

SŁAWOMIR FILIPEK

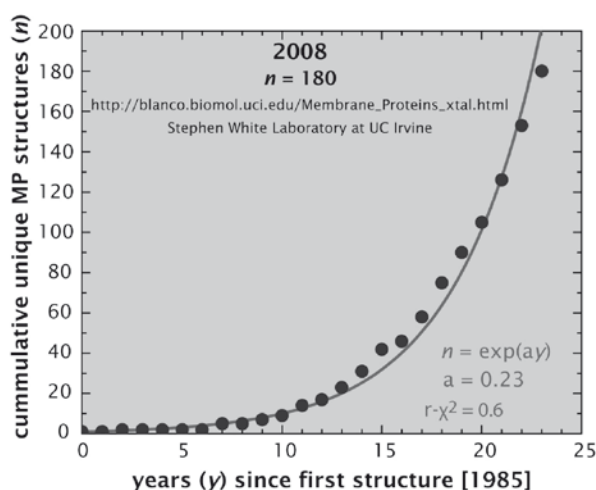
Międzynarodowy Instytut Biologii Molekularnej i Komórkowej
Ks. Trojdena 4, 02-109 Warszawa
Wydział Chemii Uniwersytetu Warszawskiego
Pasteura 1, 02-093 Warszawa
E-mail: sfilipek@iimcb.gov.pl

BIOINFORMATYKA STRUKTURALNA BIAŁEK TRANSBŁONOWYCH

WPROWADZENIE

Białka błonowe stanowią około 20–30% naszego genomu. Jednak z powodu dużych trudności w ich krystalizacji obecnie znanych jest zaledwie 187 struktur białek błonowych (z mutantami tych samych białek 457) (stan na kwiecień 2009 r.). Wykres przedstawiający wzrost liczby struktur białek błonowych, zdeponowanych w bazie Protein Data Bank (PDB), jest przedstawiony na Ryc. 1. Jak widać z tego wykresu, wzrost jest wykładniczy, co pozwala mieć nadzieję na szybkie uzyskanie znacznej liczby nowych struktur białek błonowych. Wzrost ten jest wynikiem zastosowania w ostatnich latach nowych technik krystalograficznych, w szczególności mikroogniskowania, co pozwoliło zmniejszyć rozmiar monokryształu białka nawet do kilku mikrometrów. Nawet wtedy można otrzymać struktury białek o rozdzielczościach poniżej 2 Å, wymaganych do prawidłowego zidentyfikowania położenia wszystkich atomów ciężkich (czyli bez atomów wodoru) w białku. Mikroogniskowanie w mikrokrystalografii umożliwia posługiwanie się wiązką synchrotronową o średnicy zaledwie 1 μm. W ten sposób uzyskano np. strukturę białka z rozdzielczością 1,5 Å dla kryształu o objętości 20 μm³ odpowiadającej 2 × 10⁸ komórek elementarnych. Dla tak małych kryształów prawie nie występuje efekt zniszczenia sieci krystalicznej przez olbrzymie natężenie promieniowania rentgenowskiego prawdopodobnie dlatego, że wzbudzone fotolektrony mogą łatwo wydostać się z małego obszaru naświetlania.

Spektakularnym osiągnięciem nowej techniki było uzyskanie w latach 2007–2008 struktur trzech receptorów GPCR (ang. G protein coupled receptors – receptory sprzężone z białkiem G). Pierwszym, skrytalizowanym już w 2000 r., receptorem z tej rodziny (jak się szacuje istnieje ponad 800 typów tych receptorów) była rodopsyna (PALCZEWSKI i współaut. 2000). Rodopsyna występuje w dużych ilościach w komórkach pręcikowych siatkówki oka, zatem otrzyma-



Ryc. 1. Kumulatywna liczba struktur białek błonowych zdeponowanych w bazie Protein Data Bank w poszczególnych latach. Wg. http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html

nie czystej rodopsyny było stosunkowo proste; trudność sprawiła jej krystalizacja. Inne receptory GPCR występują w bardzo małych stężeniach w błonach komórkowych, przeważnie w rejonach synaps, dlatego też można je uzyskać w większych ilościach tylko na drodze biotechnologicznej. Receptory te bardzo słabo krystalizują, ale ta trudność została pokonana przez mikrokrytalografię, natomiast biotechnologia dostarczyła metod pozwalających uzyskać mutanty badanych białek, które są bardziej stabilne w detergentach (detergenty są niezbędne do usunięcia błony), a więc również bardziej skłonne do krystalizacji w postaci dobrze wykształconych mikrokryształów (SERRANO-VEGA i współaut. 2008). W ten sposób otrzymano struktury dwóch receptorów adrenergicznych (β_1 AR i β_2 AR) oraz strukturę receptora adenozynowego A_{2a} R.

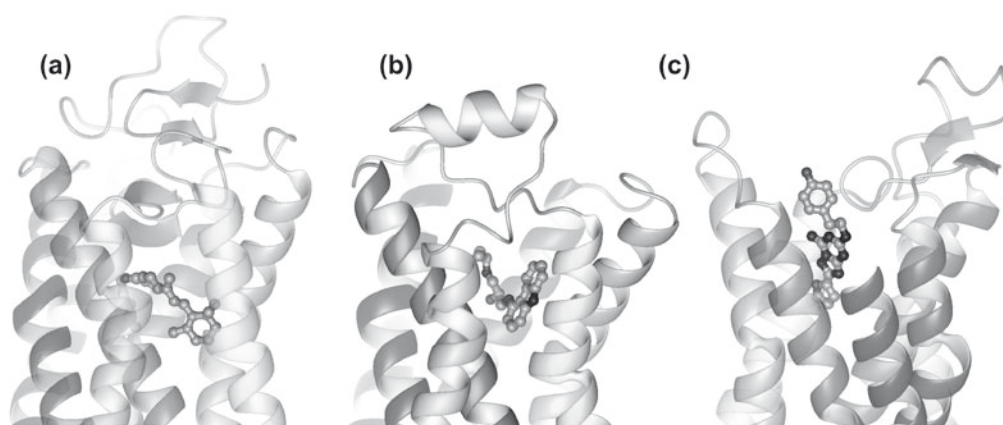
Inną obiecującą techniką do otrzymywania struktur białek błonowych jest cryo-EM (ang. cryo-electron microscopy – niskotemperaturowa mikroskopia elektronowa). Metoda ta może być zastosowana do kryształów 2D (dwuwymiarowych – np. pojedyncza warstwa molekuł białka w płaskiej błonie komórkowej), które łatwiej otrzymać niż monokrysz-

tały. Poza tym takie białka w kryształach 2D są wciąż w błonie komórkowej, a więc posiadają strukturę bardziej zbliżoną do natywnej niż w kryształach. Także kontakty między sąsiednimi molekułami białka odzwierciedlają w znacznym stopniu kontakty natywne. Niestety, typowe rozdzielczości dla tej metody są większe niż 4,5 Å, co nie pozwala wyróżnić poszczególnych aminokwasów, a bardziej mobilne części helis i pętle są niewidoczne z powodu nieuporządkowania. Wobec tego, nieregularności w strukturze białka mogą być łatwo przeoczone (patrz dalej – problemy ze strukturą akwaporyny). Na szczęście, poprzez zastosowanie metod obliczeniowych wspomaganych dodatkowo przez dane biochemiczne, można otrzymać nawet pełnoatomowe modele badanych białek (FLEISHMAN i współaut. 2006). Metoda cryo-EM może być nawet zastosowana do pojedynczych obiektów molekularnych pod warunkiem, że są odpowiednio duże. W ten sposób udało się uzyskać ogólny kształt kompleksu γ -sekreazy, ludzkiej proteazy błonowej, składającej się z czterech białek transbłonowych. Ponieważ struktura żadnego z tych białek jak dotąd nie jest znana, nie jest także możliwe dopasowanie ich do wyznaczonego kształtu kompleksu.

TEORETYCZNE METODY WYZNACZANIA STRUKTUR BIAŁEK BŁONOWYCH

Białka błonowe, a w szczególności receptory GPCR, w większości mają potencjalnie bardzo duże znaczenie terapeutyczne. Stąd też duży nacisk na otrzymanie możliwie szybko dużej liczby struktur tych receptorów. Oprócz zastosowania krystalografii, struktury te można otrzymać również metodami modelowania przez homologię/podobieństwo do struktury wzorca. Przez długi czas takim wzorcem dla receptorów GPCR była bakteriorodopsyna, chociaż nie należy ona do tej rodziny receptorów, a nawet nie jest receptorem, lecz pompą protonową usuwającą jony wodorowe z wnętrza komórki bakteryjnej. Tym niemniej budowa jej części transbłonowej odpowiada budowie receptorów GPCR. Część ta składa się z wiązki siedmiu α -helis, które na przemian przechodzą przez błonę komórkową. Zawiera również retinal, podobnie jak rodopsyna. Od 2000 r. wzorcem do budowy receptorów GPCR stała się struktura rodopsyny (to już jest receptor GPCR, ale ligand jest związany kowalencyjnie z białkiem), a następnie (od 2007 r.) struktura receptora β_2 -adrenergicznego.

Jednak, jak pokazuje przykład krystalicznej struktury receptora adenozynowego A_{2a} R, sposób modelowania przez homologię do celów dokowania ligandów i projektowania leków może być bardzo zawodny. Już struktura receptora β_2 AR wykazała, że pętla przykrywająca miejsce wiązania liganda może mieć różny kształt: zamiast β -kartki występującej w rodopsynie (Ryc. 2a) otrzymano krótką α -helisę (Ryc. 2b). Jednak położenie ligandów, retinalu w rodopsynie i karazolu (odwrotny agonista), w β_2 AR było prawie identyczne. Takie samo położenie miał ligand w receptorze β_1 AR. Tym większe było więc zakłopotanie teoretyków, kiedy okazało się, że ligand w receptorze A_{2a} R leży prostopadle do spodziewanego kierunku ułożenia liganda – spodziewanego według struktur wzorców z innych receptorów GPCR. Ponadto, ligand ten wystawał na zewnątrz receptora, choć nie był wcale taki duży i w modelach mieścił się dobrze wewnątrz białka. Drobnym już szczegółem był fakt, iż pętla przykrywająca ligand w miejscu wiązającym była zupełnie różna od dwu poprzed-



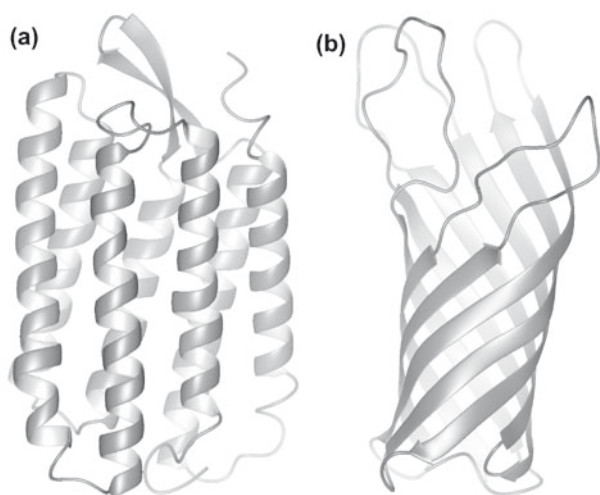
Ryc. 2. Miejsca wiążące ligand w receptorach GPCR.

(a) retinal w rodopsynie (kod 1U19 w bazie PDB); (b) odwrotny agonista karazolol w receptorze β_2 AR (kod 2RH1 w bazie PDB); (c) antagonista ZM241385 w receptorze A_{2A} R (kod 3EML w bazie PDB).

nich typów i tak naprawdę tym razem nie przykrywała liganda (Ryc. 2c).

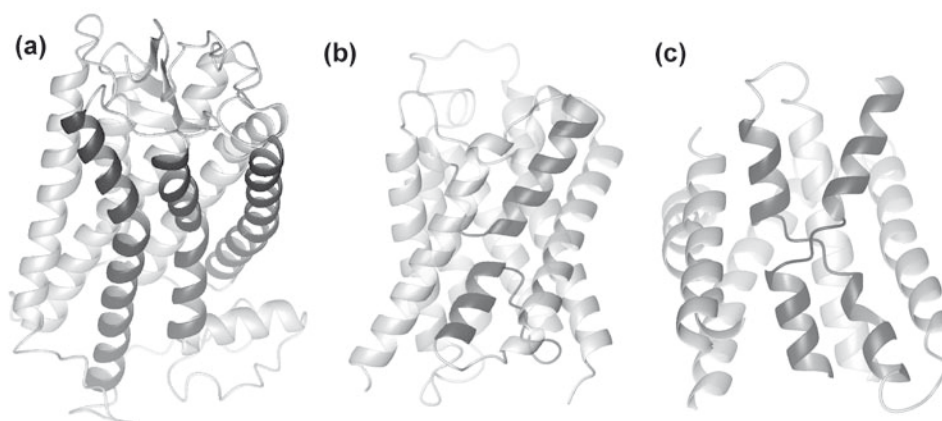
Części transbłonowe białek błonowych dzielą się na dwa zasadniczo różne typy: jest to albo wiązka α -helis (Ryc. 3a) albo też cylinder złożony ze zwiniętych β -kartek, tzw. β -beczka (Ryc. 3b). Ten drugi typ białek przeważa w zewnętrznej błonie komórek bakteryjnych (i jest zapewne związany ze specyficzną budową tych błon), podczas gdy białka, których część błonowa składa się z α -helis reprezentują wszystkie znane typy białek błonowych organizmów wyższych. Nie stwierdzono do tej pory typów mieszanych,

zawsze w części błonowej jest β -beczka lub wiązka α -helis. Nie są to jednak idealne α -helisy, bowiem niektóre z nich mogą być „pozaginane”, jak np. w strukturze rodopsyny (Ryc. 4a) i strukturach innych receptorów GPCR. Te zgięcia są istotne dla funkcji tych białek, bowiem umożliwiają zmianę struktury receptora po jego aktywacji, przez związanie liganda lub absorpcji kwantu światła. Oprócz „pogiętych” helis występują także inne odstępstwa od idealnego wyobrażenia o helisie w wiązce helis transbłonowych; w szczególności mogą występować „połówkowe” helisy. Na pierwszy rzut oka w strukturze akwaporyny występuje siedem α -helis transbłonowych, podobnie jak w rodopsynie lub bakteriorodopsynie – tak wynika ze struktur krystalicznych uzyskanych z niską rozdzielczością. Dokładniejsze badania wykazały jednak, że jedna z helis jest „pęknięta” i składa się z dwu helis połówkowych (Ryc. 4b). Taka helisa dochodzi do środka błony, podobnie jak helisa transbłonowa, lecz wraca z powrotem do tej samej części błony komórkowej w postaci luźnego łańcucha. Te luźne części są schowane we wnętrzu białka i nie mają kontaktu z fosfolipidami w błonie. Podobny „defekt” można znaleźć w strukturze antyportera sodowo-wodorowego (Na^+/H^+) NhaA, który transportuje jony sodowe na zewnątrz komórki, wykorzystując gradient stężenia jonów wodorowych, które są transportowane w tym samym czasie w drugą stronę. Białko to składa się z 12 helis transbłonowych, ale dwie z nich są rozplecione w połowie swej długości i krzyżują się ze sobą w tym miejscu (Ryc. 4c). Odpowiednie aminokwasy w



Ryc. 3. Dwa podstawowe typy architektury białek błonowych.

(a) wiązka α -helis na przykładzie bakteriorodopsyny (kod 1BRR w bazie PDB); (b) β -beczka na przykładzie białka OmpA (kod 1BXW w bazie PDB).



Ryc. 4. Struktury białek błonowych ze specyficznymi helisami.

(a) rodopsyna (kod 1U19 w bazie PDB), ciemnym wyróżniono helisy TM6, TM7 i TM1 (w kolejności od lewej do prawej); (b) akwaporyna (kod 2F2B w bazie PDB); (c) antyporter Na^+/H^+ (kod 1ZCD w bazie PDB), niektóre pętle i helisy zostały usunięte dla przejrzystości rysunku.

tym rejonie białka zapewniają selektywny transport, a więc warunkują jego prawidłową funkcję.

W bazie PDB jest zdeponowanych tylko kilka struktur ludzkich białek błonowych. Wiąże się to z niestabilnością termiczną tych białek, w nieobecności błon komórkowych. Dużo łatwiej uzyskuje się kryształy białek bakteryjnych, lecz większość białek błonowych w komórkach eukariotycznych nie ma swoich odpowiedników (homologów) bakteryjnych. Jeśli więc czekanie na odpowiednie struktury krystaliczne trwa zbyt długo, można użyć metod teoretycznych bazując na strukturze właściwego białka uzyskanego w niskiej rozdzielczości. Obie metody mogą być zawodne, jak to pokazano na powyższych przykładach. Dlatego też dla sprawdzenia poprawności teoretycznego modelu struktury białka można zaprojektować odpowiedniego mutanta, dającego przewidywalny efekt, dla sprawdzenia w badaniach biochemicznych.

Metody teoretyczne uzyskiwania struktur białek błonowych można podzielić na dwie główne klasy: metody modelowania przez homologię i metody ab-initio (zwane też modelowaniem first-principle) – podobnie jak dla białek cytoplazmatycznych. W obu metodach należy rozpocząć od odszukania w sekwencji nowego białka jego części błonowej czyli *de-facto* od znalezienia topologii. Jeśli znamy białko homologiczne do białka badanego, to zwykle wystarczy wykonać dopasowanie obu sekwencji. Odpowiednie, silnie hydrofobowe obszary, świadczące o helisach

transbłonowych, zostaną nałożone na siebie automatycznie i w ten sposób zidentyfikowana zostanie część błonowa. Aminokwasy należące do tej części także ulegają mutacjom w czasie ewolucji, lecz aby zatrzymać białko w błonie mutacje te muszą być konserwatywne, tzn. dawać w wyniku także aminokwasy hydrofobowe, dzięki czemu fragmenty transbłonowe są bardzo charakterystyczne w sekwencji. Do znajdowania tych fragmentów w metodach ab-initio służą metody, takie jak np. TopPred^{AG} czy SCAMPI (A Scale-Based Method for Prediction of Integral Membrane Proteins). Ponieważ, jak wspomniano wcześniej, niektóre helisy transbłonowe mogą być o niepełnej długości (mniej niż 20 aminokwasów, co odpowiada 30 Å szerokości błony komórkowej), wyniki tych przewidywań mogą być niekiedy mylące (BERNSEL i współaut. 2008).

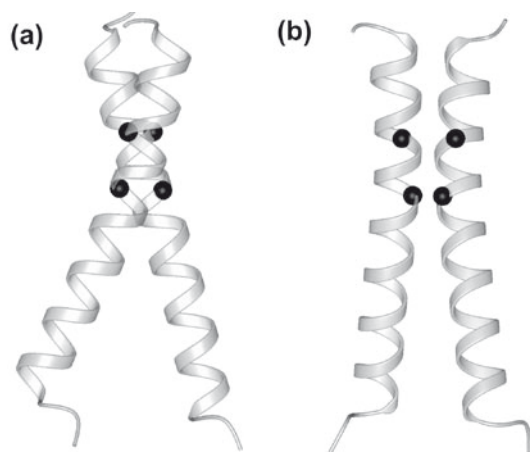
Po wykonaniu dopasowania w metodzie modelowania przez homologię następuje przepisanie współrzędnych struktury wzorca homologicznego białka na strukturę szukaną; jeśli występują różne aminokwasy, to są one zmieniane, lecz z zachowaniem położenia atomów C_α , C_β i dalszych, jeśli także się pokrywają. W ten sposób struktura całej części błonowej nowego białka jest gotowa. Tak uzyskana struktura jest zwykle poprawna i to nawet przy 20–30% podobieństwie sekwencji (zakładając, że wzorec i badane białko należą do tej samej rodziny). Dzieje się tak dlatego, że struktura białka jest zachowana w toku ewolucji bardziej niż jego sekwencja. Bez odpowiedniej struktury białko

nie mogłoby pełnić prawidłowo swojej funkcji, podczas gdy tylko kilka aminokwasów w miejscu aktywnym enzymu lub w miejscu wiążącym ligand w receptorze musi pozostać bez zmian. W przypadku receptorów GPCR, nawet miejsce wiążące ligand może zmieniać się podczas ewolucji (powstają wtedy nowe typy tych receptorów), natomiast wiązanie z białkiem G, przenoszącym sygnał na białka efektorowe, musi pozostać bez zmian. Stąd też znanych jest kilkaset typów receptorów GPCR, zaś tylko kilkanaście typów białek G. Gdy część błonowa jest już gotowa, pozostaje tylko (albo aż) wykonać modelowanie pętli zewnątrzkomórkowych oraz N- i C-końcowych rejonów. Jeśli sekwencje dla tych rejonów są krótkie, to większość programów do modelowania homologicznego (np. Modeller, SwissModel czy ostatnio także Yasara) upora się z tym problemem. W przypadku dłuższych sekwencji trzeba posłużyć się metodami modelowania *ab-initio* stosując je do wybranych, pozabłonowych rejonów białka.

Głównym czynnikiem kształtującym rejon błonowy białek jest błona komórkowa, czyli oddziaływanie z fosfolipidami. Do modelowania tych rejonów metodami *ab-initio* nie można było więc użyć metod wykorzystywanych dla białek rozpuszczalnych, ponieważ nie uwzględniały tego rodzaju oddziaływań. W związku z tym opracowano nowe metody, do wyznaczania struktur poszczególnych klas białek błonowych (wciąż nie ma ogólnej metody do wszystkich typów białek błonowych). Ponieważ najważniejszym typem oddziaływań w błonie jest oddziaływanie typu helisa-helisa, większość tych metod skupia się na przewidywaniu ułożenia helis w wiązce transbłonowej. Oddziaływania pomiędzy helisami pełnią krytyczną rolę w zwijaniu białek błonowych, ich stabilności oraz funkcji. Na poziomie molekularnym są one realizowane przez jeden lub kilka kontaktów pomiędzy resztami z różnych helis. Metody do modelowania wzajemnego ustawienia tych helis opierają się zwykle na wzorcach układania helis, aby wzajemne nakładanie łańcuchów bocznych było jak najmniejsze (tzw. knob-into-hole i ridge-into-groove) (LANGOSCH i HERRINGA 1998). W innych metodach poszukuje się specjalnych motywów na helisach, które mają ułatwiać pakowanie (WALTERS i DEGRADO 2006). Można także zacząć od poszukiwania kontaktów pomiędzy helisami na podstawie sekwencji. W metodzie TMhit (<http://biocluster.iis.sinica.edu.tw/TMhit/>) używa się do tego celu zestawu klasyfikatorów (jest to me-

toda typu machine-learning), natomiast ich wiedza budowana jest stopniowo w czasie fazy treningowej z informacji sekwencyjnej i strukturalnej na zestawie kilkudziesięciu białek błonowych. Zdolność poprawnej klasyfikacji przetestowano na niezależnym zestawie tych białek. Uzyskane kontakty są następnie wykorzystywane do wzajemnego umiejscowienia dwu helis. Na obecnym etapie rozwoju metod *ab-initio* potrafią one trafnie przewidzieć wzajemne ułożenie tylko czterech helis transbłonowych.

Białka, których część błonowa składa się z wiązki α -helis, charakteryzują się bardzo dużą różnorodnością funkcji biologicznych. Z tego powodu zrozumienie natury oddziaływań jakie istnieją między helisami oraz sił, które powodują właściwe składanie tych helis w procesie zwijania białka ma olbrzymie znaczenie. Niestety, jesteśmy dopiero na początku drogi do pełnego zrozumienia tych procesów. I tak, chociaż zdefiniowano wymagania strukturalne dla oddziaływania dwu helis transbłonowych oraz odkryto motywy sekwencyjne sprzyjające dimeryzacji, wiele kwestii pozostaje wciąż otwartych (SCHNEIDER i współaut. 2007). Okazuje się, że istnieją podobieństwa pomiędzy parami regularnych helis transbłonowych oraz odpowiednich par helis w białkach pozabłonowych dla parametrów takich jak długość helis, powierzchnia kontaktu, kąt pakowania, czy odległość pomiędzy środkami helis; chociaż już dla nieregularnych helis takich podobieństw brak. Większa powierzchnia kontaktu oraz mniejsza odległość pomiędzy środkami helis jest warunkowana przez obecność zwiększonej liczby małych aminokwasów, czyli alanin i glicyn, szczególnie jeśli są one zlokalizowane po jednej stronie α -helisy, tj. w motywach AxxxA, AxxxG i GxxxG (oddalone o 4 aminokwasy), bowiem skok helisy wynosi też ok. 4 aminokwasy (Ryc. 5). Helisy w białkach błonowych nie są upakowane w większym stopniu niż helisy białek pozabłonowych, natomiast są w stanie tworzyć ściślejsze pary helis. To powoduje, że helisy mogą pozostawać w kontakcie przez prawie całą szerokość błony komórkowej. Ścisłe podobieństwo strukturalne pomiędzy niektórymi parami helis w białkach błonowych i pozabłonowych (rozpuszczalnych) wskazuje, że w pewnych przypadkach metody *ab-initio* dla białek rozpuszczalnych mogą być zastosowane również do modelowania upakowania białek błonowych. Wydaje się także, że nie ma zasadniczych różnic pomiędzy wielkościami sił wpływających na pakowanie helis w



Ryc. 5. Struktura błonowej części dimeru glikoforyny A (GpA) oparta na motywie GxxxG (kod 1AFO w bazie PDB); zaznaczono atomy węgla C_{α} glicyn tworzących interfejs.

(a) widok na skrzyżowanie helis; (b) po obrocie o 90° .

obu typach tych białek (GIMPELEV i współaut. 2004). Na pakowanie helis białek błonowych ma niewątpliwie wpływ błona komórkowa, bowiem aminokwasy hydrofobowe muszą znajdować się na zewnątrz białka błonowego, co jest niemożliwe w przypadku białek rozpuszczalnych. Tym niemniej wewnątrz białek błonowych jest jeszcze bardziej hydrofobowe niż ich powierzchnia w części błonowej, a więc reguły pakowania białek błonowych nie ulegają odwróceniu w stosunku do białek rozpuszczalnych. W związku z tym programy do modelowania *ab-initio* dla białek pozabłonowych z wykorzystaniem doświadczalnych danych dotyczących części struktury (np. więzy na odległości między pewnymi resztami aminokwasowymi) mogą generować prawidłowe struktury białek błonowych. Przeszkodą jest oczywiście wielkość białek błonowych, które z reguły mają większą masę cząsteczkową niż białka rozpuszczalne (FLEISHMAN i BEN-TAL 2006).

SYMULACJE DYNAMIKI BIAŁEK BŁONOWYCH

Metody teoretyczne znalazły bardzo szerokie zastosowanie w badaniu dynamiki białek w błonach komórkowych. W przypadku reprezentacji pełnoatomowych (uwzględniając metodę zjednoczonych atomów) opracowano, a właściwie rozszerzono o parametry dla fosfolipidów, metody służące do mechaniki i dynamiki molekularnej białek rozpuszczalnych, takie jak NAMD, GROMACS, CHARMM czy AMBER. Ponieważ symulacje dynamiki molekularnej, nawet dla małych białek błonowych, muszą zawierać błonę komórkową, czasy tych symulacji nie są imponujące – zwykle kilka do kilkudziesięciu nanosekund. Metody doświadczalne do badania dynamiki białek błonowych w tak krótkich odcinkach czasowych jeszcze nie istnieją, dlatego wyniki symulacji porównuje się do statycznych struktur przed i po zajściu procesu np. przed i po aktywacji receptora.

Procesy aktywacji/deaktywacji w białkach błonowych trwają znacznie dłużej niż obecnie możliwe najdłuższe symulacje a mianowicie, od ok. 1 milisekundy do kilku sekund, tym niemniej poszczególne etapy tego procesu dają się już śledzić metodami teoretycznymi. Wynika to z faktu, że przejścia pomiędzy kolejnymi etapami aktywacji receptora mogą być bardzo szybkie i wiążą się np. z zerwaniem pojedynczego wiązania wodorowego (ewentualnie oddziaływania jonowego czyli

mostka solnego) lub też ze zmianą konformacji jednego aminokwasu (np. Trp6.48 na helisie TM6 w receptorach GPCR). Takie kluczowe dla aktywacji reszty aminokwasowe w receptorze tworzą tzw. przełączniki molekularne. Chociaż występują w większości receptorów z tej samej rodziny, to nie we wszystkich receptorach następuje zmiana stanu przełącznika pod wpływem związania liganda z miejscem wiążącym w receptorze. Okazuje się, że różne ligandy mogą aktywować różne przełączniki (KOBILKA i DEUPI 2007) i w konsekwencji prowadzić do odmiennego sposobu aktywacji receptora. W przypadku receptorów GPCR związanie odpowiedniego liganda prowadzi do aktywacji białka G_s (stymulatorowego) lub G_i (inhibitorowego). Ostatnio wykazano, że te same receptory GPCR mogą także uczestniczyć w przekazywaniu sygnału bez pośrednictwa białka G, a mianowicie przez arestynę. Arestyna była długo uważana za białko, które tylko wygasza aktywne receptory – stąd jej nazwa. Wygaszanie ma na celu umożliwienie powtarzania sygnalizacji, gdyż inaczej receptor byłby aktywny przez bardzo długi czas, a to spowodowałoby chaos w przekazywaniu sygnału w komórce. Pobudzenie arestyny do sygnalizacji wiąże się oczywiście z odmiennym ustawieniem przełączników molekularnych przez związany ligand. Pewną ciekawostką stanowi fakt, że

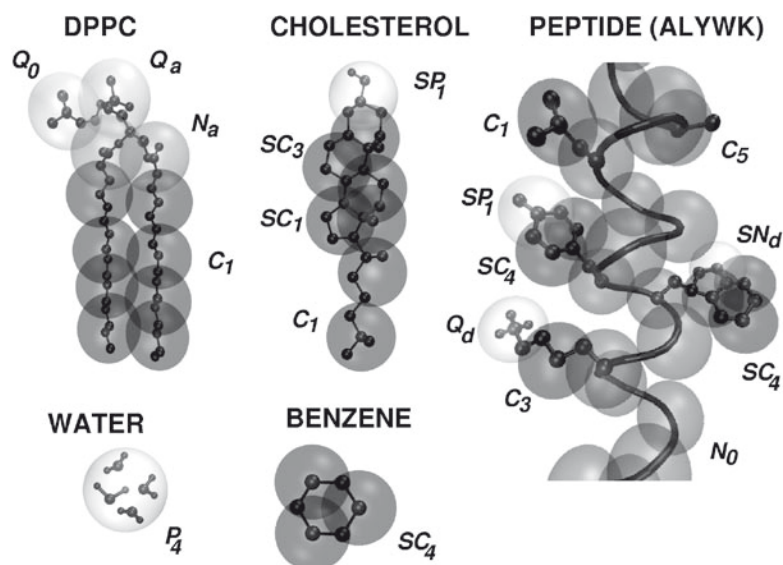
w otrzymanych ostatnio strukturach receptorów GPCR (z antagonistami i odwrotnymi agonistami), mostek solny związany z motywem E/DRY na helisie TM3 jest zerwany. Teoretycznie powinno to umożliwić związanie białka G (tak sugerowano na podstawie struktury rodopsyny), a tymczasem receptor jest nieaktywny. Być może taka struktura jest predysponowana do sygnalizacji komórkowej właśnie przez arestynę.

Niektóre pełnoatomowe metody symulacyjne zostały efektywnie zrównoleglone i jeśli są uruchomione na kilku tysiącach procesorów, to uzyskiwane czasy symulacji wchodzą w zakres mikrosekundowy. Ostatnio wykonano takie pełnoatomowe symulacje dla receptora β_2 AR z rodziny GPCR w programie DESMOND z polem siłowym CHARMM (DROR i współaut. 2009). Wykazano, że przełącznik molekularny związany z motywem E/DRY może występować w nieaktywnej postaci tego receptora w dwu pozycjach, niezależnie od związanego liganda. Przeprowadzono także symulacje dla antyportera Na^+/H^+ wykazując mechanizm jego działania, a w szczególności efekt protonacji centralnych reszt kwasowych (ARKIN i współaut. 2007). Wyniki teoretyczne zostały potwierdzone przez skonstruowanie i zbadanie odpowiednich mutantów. Z kolei stosując pełnoatomowe symulacje dynamiki molekularnej dla okta-meru akwaporyny na złączeniu dwu błon komórkowych pokazano, że to złączenie nie wpływa na transport wody przez akwaporyny (JENSEN i współaut. 2008).

Aby zbadać szczególne efekty działania białek, łączy się czasami różne metody ze sobą. Jeśli chcemy prześledzić działanie enzymu na substrat, czyli obserwować tworzenie

lub rozpad wiązań chemicznych, to część białka (miejsce aktywne enzymu) należy opisać kwantowo-chemicznie, natomiast pozostałą część układu klasycznie. Takie metody noszą nazwę QM/MM (ang. quantum mechanics/molecular mechanics). Można także działać w drugą stronę: aby zwiększyć długość symulacji w pracy (NERI i współaut. 2008) wykorzystano połączenie metody pełnoatomowej (pole siłowe z programu GROMOS) z metodą gruboziarnistą (model typu Go – jedno ziarno reprezentuje cały aminokwas). Metodą pełnoatomową opisano miejsce aktywne enzymu (bakteryjna proteaza błonowa OmpT), natomiast całą resztę modelem Go. W ten sposób udało się osiągnąć czas symulacji 1 μs zamiast 10–100 ns w typowych symulacjach pełnoatomowych. Takie modelowanie typu MM/CG (ang. molecular mechanics/coarse-grain) pozwoliło na zaobserwowanie długookresowych fluktuacji białka podobnych do fluktuacji w proteazach aspartylowych.

Metody gruboziarniste zastosowane do całości układu pozwalają na wydłużenie czasów symulacji układów błonowych o kolejny rząd wielkości (powyżej 10 mikrosekund). Taką metodą jest pole siłowe MARTINI w programie GROMACS. W tej metodzie wykorzystuje się mapowanie 4:1 tzn. cztery ciężkie atomy na jedno ziarno. W ten sposób łańcuch lipidowy o długości 16 atomów węgla (49 atomów w reprezentacji pełnoatomowej i 16 atomów w reprezentacji zjednoczonych atomów) jest reprezentowany przez 4 ziarna. Molekuły wody nie są pomijane w obliczeniach, ale także podlegają skalowaniu tj. zamiast 4 cząsteczek wody otrzymuje się jedno ziarno rozpuszczalnika (Ryc. 6). Ponieważ stosowane wzory na obliczanie oddziaływań



Ryc. 6. Przykłady mapowania struktur pełnoatomowych na odpowiednie modele gruboziarniste w polu siłowym MARTINI (rysunek ze strony internetowej <http://md.chem.rug.nl/~marrink/coarsegrain.html>). Każdemu rodzajowi ziarna przypisany jest typ (nazwy podane obok ziaren) warunkujący odpowiednie oddziaływanie.

(wiązących i niewiązących) są identyczne jak w metodach pełnoatomowych, zmienia się tylko zestaw parametrów, zatem można stosować te same programy do obliczeń typu CG co do obliczeń pełnoatomowych. Pole siłowe MARTINI zostało skonstruowane w celu symulacji zachowania błon komórkowych w długich skalach czasowych. Dzięki niemu uzyskano ciekawe wyniki tworzenia domen w mieszanej błonie złożonej z fosfolipidów z nasyconymi i nienasyconymi łańcuchami lipidowymi (diC₁₆-PC/diC_{18:2}-PC) oraz cholesterolu. Domeny zaczynały tworzyć się już po czasie 3 μ s. Cholesterol przemieszczał się do domen z nasyconymi łańcuchami lipidowymi i porządkował ich ułożenie w błonie (RISSELADA i MARRINK 2008). Udało się także wymodelować przenikanie fullerenów przez błonę komórkową i rozpraszanie się w niej. Nawet znaczne ilości fullerenów nie powodowały rozpadu błony (WONG-EKKABUT i współaut. 2008). Dodanie parametrów dla aminokwasów (MONTICELLI i współaut. 2008) znacznie rozszerzyło możliwości tej metody i umożliwiło badanie dynamiki tworzenia kompleksów błonowych, np. powstawanie kanału błonowego złożonego z czterech jednostek białka M2 wirusa grypy A, badania deformacji błon pod wpływem białek błonowych i toksyn oraz tworzenia oligomerów białek błonowych. Z powodu bardziej wygładzonej hiperpowierzchni energii potencjalnej (znacznie mniej atomów) dochodzenie układu do stanu równowagi jest dużo szybsze niż w metodach pełnoatomowych. Z tego też powodu w metodach gruboziarnistych nie obserwuje się zmian konformacyjnych, widocznych gdy wykona się wiele krótszych, ale pełnoatomowych symulacji. Z kolei niektóre przejścia wymagają znacznie dłuższych skal czasowych dostępnych w metodach CG.

Zmiany zachodzące w białku można także wywoływać poprzez zastosowanie zewnętrznej siły zarówno w metodach doświadczalnych, jak i teoretycznych. W ten sposób można osiągnąć np. efekt mechanicznego rozwijania białka, wyciągając go z błony za jeden z końców. Korzystając z metody doświadczalnej, spektroskopii sił dla pojedynczych molekuł (ang. Single Molecule Force Spectroscopy, SMFS) (SAPRA i współaut. 2006) otrzymuje się zestaw krzywych siła-przesunięcie (ang. Force-Distance, F-D), z których można wnioskować o położeniu mikrodomen (nawet w białkach jednodomenowych) warunkujących stabilność w białku oraz badać zmiany stabilizacji białka pod wpływem mutacji lub

czynników zewnętrznych np. temperatury, pH czy jonów metali (PARK i współaut. 2007). W przypadku rozwijania bakteriorodopsyny lub innych białek siedmiohelikalnych (GPCR) otrzymuje się trzy duże piki sił, co może wskazywać, że podczas rozwijania białka helisy transbłonowe są wyciągane z błony parami. Niestety, w tej metodzie nie ma możliwości bezpośredniego potwierdzenia tego czy innych wniosków. Z pomocą przychodzą metody teoretyczne, pozwalające zastosować zewnętrzną siłę do rozwijania modelu białka w modelu błony komórkowej. Podobnie jak w metodach doświadczalnych, otrzymane teoretyczne krzywe są różne dla tego samego białka rozwijanego w identycznych warunkach. Wskazuje to na istnienie różnych dróg rozwijania tego samego białka, które jest dynamiczną strukturą i w danym momencie wybierana jest ta droga, która z pewnych powodów jest optymalna (wymaga przyłożenia najmniejszej siły). Ponieważ białko ulega rozwijaniu w trakcie symulacji, wymagane jest użycie bardzo dużych układów, aby po rozwinięciu mieściło się wciąż w opisywanej komórce (ang. periodic-box) co uniemożliwia zastosowanie metod pełnoatomowych. Korzystając z metod uproszczonych wykonano rozwijanie bakteriorodopsyny (SEEBER i współaut. 2006), z modelu oligomeru otrzymując krzywe F-D jakościowo zgodne z krzywymi doświadczalnymi. W pracy tej zastosowano metodę, w której zarówno woda, jak i błona były ośrodkami ciągłymi (metoda IMM1 w programie CHARMM). Z kolei wykorzystując zmodyfikowaną metodę modelu Go (ang. Go-like model) wykonano symulowane mechaniczne rozfałdowanie bakteriorodopsyny i halorodopsyny w różnych temperaturach (CIEPLAK i współaut. 2006) (Ryc. 7). Metody oparte na modelu Go różnią się od metod mechaniki i dynamiki molekularnej, bowiem są oparte na sieci połączeń pomiędzy najbliższymi (w sensie przestrzennym a nie tylko po wiązaniach) atomami C _{α} badanego białka. Nie jest to jednak nierozzerwalna struktura, jak np. w modelach sieci elastycznych, bowiem połączenia te nie są opisywane potencjałem harmonicznym i uwzględniają dysocjację wiązania przy odpowiednim zwiększeniu odległości pomiędzy atomami. Chociaż szybkość rozciągania białka w metodach symulacyjnych jest wciąż kilka rzędów wielkości (5-7) większa niż szybkość w metodach doświadczalnych, to jakościowa zgodność uzyskanych krzywych F-D z doświadczeniem pozwala mieć nadzieję na poprawność



Ryc. 7. Struktura bakteriorodopsyny w trakcie symulowanego wyciągania białka z błony przy ciągnięciu za N-koniec w metodzie modelu Go. Błona została pominięta dla przejrzystości rysunku. Dwie pierwsze helisy uległy wyciągnięciu z błony i rozwijają się przy niezmięnionej strukturze pozostałej części białka.

interpretacji i przewidywań wpływu różnych czynników na stabilność białek błonowych.

Intensywny rozwój zarówno metod doświadczalnych, jak i teoretycznych badania białek błonowych w ostatnich kilku latach sprawił, że białka te stopniowo odkrywają przed nami swoje tajemnice. Wielkość białek błonowych powoduje, że przy użyciu samych tylko metod teoretycznych nie jesteśmy w stanie wymodelować poprawnych ich struktur, a czasami nawet określić liczby helis transbłonowych. Jednak, gdy struktura białka błonowego jest dostępna, metody symulacyjne pozwalają na zbadanie zmian konformacyjnych w białkach w różnych skalach czasowych (im dłuższy czas symulacji, tym zastosowana metoda jest mniej dokładna), co jest często niemożliwe w przypadku stosowania metod doświadczalnych. Wzajemne uzupełnianie się obu rodzajów metod może pozwolić na przewidywanie struktur białek błonowych i ich kompleksów z ligandami czy też wpływu mutacji, jak też na pełniejszą analizę obserwowanych zjawisk.

STRUCTURAL BIOINFORMATICS OF MEMBRANE PROTEINS

Summary

Our genome is composed of 20–30 % of membrane proteins but number of structures of these proteins known and deposited in public databases is still small. However, new achievements in experimental techniques, especially microfocusing of X-ray beam enabling diffracting of microcrystals, as well as mutagenesis leading to obtaining of thermostable mutants are real hope for quick emerging of new structures. Theoretical methods for determination of structure of membrane proteins are still in infant phases. Usage of homology modeling is limited by

small number of membrane proteins which are necessary to serve as templates whereas ab-initio methods are confined to predicting of small membrane proteins or parts of larger ones only. The area which the bioinformatics is foremost in is prediction of dynamical behavior of proteins in lipid bilayer which is still mostly inaccessible to experimental methods. Full-atom as well as coarse-grain molecular dynamics methods are used to describe investigated systems in different time scales and with different accuracy.

LITERATURA

- ARKIN I. T., XU H. F., JENSEN M. O., ARBELY E., BENNETT E. R., BOWERS K. J., CHOW E., DROR R. O., EASTWOOD M. P., FLITMAN-TENE R., GREGERSEN B. A., KLEPEIS J. L., KOLOSSVARY I., SHAN Y. B., SHAW D. E., 2007. *Mechanism of Na⁺/H⁺ antiporting*. *Science* 317, 799–803.
- BERNSEL A., VIKLUND H., FALK J., LINDAHL E., VON HEIJNE G., ELOFSSON A., 2008. *Prediction of membrane-protein topology from first principles*. *Proc. Natl. Acad. Sci. USA* 105, 7177–7181.
- CIEPLAK M., FILIPEK S., JANOVJAK H., KRZYSKO K. A., 2006. *Pulling single bacteriorhodopsin out of a membrane: Comparison of simulation and experiment*. *Biochim. Biophys. Acta-Biomembr.* 1758, 537–544.
- DROR R. O., ARLOW D. H., BORHANI D. W., JENSEN M. O., PIANA S., SHAW D. E., 2009. *Identification of two distinct inactive conformations of the beta(2)-adrenergic receptor reconciles structural and biochemical observations*. *Proc. Natl. Acad. Sci. USA* 106, 4689–4694.
- FLEISHMAN S. J., BEN-TAL N., 2006. *Progress in structure prediction of alpha-helical membrane proteins*. *Curr. Opin. Struct. Biol.* 16, 496–504.
- FLEISHMAN S. J., UNGER V. M., BEN-TAL N., 2006. *Transmembrane protein structures without X-rays*. *Trends Biochem. Sci.* 31, 106–113.
- GIMPELEV M., FORREST L. R., MURRAY D., HONIG B., 2004. *Helical packing patterns in membrane and soluble proteins*. *Biophys. J.* 87, 4075–4086.
- JENSEN M. O., DROR R. O., XU H. F., BORHANI D. W., ARKIN I. T., EASTWOOD M. P., SHAW D. E., 2008. *Dynamic control of slow water transport by aquaporin 0: Implications for hydration and junction stability in the eye lens*. *Proc. Natl. Acad. Sci. USA* 105, 14430–14435.
- KOBILKA B. K., DEUPI X., 2007. *Conformational complexity of G-protein-coupled receptors*. *Trends Pharmacol. Sci.* 28, 397–406.
- LANGOSCH D., HERINGA J., 1998. *Interaction of transmembrane helices by a knobs-into-holes packing*

- characteristic of soluble coiled coils.* Proteins 31, 150–159.
- MONTICELLI L., KANDASAMY S. K., PERIOLE X., LARSON R. G., TIELEMAN D. P., MARRINK S. J., 2008. *The MARTINI coarse-grained force field: Extension to proteins.* J. Chem. Theory Comput. 4, 819–834.
- NERI M., BAADEN M., CARNEVALE V., ANSELMINI C., MARI-TAN A., CARLONI P., 2008. *Microseconds dynamics simulations of the outer-membrane protease T.* Biophys. J. 94, 71–78.
- PALCZEWSKI K., KUMASAKA T., HORI T., BEHNKE C. A., MOTOSHIMA H., FOX B. A., LE TRONG I., TELLER D. C., OKADA T., STENKAMP R. E., YAMAMOTO M., MIYANO M., 2000. *Crystal structure of rhodopsin: A G protein-coupled receptor.* Science 289, 739–745.
- PARK P. S. H., SAPRA K. T., KOLINSKI M., FILIPEK S., PALCZEWSKI K., MULLER D. J., 2007. *Stabilizing effect of Zn²⁺ in native bovine rhodopsin.* J. Biol. Chem. 282, 11377–11385.
- RISSELADA H. J., MARRINK S. J., 2008. *The molecular face of lipid rafts in model membranes.* Proc. Natl. Acad. Sci. USA 105, 17367–17372.
- SAPRA K. T., PARK P. S. H., FILIPEK S., ENGEL A., MULLER D. J., PALCZEWSKI K., 2006. *Detecting molecular interactions that stabilize native bovine rhodopsin.* J. Mol. Biol. 358, 255–269.
- SCHNEIDER D., FINGER C., PRODOHL A., 2007. *From interactions of single transmembrane helices to folding of alpha-helical membrane proteins: Analyzing transmembrane helix-helix interactions in bacteria.* Curr. Protein Pept. Sci. 8, 45–61.
- SEEBER M., FANELLI F., PACI E., CAFLISCH A., 2006. *Sequential unfolding of individual helices of bacterioopsin observed in molecular dynamics simulations of extraction from the purple membrane.* Biophys. J. 91, 3276–3284.
- SERRANO-VEGA M. J., MAGNANI F., SHIBATA Y., TATE C. G., 2008. *Conformational thermostabilization of the beta 1-adrenergic receptor in a detergent-resistant form.* Proc. Natl. Acad. Sci. USA 105, 877–882.
- WALTERS R. F. S., DEGRADO W. F., 2006. *Helix-packing motifs in membrane proteins.* Proc. Natl. Acad. Sci. USA 103, 13658–13663.
- WONG-EKKABUT J., BAOUKINA S., TRIAMPO W., TANG I. M., TIELEMAN D. P., MONTICELLI L., 2008. *Computer simulation study of fullerene translocation through lipid membranes.* Nat. Nanotechnol. 3, 363–368.