Курсовая работа по теме:

Предсказание специфичности в семействе ОРА

Выполнил: студент 4 курса ФББ МГУ П.В. Мазин Научный руководитель: старший преподаватель Факультета биоинженерии и биоинформатики О.В. Калинина доцент Факультета Биоинженерии и Биоинформатики к.б.н. А.Б. Рахманинова

Москва, 2007 г.

Введение

Обзор литературы

Суперсемейство MFS

Семейство ОРА

Состав семейства

Функционально важные остатки в UhpT E.coli

Структура и каталитический механизм транспортера GlpT E.coli

Материалы и методы

Выравнивание

Методы

Результаты

SDP

Предсказание специфичности

Обсуждение

Выводы

Список литературы

Введение

Суперсемейство мембранных транспортеров MFS содержит более 1000 последовательностей, обладающих низким уровнем сходства и широким спектром Большинство членов MFS схожую специфичности. имеют пространственную организацию и каталитический механизм. Аннотация специфичности белков из семейства MFS затруднена в связи с низкой степенью сходства последовательностей. Разработанный ранее метод SDPlight (Мазин, Калинина. 2007) позволяет быстро находить позиции, определяющие специфичность - SDP (Specificity Determining Position) и может быть применен для предсказания специфичности. В данной работе мы используем новый итерационный метод для предсказания SDP и специфичности белков из семейства OPA, входящего в MFS.

Используемые сокращения

ТМС – трансмембранный сегмент (α-спираль, пересекающая мембрану)

ТМС1-ТМС12 – ТМС с первого по двенадцатый в рассматриваемых ниже структурах

L1-2 - L11-12 - петли, соединяющие ТМС

ФЕП – фосфоенол-пируват

РСА – рентгеноструктурный анализ

 P_i – неорганический фосфат (H₂PO₄⁻, HPO₄²⁻)

UhpT - UhpT *E.coli*; TC #2.А.1.4.1; гексозофосфат/Р_i антипортер.

GlpT - GlpT *E.coli*; TC #2.А.1.4.3; глицерин-3-фосфат/Р_i антипортер

SDP - Specificity Determining Position, позиция в выравнивании определяющая субстратную специфичность.

Обзор литературы

Суперсемейство MFS

Суперсемейство Major Facilitator Superfamily (MFS, 2.А.1 по систематике Сайера, Saier, 2000) является большой группой вторичных транспортеров, осуществляющих транспорт по механизмам унипорта, симпорта и антипорта. MFS содержит более 1000 последовательностей, обладающих низким сходством и широким спектром специфичностей. В соответствии с предсказаниями положения мембранных сегментов большинство членов MFS содержат 12, 14 или 24 трансмембранных α-спиральных сегмента (www.tcdb.org).

Предполагается, что структура из 12 ТМС образовалась в результате двух дупликаций домена, состоящего из 3 ТМС (www.tcdb.org). Группой Р.С.Маloney была предсказана двухдоменная структура UhpT на основании небольшого сходства последовательностей в ТМС2 и ТМС8. Авторы предположили, что такая структура образовалась в результате дупликации домена содержащего 6 ТМС (M.C.Fann et. al., 1998). TetL (TC #2.A.1.3.16) содержит 14 ТМС и способен осуществлять антипорт тетрациклин/Н⁺, а также антипорт моновалентных ионов. Показано, что мутанты транспортера, лишенные 7-ого и 8-ого ТМС, сохраняют способность к антипорту моновалентных ионов (Jin et. al., 2001). То есть TetL, содержащий 12 ТМС, остается активным. Показано, что нитратный транспортер NarK (TC #2.A.1.8.11) из *P.pantotrophus*, содержащий 24 ТМС, состоит из 2-х доменов NarK1 и NarK2, включающих по 12 ТМС. Оба домена осуществляют транспорт нитрата, при этом NarK1 является протонзависимым NO_3^-/H^+ антипортером, а NarK2 – протон-независимым NO_3^-/NO_2^- антипортером (Wood et. al., 2002).

В настоящий момент доступны три пространственные структуры членов MFS. Две из них получены при помощи PCA и имеют высокое разрешение: глицерин-3-фосфат/P_i антипортер GlpT (3.3 Å, Huang et. al., 2003) и лактоза/H⁺ симпортер LacY (3.5 Å, Abramson et. al., 2003). Также при помощи электронной кристаллографии 2-мерных кристаллов была получена карта оксалат/формиатного антипортера (OxIT TC #2.A.1.11.1) с разрешением 6 Å (Heymann et. al., 2001). Все три структуры, несмотря на низкое сходство последовательностей имеют 12 одинаково расположенных TMC, двухдоменную организацию и псевдосимметрию второго порядка. Среднеквадратичное отклонение для 64% С_{α} атомов при наложении структур LacY и GlpT составляет 2.84 Å (Lemieux et. al., 2004).

Эти данные позволяют предположить, что структура, состоящая из 12 ТМС, является минимальным функциональным остовом транспортеров суперсемейства MFS. Таким образом, кажется вероятным, что двухдоменная структура является общей для всех членов MFS и существует общий транслокационный механизм для транспортеров этого суперсемейства (Lemieux et. al., 2004).

Разнообразие специфичностей с одной стороны, и относительно низкое сходство последовательностей с другой, делают интересным изучение детерминант специфичности для данного семейства. В данной работе мы провели детальное изучение детерминант специфичности в семействе ОРА (TC #2.A.1.4), к которому относится одна из двух известных на данный момент пространственных структур - GlpT и один из наиболее экспериментально изученных представителей MFS - UhpT. Также мы сделали новые предсказания субстратной специфичности других представителей семейства ОРА, которые хорошо согласуются с доступными данными.

Семейство ОРА

Семейство OPA (Organophosphate/P_i Antiporter) включает белки, транспортирующие в клетку сахара в обмен на экспорт P_i за счет осмотического потенциала неорганического фосфата.

Состав семейства

Семейство содержит несколько сотен транспортеров, разбитых на 6 групп, обладающих различной субстратной специфичностью, 5 из них содержат бактериальные белки (www.tcdb.org). Более подробно эти группы описаны в Таб.1.

5

Группа	Субстратная специфичность	Типичные	Дополнительная информация
		представители	
2.A.1.4.1	Гексозо-6 или 1-фосфаты	UhpT. E.coli	Транспортирует с низкой
	HO		специфичностью глюкозо-,
	P		фруктозо-, манозо, 2-
			диоксиглюкозо-6-фосфат. Не
	,		транспортирует ФЕП и
	нолик Уличон		фосфоглицерат (Ambudkar et. al,
			1986; Fann et. al., 1998; Hall et. al.,
	но он		1998; Schwöppe et. al., 2002).
2.A.1.4.2	3-фосфоглицерат	PgtP. S. typhimurium	Транспортирует ФЕП и 3-
	O O OH		фосфоглицерат с равной
			эффективностью. (Varadhachary
	но Т о он		et. al., 1990)
	ОН		
2.A.1.4.3	Глицерофосфат	GlpT. E.coli	Известна пространственная
	O \\OH		структура (Huang et. al., 2003).
	HO		
	ОН		
$2 \wedge 1 \wedge 4$		UhnT E coli	Реновлати и болон 2
2.A.1.4.4	тлюкозо-о-фосфат	011p1. <i>E.cou</i>	гецепторный ослок 2-
			(Wester et al. 1987)
			(weston et. al., 1987).
			высокоспецифично распознает
			тлюкозо-о-фосфат и способен к
			ero транспорту (Schwoppe et. al.,
			2003)
2.A.1.4.5	Глюкозо-6-фосфат	GSD1b. <i>H. sapiens</i>	Микросомальный транспортер
2.A.1.4.6	Глюкозо-6-фосфат	Hpt. C. pneumoniae	Не способен транспортировать
			фруктозо- и маннозо- 6-фосфат,
			но может переносить эритрозо-4-
			фосфат (Schwöppe et. al., 2002).

Таб. 1. Группы субстратной специфичности семейства OPA (www.tcdb.org).

Как видно по Таб.1, члены семейства ОРА могут обладать как широкой, так и узкой специфичностью. В данной работе мы разбили все семейство на две группы: гексозные (UhpT, UhpC,HPT) и триозные (GlpT и PgtP) транспортеры.

Функционально важные остатки в UhpT

За более чем 10 лет исследований группа Р.С.Маloney экспериментально охарактеризовала большое количество функционально важных остатков в UhpT. При помощи сканирования цистеином и лизином (т.е. получения мутантных белков моноцистеин (или лизин)-замещенных по интересующим позициям) было показано, что из 14 аргининов, содержащихся в UhpT, только два (Arg46 и Arg275) не могут быть заменены ни на лизин, ни на цистеин без потери активности. Arg20 и Arg325 могут быть заменены на лизин, но не на цистеин, без потери активности (Fann et. al., 1998).

При помощи сканирования цистеином с последующим ингибированием цистеинпара-хлор-Нд-бензо-сульфонововой чувствительным агентом кислотой (pchloromercuribenzosulfonate - PCMBS), по заряду и размеру похожей на глюкозо-6-фосфат, было показано, что ТМС7 (Yan et. al., 1993; 1995) и ТМС11 (Hall et. al., 2001) участвуют в транслокации субстрата. В ходе сканирования создавались моно-цистеин-замещенные мутанты для позиций, полный список которых приведен в Таб. 2. Большинство из таких мутантов сохраняло достаточно высокую активность. При введении цистеина в некоторые позиции получались мутанты, ингибирующиеся PCMBS конкурентно. Считается, что такие позиции находятся в активном центре. Для ТМС7 было показано существование трех групп остатков: доступных для PCMBS только со стороны цитоплазмы (домен А), только со стороны периплазмы (домен В) и доступных с обоих сторон (домен С). Также были исследованы все заряженные остатки находящиеся внутри ТМС, и способные образовывать солевые мостики. Было показано, что белок, в котором любой из таких остатков замещен на цистеин или на другой остаток, даже с таким же зарядом (лизинаргинин, глутаминовая кислота – аспарагиновая кислота) теряет активность. При замене обоих остатков, способных образовать солевой мостик, на цистеин все мутанты так же оказывались неактивны, кроме мутантов с замещенными позициями 388D→C и 391K→C, которые сохраняли активность (Hall et al, 1999). Также было показано, что мутанты UhpT содержащие на месте солевого мостика 388-391 положительный заряд теряют способность к транспорту гексозо-6-фосфатов, однако становятся способны транспортировать ФЕП и $\Phi\Gamma$ (Hall et al, 2002). Основываясь на своих наблюдениях и на том, что транспортер $\Phi E\Pi$ PgtP S. typhimurium содержит в аналогичной позиции аргинин, не входящий в состав солевого мостика, авторы предположили, что смена субстратной специфичности связанна с тем, что транслокация происходит в электронейтральном состоянии. Поэтому избыточно заряженный мутант UhpT способен транспортировать только субстраты, имеющие меньший заряд, чем глюкозо-6-фосфат, такие как ФЕП.

В работе (Fann et. al., 2003) при помощи сканирования всех цистеинов GlpT *E.coli* серином и ингибирования полученных мутантов PCMBS было обнаружено, что Cys163

7

доступен для PCMBS с обоих сторон мембраны, и что этот остаток функционально важен. На этом основании авторы предположили что TMC5 (в котором расположен остаток Cys163) участвует в транслокации субстрата.

Номер в	Номер в	TMC	Краткое описание	Источник
UhpT E.coli	GlpT E.coli			
R46	R45	1	Аргинин участвующий в связывании	Lemieux et al, 2004
			субстрата.	Fann et. al., 1998
K82	K80	2	Лизин, важный для функционирования	Fann et. al., 1998
			Uhpt. Возможен солевой мостик с E305	
			(по UhpT)	
H168	H165	5	Лежит между Arg46 и Arg269 (по GlpT	Lemieux et al, 2004
			E.coli)	
I166	C163		Лежит на пути транслокации	Fann et. al., 2003
			субстрата.	
	232-239	L6-7	Претерпевают конформационные	Lemieux et al, 2004
			изменения при связывании субстрата.	
I261, L264	L255, I258	7	Лежат на пути транслокации субстрата	Yan et. al., 1995
			со стороны цитоплазмы. Не входят в	
			активный центр.	
R275	R269		Аргинин участвующий в связывании	Lemieux et al, 2004
			субстрата.	Fann et. al., 1998
C265, N268,	A259, N262,		Лежат на пути транслокации субстрата.	Yan et. al., 1993; 1995
I269, L271,	V263, V265,		Входят в активный центр.	
V273, I276	L267, Y270			
D279	L273		Асп. к-та, важная для	Fann et. al., 1998
			функционирования Uhpt. Возможен	
			солевой мостик с R275 (по UhpT)	
S282, T283,	S276, P277,		Лежат на пути транслокации субстрата	Yan et. al., 1995
V284, A286	T278, L280		со стороны периплазмы. Не входят в	
			активный центр, кроме 282 по UhpT.	
E305	E299	8	Глут. к-та, важная для	Fann et. al., 1998
			функционирования Uhpt. Возможен	
			солевой мостик с К82 (по UhpT)	
D388, K391	A385, T388	11	Солевой мостик в UhpT. При	Hall et. al., 1999
			появлении некомпенсированного	
			положительного заряда меняет	
			специфичность на ФЕП.	
A395, L397,	G392, L394,		Лежат на пути транслокации. Входят в	Hall et. al., 2001
I398, F402	G395, A399		активный центр.	
D400, K404	S397, S401		Возможный солевой мостик в UhpT.	Fann et. al., 1998

Таб. 2. Остатки в GlpT и UhpT *E.coli* охарактеризованные экспериментально.

Важны для функционирования UhpT.	
----------------------------------	--

Структура и каталитический механизм транспортера GlpT

Структура фосфоглицеринового транспортера GlpT *E.coli*, определена в 2003 году (Huang et al., 2003) и в настоящий момент хорошо изучена (Hang et al, 2003; Lemieux et al, 2004). Общая структура и расположение транспортера в мембране клетки изображено на (Рис. 1а). С- и N- концы полипептидной цепи находятся в цитоплазме. Arg45 и Arg269, соответствующие Arg46 и Arg275 из UhpT, экспонированы в просвет поры и, вероятно, участвуют в связывании субстрата (Рис. 1с). Топология ТМС (Рис. 1d) указывает на



Рис 1. Структура GlpT. (а) Положение транспортера в мембране. N-концевой домен окрашен в зеленый цвет, C-концевой в розовый. (b) Периферийные TMC3, TMC6, TMC9, TMC12 обозначены зеленым; TMC2, TMC5, TMC8, TMC11 образующие центральную пору обозначены желтым; центральные TMC1, TMC4, TMC7 и TMC10 обозначены пурпурным. (c) Предполагаемый центр связывания субстрата (Arg45 из TMC1 и Arg269 из TMC7) с молекулой субстрата. Расстояние между гуанидиновыми группами Arg45 и Arg269 9.9 Å. (d) Последовательность и топология TMC GlpT. Arg45 и Arg269 выделены жирным шрифтом. Рисунок взят из статьи (Lemieux et al, 2004)

двухдоменную организацию транспортера: N-концевой домен TMC1-6 и, соединенный с ним петлей L6-7 (существенно более длинной, чем остальные петли в этой структуре), С-концевой домен TMC7-12. На основании структурных и биохимических данных Lemieux и сотр. предложили возможный транслокационный механизм, присущий, как считают авторы, многим вторичным транспортерам семейства MFS. Для GlpT, в соответствии с

этим механизмом, транспортируемый субстрат связывается с Arg46 и Arg275. Расстояние между гуанидиновыми группировками последних составляет 9.9 Å, что на 1.4 Å больше, чем оптимальное расстояние для кулоновского взаимодействия, между ними и фосфатной группировкой субстрата. Под воздействием связанного субстрата С- и N- концевые домены поворачиваются друг относительно друга. При этом пора (открытая ранее цитоплазму) открывается в периплазму (Рис 2). Направленность транспорта при этом осуществляется за счет градиента P_i и существенно большего сродства транспортера к углеводо-фосфатам. В работе (Fann et al., 1998) показано, что периплазматическая и цитоплазматическая стороны UhpT связываются с транспортируемыми субстратами с равной эффективностью, что хорошо согласуется с предположением о едином связывающем центре для обоих субстратов (Р_i и углеводо-фосфат).



Рис.2. Предполагаемый механизм действия GlpT *E.coli*. (а) Обозначены остатки Arg46 и Arg275, P_i обозначен как круг, глицерофосфат как треугольник. (b) Схематическое изображение уровня свободной энергии в ходе транслокации. Рисунок взят из статьи (Lemieux et al, 2004) Следует отметить, что структурные данные хорошо согласуются с результатами

биохимических исследований группы Р.С.Maloney. Предсказанные ими как критичные для функционирования транспортера остатки Arg46 и Arg275 (Arg45 и Arg269 в GlpT) действительно экспонированы в просвет центральной поры белка и, по результатам докинга субстрата, взаимодействуют с фосфатной группировкой (Lemieux et al, 2004).

Предсказанные как участвующие в транспорте ТМС7 и ТМС11 (Yan et. al., 1993; 1995; Hall et. al., 2001) и ТМС5 (в GlpT *E.coli*), образуют центральную пору (ТМС5 и ТМС11) и содержат центр связывания субстрата (ТМС7). Однако, по всей вероятности, далеко не все функционально важные остатки, были изучены.

Также в работе (Lemieux et al, 2004) было показано, что остатки 232-239 (по GlpT *E.coli*) принадлежащие центральной петле L6-7, могут играть роль в конформационном переходе происходящем при связывании субстрата.

Материалы и методы

Выравнивание

Полная выборка последовательностей была составлена на основе списка белков в записи InterPro IPR000849 (http://www.ebi.ac.uk/interpro/), а также ортологических рядов БД KEGG (http://www.genome.jp/kegg/). Следует отметить, что 10 последовательностей, описанных в IPR000849, аннотированы как транспортеры галактоната и кроме подписи GlpT transporter, содержат подпись Dgal transporter (IPR004744). С другой стороны часть последовательностей из ортологических рядов KEGG содержат лишь общую подпись суперсемейства MFS. По-видимому, это связано с тем, что суперсемейство MFS трудно делится на отдельные семейства. Поэтому в качестве исходной выборки были взяты все последовательности-кандидаты Затем из эукариотических белков оставили 6 в качестве внешней группы, и, наконец, исключили чересчур близкие последовательности (из каждых 2-х последовательностей, совпадающих более, чем на 98%, в выборку вошла одна). В обучающую выборку белки-прототипы (экспериментально вошли охарактеризованные) и их достоверные ортологи. Достоверным ортологом считался белок удовлетворяющий двум условиям:

1. Аминокислотная последовательность является лучшим двусторонним хитом (ВЕТ) для белка-прототипа, совпадение последовательностей более 30%.

2. Ген локализован в опероне или кластере генов, гомологичном оперону (кластеру) гена белка-прототипа.

Поиск ВЕТ и определение кластеров генов проводилось при помощи ресурсов энциклопедии KEGG (<u>http://www.genome.jp</u>).

Обучающая выборка была разбита на транспортеры триозо- и гексозо-фосфатов:

Транспортеры триоз	Транспортеры гексоз
GlpT_ECOLI P08194	UhpC_ECOLI P09836
GLPT_BACLD Q65CZ6	UHPC_MANSM Q65Q66 Q65Q66_MANSM
GlpT_Bacsu P37965	UHPC_PASMU Q9CLA1 Q9CLA1_PASMU
GLPT_ERWCT Q6CZI5 Q6CZI5_ERWCT	UHPC_PHOPR Q6LPP6 Q6LPP6_PHOPR
GLPT_FRATH Q2A293 Q2A293_FRATH	UHPC_SALTY P27669 UHPC_SALTY
GLPT_HAEIN P96335 GLPT_HAEIN	UHPC_VIBF1 Q5E1I7 Q5E1I7_VIBF1
GLPT_HSM 195_0220	UHPC_VIBPA Q87HJ7 Q87HJ7_VIBPA

CLET PASMULOOCI 05/00CL05 PASMU	
GLPT_PHOLL Q7MZY8 Q7MZY8_PHOLL	HPT_CHLAB Q5L737 Q5L737_CHLAB
GLPT_SALTY Q8ZNG6 Q8ZNG6_SALTY	HPT_CHLCV Q824R5 Q824R5_CHLCV
GLPT_VIBCH Q9KN29 Q9KN29_VIBCH	HPT_CHLFF Q252T6 Q252T6_CHLFF
GLPT_VIBPA Q87M76 Q87M76_VIBPA	HPT_CHLMU Q9PJJ8 UHPT_CHLMU
GLPT VIBVU Q8D6W8 Q8D6W8 VIBVU	HPT CHLPN Q9Z7N9
PGTP_ECOL6 Q8FAP1 Q8FAP1_ECOL6	HPT_CHLTA Q3KLF3 Q3KLF3_CHLTA
PGTP_SALTY P12681 PGTP_SALTY	HPT_PARUW Q6ME88 Q6ME88_PARUW
	UHPT_ECOLI P0AGC0
	UHPT_MANSM Q65Q69 Q65Q69_MANSM
	UHPT_PASMU Q9CL98 Q9CL98_PASMU
	UHPT PHOPR Q6LPP3 Q6LPP3 PHOPR
	UHPT SALTY P27670 UHPT SALTY
	UHPT_VIBF1 Q5E1I4 Q5E1I4_VIBF1
	UHPT VIBPA Q87HK0 Q87HK0 VIBPA
	UHPT_VIBVU Q8D6L9 Q8D6L9_VIBVU

Методы

Предсказание SDP

Для определения SDP использовался метод SDPlight (Мазин, Калинина. 2007).

Предсказание специфичности

Для предсказания специфичности производилось взвешивание каждой последовательности к каждой группе специфичности. Вес последовательности *S* к группе *j* определялся по формуле:

$$W(S) = \sum_{i \in SDP} w_j(i, S_i);$$

где:

S_i – аминокислота в последовательности S в позиции i,

 $w_i(i, \alpha)$ – вес аминокислоты α в позиции *i* в группе *j*,

Вес *w* вычислялся по формуле:

$$w_j(i,\alpha) = \frac{\ln f_i(j,\alpha) - M(\ln f_i(j,\alpha))}{\sqrt{D(\ln f_i(j,\alpha))}};$$

где:

 $f_i(j, \alpha)$ - доля аминокислоты α в группе *j* в позиции *i*.

M() и D() – матожидание и дисперсия равные соответственно:

$$M(\ln f_i(j,\alpha) = \sum_{\alpha=1}^{20} q(\alpha) \times \ln f_i(j,\alpha);$$
$$D(\ln f_i(j,\alpha) = \sum_{\alpha=1}^{20} q(\alpha) \times \ln^2 f_i(j,\alpha) - M^2(\ln f_i(j,\alpha);$$

где:

 $q(\alpha)$ - частота аминокислоты α во всем исследуемом выравнивании

Для каждой последовательности также вычислялся фоновый вес (т.е. вес ко всему выравниванию: $f_i(j, \alpha) = f_i(\alpha)$, где $f_i(\alpha) - доля$ аминокислоты α во всем выравнивании)

Для разбиения семейства на группы специфичности использовалась следующая итерационная процедура:

1. Инициация: обучающая выборка. Остальные последовательности выравнивания имеют нулевую специфичность.

2. Шаг итерации: SDP определяется методом SDPlight. Все последовательности выравнивания взвешиваются относительно профилей для каждой группы. Последовательности перегруппируются в соответствии с максимальным весом (последовательности, для которых максимальный вес соответствует фону, получают нулевую специфичность и не относятся ни к какой группе специфичности).

3. Завершение: шаг итерации не приводит к перемещению последовательностей.

Результаты

SDP

График бернуллиевской вероятности имеет три минимума примерно одного порядка. Предсказанные SDP приведены на таб.3. Проекция предсказанных SDP (до первого минимума) на пространственную структуру GlpT приведена на Puc. 3. Туг270, Leu273, Tyr298, Tyr362 непосредственно окружают Arg269. Gly392 непосредственно контактирует с консервативным Tyr393, входящем в окружение Arg45. Ala301 и Ile303 находятся в месте изгиба TMC8 в непосредственной близости от Tyr298. Met83 и Met136 находятся на расстоянии менее 5 Å друг от друга и экспонированы в мембрану. Trp149 экспонирован цитоплазму и находится в непосредственной близости петли L6-7, принимающей участие в конформационных изменениях при связывании субстрата. Все предсказанные SDP находятся в внутренних TMC, участвующих в образовании поры.



Рис. 3. Пространственная структура GlpT. Синими и красными точками изображены цитоплазматическая и периплазматическая поверхности мембраны соответственно. Белок представлен в ленточной модели. Предсказанные SDP изображены в проволочной модели и выделены красным цветом, Arg45 и Arg269 – синим. (А) вид сбоку, (В) вид с цитоплазматической стороны. (С) Проекция SDP на вторичную структуру GlpT. SDP обозначенны красным цветом.

Δ	
	١
V	2
٩	2
- 5	5
	5
2	
50	3
5	3
2	5
Ē	1
4	2
È	
3	2
C	5
, a	3
Ŀ	

Ta6.3 I	Іредсказанные SDP.			
Позиция по выравниванию	Аминокислотный состав в группе триозных транспортеров	Аминокислотный состав в группе гексозных транспортеров	Позиция по Glpt	Z- score
470	XSSSSXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX	XX3333333434343434343434343434343434343	362	8,14
386	IIIIIAMMIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	EFFFFFELLLLLLLLLLLLLFFFFFFFFFFFFFFFLFLLLLLL	303	7,68
384	AAAAPAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	000000000000000000000000000000000000000	301	7,19
177	WEEWEWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW	ΧΑΧΑΧΑΧΑΤΠΤΠΙΠΙΠΙΑΥΑΥΥΥΥΥΥΥΥΥΥΥΥΥΥΥΥΥΥΤΙΥΙΤΙΥΥΥΥΥΥΥ	149	7,05
381	XXXXEXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX	EREFERERERERERERERERERERERERERERERERERE	298	6,68
351	TATTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAAAAAAA	SNDDDDDDDDNNNNNNNNNNNNNNNDDDDDDDDDNNNNNN	273	6,31
106	SUNUNUN SM SMUMMAMAMATANA MAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMA	SSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS	83	5,93
164	MINIMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMA	WWWWEWWWTTTTTTTTTWWEWWWWWWWWWWWWWWWWWWW	136	5,71
501	GGGGGGGGGGGGSSGGGGGGGGGGGGGGGGGGGGGGGGG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGGAAAA	392	5,66
348	XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX	TTITITITITITITTTTTTTTTTTTTTTTTTTTTTTTT	270	5,6
	Первый миниму	Wi		
473	<u>νυνυνυνυνυνοφοννυνυνυνυνυνυνυνυνυνυνυνυν</u>	0H000000000000000000000000000000000000	365	5,06
55	IIIITAIIIIIIIVVIIIIIIIIIIIIIIIIIIIIIIII	TUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUU	34	4,88
463	IIIISTTTLIVIVAAIVIVIIIFVVVIIITTIAVVVVVVVVVVVSASSSSGSLAGIASSSSV	EFFFEFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF	355	4,83
336	IIIIIIIAIIIIAIIIIIIIIIIIIIIIIIIIIIIIII	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	258	4,77
494	AAAAAAAAAAAAVVAAAAAAAAAAAAAAAAAAAAAAAA	TTTTTTTTDDDDDDDSSTSSSSDTTTTTTTTTTDDTNTNNNNNTT	385	4,63
387	PPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP	IFMVIVVVVVVVVVFFFFFFFFFFFFFFFFFFFFFFFFF	304	4,54
512	IILIVVVVIIIVVLLVVIVVIAALALVVVVAIAAAAAAAA	PPPPPPPGGGGGGPPPPPPPGPPPPPPPPPPPPGGPGGGG	403	4,53
	Второй миниму	W		
504	CCCCTCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	TLLMMMMIIIIIIEFFFFFLLLLMMMLLLIILLLFFFFFVI	395	4,33
538	GGGGVGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	FFFFFFFTTTTTTTFFFFFFFFFFFFFFFFFFFTTTTTWWWYTTVV	416	4,22
71	LFLLVLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLL	AFYYAYYYIIIVIVVVFFFFFFFFFFFFFFFFFFFIIFAFAAAAAAA	50	4,1
272	CCVCVCCCCCCCEECCCCCVVCVCQQCCCCVCCCCCCCCQEVEEEEEVAKVEVVVVVT	TLMLIMMQYYYYYYYLLLLLLYMLLLMMMMMMIYYQ	215	4,01
420	MMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMM	AAAAAAALLLLLLLLSSTSSSSLATTTSSSAAAVLLALALMMWVVVAC	329	3,72
378	YYYYFYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY	VLLLVLLFFFFFFFVVVVVVFIVVVLLLVVVVFFLILIIIIIIV	295	3,64
482	LILLLLLLLLLLLMMLLLLLLLLLLLLLLLLLLLLLLLL	AAAAAAVVVVVVAAAAAAAAAAAAAAAAAAAAAAAAAAA	373	3,55
173	MMLMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMM	TLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLL	145	3,52
95	IVLILAAAIIVVVMMVVVVVIIIVVIIVVAAALVVVVVVVVVV	FFFFFFFFFFFFFLLLLLLFFLLLFFFFFFFFFFFFFF	72	3,47
191	IKKIKIIIIIIIVVVLILIIIVVIVVKKIIIRTKKKKKKKKKKKKKKKKKKKKKKKKK	WWWWRLLFFYYFWWWWWWWWWWWWWWWWFFWYWYYYYWWWW	157	3,42
65	VLIVGVVVVVGVVVVVVVVVVVVVLLVLLLLLVVVVVVVV	TTSTTTTTIVIIIIIITTTTTTTTTTTTTTIIIIIIIIII	44	3,4
155	IMLIMIIILLLLLVVLLLLLLLLLLLLLIIIIMMQQQQQQQQQQ	XXMMMMMMMXXMMMMMMMMMMMMMMMXXXXXXXXXXXMMMM	127	3,38
510	SNNSQSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS	GGGGGGGGKKKKKKKKGGGGGGGGGGGGGGGGGSSSSGKKGKGKKKKKGG	401	3,38
192	VMTVMVVVVVVGGVVVVVVVVVVVVVVVVMMMMMMMMMM	МММММММТГГГГГГГТИМММММММММММММММТГИГГГГГЛА	158	3,34

Предсказание специфичности

Филогенетическое дерево для изучаемой выборки транспортеров ОРА приведено на Рис.4. Синим и красным обозначены ветви соответсвующие белкам, вошедшим в обучающую выборку: транспортерам гексозофосфатов и триозофосватов соответсвенно. Голубым и оранжевым обозначены результаты предсказаний. Черным - белки не вошедшие ни в одну группу. В желтый круг обведена группа эукариотических транспортеров, в синий - транспортеров галактозы.



Рис.4 Филогенетическое дерево семейства ОРА. Дерево построено по изучаемой выборке (процедуру составления выборки см выше). Синим и красным обозначены ветви соответсвующие белкам, вошедшим в обучающую выборку: транспортерам гексозофосфатов и триозофосватов соответсвенно. Голубым и оранжевым обозначены результаты предсказаний. Черным - белки не вошедшие ни в одну группу.

Как видно на рисунке, эукариотические транспортеры и транспортеры галактозы (и группа близких к ним транспортеров из семейства OPA) не были предсказаны как принадлежащие к какой бы то ни было группе специфичности. Вероятно они имеют несколько отличный от характерного для семейства OPA транслокационный механизм или механизм распознавания субстрата. В остальном результаты разбиения соответствуют филогенетическому дереву.

Обсуждение

Как видно (Рис. 3) большая часть предсказанных SDP группируется около Arg269. Два из них (Gly392 и Tyr270) по экспериментальным данным находятся в активном центре. Для предсказанного как SDP Leu273 (Asp279 по UhpT) экспериментально показано участие в функционировании транспортера. Ala385 и Thr388, по экспериментальным данным, определяют специфичность транспортера по отношению к заряду субстрата, а не к его размеру, и поэтому они не предсказаны как SDP (Ala385 находится между первым и вторым минимумами бернуллиевской вероятности). Про остальные SDP экспериментальные данные отсутствуют.

Интересным фактом является то, что большинство предсказанных нами SDP Arg269. Особый интерес представляют группируются около находящиеся В непосредственной близости (менее 3.5 Å) предсказанные как SDP остатки тирозина: Туг362, Туг298, Туг270 (Рис. 5). Из них только Туг298 находится достаточно близко для образования водородной связи. При этом связь образуется не с гуанидиновой группировкой, а с карбоксильным кислородом остова. Интересно, что в группе гексозных транспортеров остатки тирозина окружающие Arg269 заменяются на апротонные и, в основном, неполярные остатки (Таб. 3). При этом характерный для группы триозных транспортеров предсказанный как SDP неполярный остаток Leu273 у гексозных заменяется на кислую аминокислоту.



Рис. 5. Ближайшее окружение Arg45 (A) и Arg269 (B). Синим обозначены остатки аргинина (Arg45 и Arg269), зеленым и красным – остатки хоть один атом которых находтся на расстоянии менне 3.5 Å от Arg45 или Arg269. SDP обозначенны красным цветом

Иначе выглядит пространственное окружение другого функционально важного остатка аргинина - Arg45. Оно так же в основном сформировано остатками тирозина (Рис. 5). Однако Туг42, Туг38 и Туг393 (окружающие Arg45) консервативны. Кроме того, гидроксильные группы Туг393 и Туг42 находятся на расстоянии около 3 Å от гуанидиновой группировки Arg45 и, вероятно, образуют с ней водородные связи.

Выше сказанное позволяет предположить функциональную ассиметрию Arg45 и Arg269. Жестко фиксированный в пространстве Arg45 участвует в кулоновском связывании субстрата, а относительно свободный Agr269 и окружающие его вариабельные остатки отвечают за специфическое распознавание субстрата.

Результаты предсказания специфичности, приведены на Рис. 4. Как видно, предложенная процедура позволяет объединять в одну группу специфичности не только близкородственные, но и далекие группы, такие как UhpT и UhpC. Это свойство связано с тем, что формируемая для предсказания специфичности на первом шаге итераций обучающая выборка не зависит от филогенетического дерева последовательностей – в одну группу специфичности могут войти эволюционно далекие последовательности. Это может быть полезно, если существуют данные, позволяющие предполагать одинаковую специфичность для мало похожих белков.

Выводы

- 1. Разработан новый итерационный метод предсказания SDP и субстратной специфичности, использующий обучающую выборку.
- 2. Метод применен к семейству мембранных транспортеров ОРА
- 3. На основании найденных SDP сделано предсказание о функциональной ассиметрии С- и N- концевых доменов транспортеров семейства MFS
- Для всех членов семейства ОРА предсказана специфичность по отношению к транспортируемому субстрату. Семейство разбито на транспортеры тризо- и гексозофосфатов.
- 5. Принцип независимости от филогенетического дерева при формировании обучающей выборки позволяет одновременно предсказать специфичность для эволюцонно далеких групп гезсозных транспортеров UhpC и UhpT.

Список литературы

Abramson, J., I. Smirnova, V. Kasho, G. Verner, H.R. Kaback, and S. Iwata. (2003). Structure and mechanism of the lactose permease of *Escherichia coli*. Science 301: 610-615.

Ambudkar SV, Larson TJ, Maloney PC. Reconstitution of sugar phosphate transport systems of Escherichia coli. J Biol Chem. 1986 ;261(20):9083-6.

Atul Varadhachary and Peter C. Maloney. Reconstitution of the Phosphoglycerate TransportP rotein of *Salmonella typhimurium*. Jr Biol Chem. Vol. 266, No. 1, pp. 130-135,1391.

Fann M.-C., A.H. Davies, A. Varadhachary, T. Kuroda1, C. Sevier, T. Tsuchiya, P.C. Maloney. Identification of Two Essential Arginine Residues in UhpT, the Sugar Phosphate Antiporter of *Escherichia coli*. J. Membrane Biol. 164, 187–195 (1998)

Fann M.-C., Anne Busch, and Peter C. Maloney. Functional Characterization of Cysteine Residues in GlpT, the Glycerol 3-Phosphate Transporter of *Escherichia coli*. Jr of Bacter, 2003, p. 3863–3870 Vol. 185, No. 13

Fann M.-C., Peter C. Maloney. Functional Symmetry of UhpT, the Sugar Phosphate Transporter of *Escherichia coli*. Jr of Biol Chem. Vol. 273, No. 50, pp. 33735–33740, 1998

Hall J.A, Mon-Chou Fann, and Peter C. Maloney. Altered Substrate Selectivity in a Mutant of an Intrahelical Salt Bridge in UhpT, the Sugar Phosphate Carrier of *Escherichia coli*. Jr of Biol. Chem. Vol. 274, No. 10, pp. 6148–6153, 1999.

Hall J.A., Peter C. Maloney. Pyridoxal 5-Phosphate Inhibition of Substrate Selectivity Mutants of UhpT, the Sugar 6-Phosphate Carrier of *Escherichia coli*. Jr of Bactr, 2002, p. 3756– 3758 Vol. 184, No. 13

Hall J.A., Peter C. Maloney. Transmembrane Segment 11 of UhpT, the Sugar Phosphate Carrier of *Escherichia coli*, Is an a-Helix That Carries Determinants of Substrate Selectivity. Jr of Biol Chem Vol. 276, No. 27, pp. 25107–25113, 2001.

Heymann, J.A.W., R. Sarker, T. Hirai, D. Shi, J.L.S. Milne, P.C. Maloney, and S. Subramaniam. (2001). Projection structure and molecular architecture of OxIT, a bacterial membrane transporter. EMBO J. 20: 4408-4413.

Huang Y., M. Joanne Lemieux, Jinmei Song, Manfred Auer, Da-Neng Wang. Structure and Mechanism of the Glycerol-3-Phosphate Transporter from *Escherichia coli*. Science. 2003 1;301(5633):616-20.

Huang, Y., M.J. Lemieux, J. Song, M. Auer, and D.-N. Wang. (2003). Structure and mechanism of the glycerol-3-phosphate transporter from *Escherichia coli*. Science 301: 616-620.

Jin, J., A.A. Guffanti, C. Beck, and T.A. Krulwich. (2001). Twelve-transmembranesegment (TMS) version (*TMS VII-VIII) of the 14-TMS Tel(L) antibiotic resistance protein retains monovalent cation transport modes but lacks tetracycline efflux capacity. J. Bacteriol. 183: 2667-2671.

Lemieux MJ, Huang Y, Wang DN. The structural basis of substrate translocation by the Escherichia coli glycerol-3-phosphate transporter: a member of the major facilitator superfamily. Curr. Opin. Struct. Biol. 2004 Aug;14(4):405-12.

M.-C. Fann, A.H. Davies, A. Varadhachary, T. Kuroda1, C. Sevier, T. Tsuchiya, P.C. Maloney. Identification of Two Essential Arginine Residues in UhpT, the Sugar Phosphate Antiporter of *Escherichia coli*. J. Membrane Biol. 164, 187–195 (1998)

M.H. Saier, A functional-phylogenetic classification system for transmembrane solute transporters. Microbiology and Molecular biology reciews,June 2000, p. 354–411

Schwöppe C, Herbert H. Winkler, H. Ekkehard Neuhaus. Connection of transport and sensing by UhpC, the sensor for external glucose-6-phosphate in Escherichia coli. Eur. J. Biochem. 270, 2003, 1450–1457.

Schwöppe C., Herbert H. Winkler, H. Ekkehard Neuhaus. Properties of the Glucose-6-Phosphate Transporter from *Chlamydia pneumoniae* (HPTcp) and the Glucose-6-Phosphate Sensor from *Escherichia coli* (UhpC). Jr of Bactr, 2002, p. 2108–2115 Vol. 184, No. 8

Weston L. A., Robert J. Kadner. Identification of Uhp Polypeptides and Evidence for Their Role in Exogenous Induction of the Sugar Phosphate Transport System of Escherichia coli K-12. Jr of Bacter, 1987, Vol. 169, No. 8, P. 3546-3555

Wood, N.J., T. Alizadeh, D.J. Richardson, S.J. Ferguson, and J.W.B. Moir. (2002). Two domains of a dual-function NarK protein are required for nitrate uptake, the first step of denitrification in *Paracoccus pantotrophus*. Mol. Microbiol. 44: 157-170.

Yan RT, Maloney PC. Identification of a residue in the translocation pathway of a membrane carrier. Cell. 1993 Oct 8;75(1):37-44.

Yan RT, Maloney PC. Residues in the pathway through a membrane transporter. Proc. Natl. Acad. Sci. USA. Vol. 92, pp. 5973-5976, 1995 Biochemistry.

Мазин П.В., Калинина О.В. SDPlight – быстрый метод поиска позиций, пределяющих специфичность в белковых выравниваниях. Ломоносов-2007. 11 - 14 апреля 2007 г.