9RU54) на 57 %, дрожжей *Schizosaccharomyces pombe* (O42943) на 48 %, гриба *Aspergillus fumigatus* (Q4WU32) на 48 %. Преобладают белки грибов (гомология 40–60 %).

Фрагмент фрустулина 5_2 (Frustulin 5_2; ID B7FZJ8; 283 Ам; 28,265 Да; название гена – fru5_2; *Phaeodactylum tricornerutum* CCAP 1055 / 1) гомологичен фрагментам белков амёбы *Naegleria gruberi* (D2W6L6) до 40 %, слизевика *Dictyostelium discoideum* (O54LY1, Q54WQ8, Q86I54, Q55AJ1) до 37–44 %, совок *Mamestra configurata* (Q95U94) и *Helicoverpa armigera* (B6CME3) до 29 %, дрозофил *Drosophila melanogaster* (A9YKM8) и *Drosophila pseudoobscura pseudoobscura* (B5DNV2) до 34 % и 42 %, крысы *Rattus norvegicus* (Q63549) до 37 %, быков *Bos mutus grunniens* (C1KRN6, Q0Z9W5 и др.), *Bos indicus* (B2MW75) и *Bos taurus* (Q2V0E6, B2MW77, Q8SQ37 и др.) до 29–39 %, человека *Homo sapiens* (P98088, C9JMC6 и др.) на 40 % и 50 %, растений кардаминописса *Cardaminopsis arenosa* (Q6V5H0) на 26 %, резуховидки *Arabidopsis thaliana* (Q6IMG0) до 44 %, клещевины *Ricinus communis* (B9S9T1) до 42 %, кукурузы *Zea mays* (B6TRK7, B4FBH5) до 36 %, гриба *Aspergillus clavatus* (A1C839) до 40 %.

Фрагмент фрустулина 3 (Frustulin 3; ID B7FWP1; 386 Ам; 39,378 Да; название гена – Fru3; *Phaeodactylum tricornutum* CCAP 1055 / 1) гомологичен фрагментам белков амебы *Naegleria gruberi* (D2VYR3) до 24%, нематоды *Caenorhabditis briggsae* (A8WXW9) до 24%, дрозофилы *Drosophila mojavensis* (B4KQE7) до 65%, дрозофилы *Drosophila yakuba* (B4P819) до 57%, комара *Aedes aegypti* (Q16KQ9) до 24%, бабочки *Papilio xuthus* (B2DBL6) до 68%, хрущака *Tribolium castaneum* (D2A0F6) до 41%, устрицы *Crassostrea nippona* (B3ITC3) на 21%, ланцетника *Branchiostoma floridae* (C3Z4A7) до 25%, рыбы *Tetraodon nigroviridis* (Q4T7H, Q4SNA1) до 25% и 56%, лягушки *Xenopus laevis* (P17437) до 24%, мыши *Mus musculus* (O35730) до 23%, крысы *Rattus norvegicus* (Q6MGB6) до 24%, человека *Homo sapiens* (Q6ZMY3, Q6ZMY3-5, Q6ZMY3-2 и др.) до 25%, риса *Oryza sativa subsp. Japonica* (Q6K8Z3) до 26%, винограда *Vitis vinifera* (A5AV41, A5ACV1, A5BT12 и др.) до 38–63%, кукурузы *Zea mays* (B6TSC7) до 31%.

Фрагмент белка FRU1 (Putative uncharacterized protein FRU1; ID B8C1P1; 397 Ам; 40,714 Да; название гена – FRU1; *Thalassiosira pseudonana*) гомологичен фрагментам белков амебы *Naegleria gruberi* (D2W009) до 25%, слизевика *Dictyostelium discoideum* (Q75JE6) до 26%, трихоплакса *Trichoplax adhaerens* (B3RIJ8) до 23%, анемоны *Nematostella vectensis* (A7RL39) до 22%, нематоды *Caenorhabditis elegans* (Q22886) до 24%, дрозофилы *Drosophila yakuba* (B4NYY8, B4PYX7) до 23–25% и *Drosophila virilis* (B4MA01) на 35%, комара *Aedes aegypti* (Q16KQ8) до 21%, гусеницы *Spodoptera exigua* (A8HG49) до 23%, шелкопряда *Bombyx mori* (Q2F5X6) до 29%, ланцетника *Branchiostoma floridae* (C3YCB7, C3ZT12) до 24–26%, рыбы *Tetraodon nigroviridis* (Q4SDG6) до 28%, змеи *Philodryas olfersii* (A7X4K1) до 28%, мыши *Mus musculus* (Q9D7Q2, Q6PAI5) до 37%, свиньи *Sus scrofa* (A5A8W4) до 27%, быка *Bos taurus* (O18977) до 24%, человека *Homo sapiens* (P22105, P22105–3, O35452, C9J7W4 и др.) до 26%, фрагментам белков растений, таких как рис *Oryza sativa subsp. Japonica* (Q2QMS5) до 24%, томат *Solanum peruvianum* (Q4KR06, Q4KQZ5) до 36%, сорго *Sorghum bicolor* (C5YLV1) до 27%, кукуруза *Zea mays* (C0PLI0, C0PKW5 и др.) до 27%, а также грибов *Chaetomium globosum* (Q2H2I8) до 25%, *Aspergillus oryzae* (Q75R32) до 30%.

Фрагмент белка FRU2 (Putative uncharacterized protein FRU2; ID B8C3E9; 134 Ам; 13,730 Да; название гена – FRU2; *Thalassiosira pseudonana*) гомологичен фрагментам белков: амебы *Naegleria gruberi* (D2W009) до 26%, слизевика *Dictyostelium discoideum* (Q54V57, Q54YV1) до 50% и 70%, оболочника *Ciona intestinalis* (A2SY10) на 27%, моллюска *Haliotis fulgens* (D0EL74) до 44%, нематод *Caenorhabditis elegans* (Q5DTE9) и *Caenorhabditis briggsae* (A8XVK1) до 60–80%, шистосомы *Schistosoma mansoni* (C4QIR9) на 57%, дрозофил *Drosophila mojavensis* (B4KDH9) и *Drosophila simulans* (B4QQJ2) до 47% и 70%, комара *Leishmania major* (Q4QIZ4) до 73%, форели *Oncorhynchus mykiss* (C1BGN0) до 69%, лосося *Salmo salar* (B9EPX0) до 40%, лягушки *Xenopus laevis* (P10667) до 37%, мыши *Mus musculus* (Q9D9N0) на 46%, человека *Homo sapiens* (Q6ZVD8, Q6ZVD8-3, Q6ZVD8-2) на 40%, винограда *Vitis vinifera* (D1IWB7) на 40%, грибов *Magnaporthe grisea* (A4QQ53) до 47%, *Emericella nidulans* (Q5B653) до 41%.

Фрустулиновый белок (Frustulin protein; ID B8C0T6; 137 Ам; 14,914 Да; название гена – FRU3; *Thalassiosira pseudonana*). Для данного белка сервер показал только 42 протеина-гомолога (включая известные фрустулины) со степенью гомологии к фрустулину от 29 до 65%. Фрустулиновый белок имеет сходство только

с белками и фрагментами белков дрозофил *D. melanogaster* (Q9W5X1, Q8IQ18), *D. yakuba* (B4PYX7), *D. erecta* (B3NWX0), *D. sechellia* (B4I642), *D. virilis* (B4LUW4), *D. pseudoobscura pseudoobscura* (B5DJ80), *D. simulans* (B4Q310), *D. willistoni* (B4N1M4) на 36–39% и белком комара *Culex quinquefasciatus* (B0X780) до 29%. Остальные белки выделены из различных бактерий (*Saccharophagus degradans*, *Bacillus sp. P-358*, *Cellvibrio japonicus* и др.) и подобны фрустулину на 28–43%.

Необходимо отметить, что для основной части фрустулинов большинство гомологов представлено белками различных бактерий (*Jannaschia sp.*, *Escherichia coli* и др.), где уровень гомологии между отдельными участками цепей достигает 25–70%. Также многие фрустулины гомологичны (до 70%) белкам вирусов (*Choristoneura fumiferana defective polyhedrosis virus*, *Paramecium bursaria Chlorella virus FR483* и др.). Например, белок FRU2 подобен гликопротеину вируса *Human immunodeficiency virus 2* (Q9DSJ4) на 36%.

Предположительно, прямое отношение к фрустулинам может иметь часть полностью идентифицированных белков *Phaeodactylum tricornutum* CCAP 1055 / 1, *Thalassiosira pseudonana* CCMP1335, *Thalassiosira pseudonana* (B7GB28, B7G5U0, B7G575, B7G576, B7FV76, B7G3Y4, B7S3L2, B7G7C2, B7G7R2, B8BRF7, B8C9Y3, B8C774, B8C4P7, B8C8U2, B7G7C5, B7G4F8, B7FW82, B8LBV6, B8C7L2, B7FNY9, B7G8Y2, B8C9X0, B8C2K2, B8C7M9, B8CB46, B8BW49, B8LEW4, B8C4C0, B8CGR2, B8BSN8, B7FV77, B8CGR0, B8C0R9, B7G1P2, B8LDT2, B8CGD9, B8LEJ2, B8CGQ7, B8C2E4, B8CC83, B5YNT8, B8C7W1, B8BXT7, B8LEZ4, B5YLV4, B7G3S2, B8CEN6), гомология которых с известными фрустулинами достигает от 21 до 57%.

Таким образом, белки, подобные фрустулинам, широко охватывают ветви филогенетического дерева. Прямое подтверждение полученных нами данных имеется в исследовании [11], в котором при геномном анализе (с использованием базы данных по диатомовым водорослям Diatom EST Database) обнаружено большое сходство генов диатомей с генами многоклеточных.

Заключение

Данные биоинформатики указывают на родство фрустулинов диатомей и белков большого ряда различных форм жизни, стоящих на разных ступенях эволюции. Однако, вероятно, многие белки высших организмов в процессе эволюции претерпевали существенные изменения, приведшие к потере или изменению первоначальной функции. В то же время биологические функции обнаруженных гомологов фрустулинов полностью не установлены. Дальнейшее исследование структуры и функциональной активности фрустулинов и их гомологов, а также других белков биоминерализации будет способствовать расшифровке механизма биоминерализации у разных организмов и выявлению новых аспектов взаимодействия живого мира и минеральной среды.

Литература

1. Armbrust E. V., Berges J. A., Bowler C. et al. The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism // Science. 2004. Vol. 306. P. 79–86.

2. Bowler C., Allen A.E., Badger J.H., Grimwood J. et al. The Phaeodactylum genome reveals the evolutionary history of diatom genomes // Nature. 2008. Vol. 456. P. 239–244.
3. Falciatore A., Bowler C. Revealing thermoregular secrets of marine diatoms // Annu. Rev. Plant Biol., 2002. № 53. P. 109–30.
4. Fischer H., Robl I., Sumper M., Kröger N. Targeting and covalent modification of cell wall and membrane proteins heterologously expressed in the diatom *Cylindrotheca fusiformis* (BACILLARIOPHYCEAE) // J. Phycol. 1999. Vol. 35, № 1. P. 113–120.
5. Kroger N., Bergsdorf C., Sumper M. A new calcium-binding glycoprotein family constitutes a major diatom cell wall component // EMBO. 1994. Vol. 13. P. 4676–4683.
6. Kroger N., Bergsdorf C., Sumper M. Frustulins: domain conservation in a protein family associated with diatom cell wall // Eur. J. Biochem. 1996. Vol. 239. P. 259–264.
7. Kroger N., Lehmann G., Rachel R., Sumper M. Characterization of a 200-kDa diatom protein that is specifically associated with a silica-based substructure of the cell wall // Eur. J. Biochem. 1997. Vol. 250. P. 99–105.
8. Kroger N., Poulsen N. Diatoms – From Cell Wall Biogenesis to Nanotechnology // Annu. Rev. Genet. 2008. Vol. 42. P. 83–107.
9. Poll van de W.H., Vrieling E.G., Gieskes W.W.C. Location and expression of frustulins in the pennate diatoms *Cylindrotheca fusiformis*, *Navicula pelliculosa* and *Navicula salinarum* (Bacillariophyceae) // J. Phycol. 1999. Vol. 35. P. 1044–1053.
10. Scala S., Bowler C. Molecular insights into the novel aspects of diatom biology // Cell. Mol. Life Sci. 2001. Vol. 58. P. 1666–1673.
11. Zurzolo C., Bowler C. Exploring Bioinorganic Pattern Formation in Diatoms. A Story of Polarized Trafficking // Plant Physiol. 2001. Vol. 127. P. 1339–1345.