

# Chemiczne biblioteki kombinatoryczne — ich synteza i zastosowania

Piotr Kwaśnikowski  
Wojciech T. Markiewicz  
Instytut Chemii Bioorganicznej  
Polska Akademia Nauk  
Poznań

## 1. Wprowadzenie

Oddziaływania międzycząsteczkowe stanowią podstawę regulacji wszystkich procesów metabolicznych zachodzących w organizmach żywych. Poznanie zależności między strukturą a funkcją biologiczną związków ma zasadnicze znaczenie dla zrozumienia istoty procesów życiowych, ale równie istotne jest dla badań o aspekcie praktycznym, np. poszukiwania nowych leków czy poznania mechanizmu ich działania.

Środki terapeutyczne wpływają np. na metabolizm komórek. Często charakteryzują się dużą niespecyficnością aktywności biologicznej co prowadzi do niepożądanych efektów ubocznych. Specyficzność działania leków można uzyskać przez odpowiedni dobór ich pochodnych rozpuszczalnych w wodzie lub lipidach, uwzględnienie dróg ich biodegradacji i ścisłą kontrolę ostatecznego stężenia leku w ustroju. Wyjątkiem są terapeutyki wykazujące receptorowy mechanizm działania lub pobierane przy udziale receptorów, chociaż nawet wtedy mamy do czynienia zazwyczaj z więcej niż jednym rodzajem komórek poddawanych działaniu tego samego leku.

W poszukiwaniu nowych środków terapeutycznych zasadniczym problemem jest znalezienie takiej cząsteczki chemicznej, która wykazywałaby maksymalną specyficzność komórkową, tkankową lub organową i dawałaby minimalne efekty uboczne. Rozwój badań modelowych i metod teoretycznych znacznie przybliżył nas do skutecznego racjonalnego projektowania nowych terapeutyków. Wydaje się jednak mało prawdopodobne, aby metody czysto teoretyczne mogły w najbliższym czasie zastąpić tradycyjne sposoby poszukiwania leków.

Zasadniczym problemem staje się znajdowanie coraz to nowych związków chemicznych, tzw. związków wiodących (ang. *lead*), z których w toku dalszych badań można wytworzyć nowe terapeutyki. Naprzeciw tym potrzebom wycho-

dzą badania złożonych mieszanin związków chemicznych, w których mogą znajdować się cząsteczki o pożądanej aktywności, oparte na wykorzystaniu koncepcji **różnorodności molekularnej** (ang. *molecular diversity*). Nowe podejście opiera się na wykorzystaniu syntetycznych kombinatorycznych bibliotek związków chemicznych i zostało zainicjowane osiągnięciami biologii molekularnej. Wiąże się z nim nadzieje na osiągnięcie znacznego postępu w poszukiwaniach np. nowych leków, ale także w wyjaśnieniu molekularnych podstaw procesów życiowych.

W omówieniu tym przedstawione zostały najważniejsze podejścia metodyczne związane z koncepcją syntetycznych kombinatorycznych bibliotek związków chemicznych i bibliotek otrzymywanych z wykorzystaniem układów biologicznych (fagowe biblioteki peptydów i przeciwciał).

## 2. Syntetyczne kombinatoryczne biblioteki związków chemicznych

Pod pojęciem biblioteki kombinatorycznej rozumiemy zbiór cząsteczek podobnych w budowie. Bibliotekę taką bada się pod kątem wyselekcjonowania jej aktywnych składników w specyficzny sposób oddziałujących z wybraną cząsteczką docelową. Z definicji wynika, że biblioteka kombinatoryczna może być zbudowana z dowolnego typu monomerów. Jednak fakt, że w przyrodzie podstawowymi monomerami są aminokwasy i nukleotydy spowodował, iż właśnie z nich były zbudowane pierwsze syntetyczne biblioteki kombinatoryczne.

Na przykładzie heksapeptydów zbudowanych z wszystkich podstawowych aminokwasów białkowych (20) widzimy, że możliwa liczba peptydów o różnych sekwencjach jest olbrzymia i wynosi  $20^6$  czyli 64 000 000. Liczba możliwych sekwencji rośnie eksponencjalnie z długością łańcucha  $n$  i zależy od liczby monomerów  $m$  zgodnie z wzorem  $m^n$ .

Wybór dokonany w przyrodzie pod względem cząsteczek kodujących i realizujących informację biologiczną, tzn. kwasów nukleinowych i białek (peptydów) skłania do szczególnego traktowania bibliotek peptydowych i oligonukleotydowych. Mogą mieć one potencjalnie najszersze zastosowanie zarówno w badaniach podstawowych jak i stosowanych, np. w poszukiwaniach związków wiodących (1).

## 3. Podział syntetycznych kombinatorycznych bibliotek związków chemicznych

Podziału syntetycznych bibliotek kombinatorycznych można dokonać biorąc pod uwagę następujące kryteria: a) typ związku; b) kompletność biblioteki; c) homogeniczność biblioteki; d) format biblioteki; e) „czytanie” sekwencji **elementu** biblioteki; f) synteza biblioteki.



Biblioteka może być kompletna czyli zawierać pełną reprezentację wszystkich możliwych struktur (**elementów** biblioteki) związków danego typu. Biblioteka **homogeniczna** składa się ze związków zbudowanych z tej samej liczby **monomerów** (ang. *building blocks*) w odróżnieniu od biblioteki **heterogenicznej**, w której znajdują się związki zbudowane z różnej liczby monomerów. Ze względu na format biblioteki można podzielić na **zintegrowane** (ang. *integrated*) np. biblioteki **dwuwymiarowe** (2,3) (ang. *arrays*) i biblioteki **rozproszone** (ang. *dispersed*). Biblioteki rozproszone w skrajnym przypadku składają się z pojedynczych cząsteczek (4,5). Ostateczna przydatność biblioteki kombinatorycznej zależy od możliwości odczytania struktury (sekwencji) każdego związku biblioteki znalezionej w procesie przeszukiwania biblioteki (**selekcja**, ang. *selection, screening*). Biblioteki **zintegrowane** nie wymagają sekwencjonowania, ponieważ przestrzenna lokalizacja związku w bibliotece (współrzędne) określa z definicji jego budowę (historia syntezy). W przypadku bibliotek **rozproszonych** odczytywanie struktury związków jest konieczne. Bezpośrednie sekwencjonowanie jest możliwe obecnie tylko dla niektórych związków (peptydów lub fragmentów kwasów nukleinowych — np. oligodeoksyrybonukleotydów). Czytanie sekwencji w sposób pośredni odbywa się na zasadzie odczytywania **etykiat kodujących** (4,6,7) (ang. *tag*) lub poprzez sekwencjonowanie kopii związku.

#### 4. Otrzymywanie i zastosowania bibliotek kombinatorycznych

Szersze omówienie zintegrowanych bibliotek kombinatorycznych (arrays), nie będących przedmiotem tego przeglądu można znaleźć w (8).

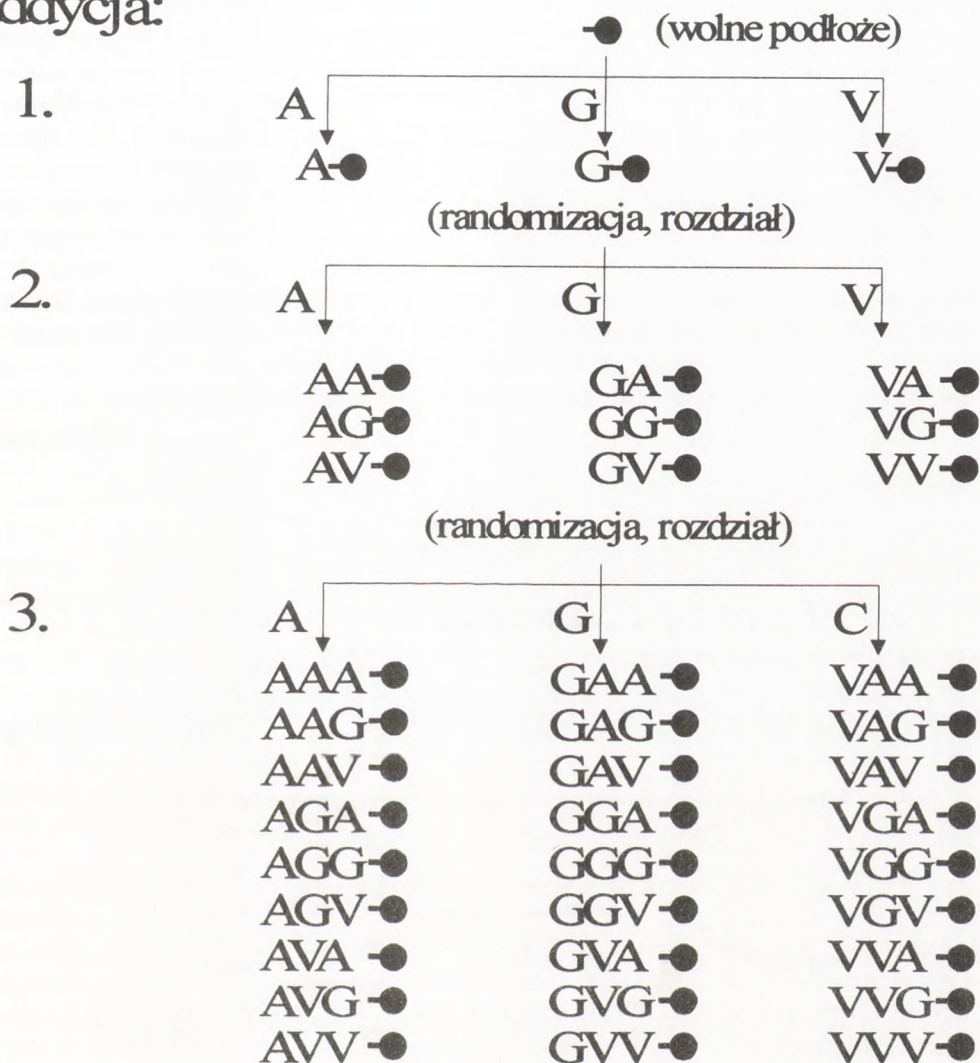
Kombinatoryczne biblioteki rozproszone można podzielić ze względu na sposób „odczytywania” (sekwencjonowania po selekcji) na **kodowane** (ang. *tagged*) i **niekodowane** (ang. *untagged*) (9). Określenia te nie odpowiadają w pełni terminom angielskim, częściowo z powodu braku pełnego uporządkowania angielskiego nazewnictwa bibliotek kombinatorycznych.

Do syntezy bibliotek niekodowanych można wykorzystać jedną z dwóch metod: pierwszą z nich jest **synteza łączona** (ang. *mixture synthesis*), a druga tzw. **synteza dzielona** (ang. *split-synthesis*). W pierwszej zakłada się, że użycie mieszaniny monomerów podczas kolejnych etapów addycji doprowadzi do uzyskania wszystkich możliwych kombinacji tych monomerów. (Należy pamiętać, że w mieszaninie musi powstawać przynajmniej tyle cząsteczek ile kombinacji strukturalnych liczy otrzymywana biblioteka, o ile chcemy uzyskać jej **kompletną** wersję.) Odmienne reaktywności monomerów użytych do syntezy biblioteki doprowadzą do zmniejszenia jej zróżnicowania. Na przykład, przy syntezie biblioteki peptydowej zbudowanej z aminokwasów białkowych, błędne byłoby założenie, że wszystkie aminokwasy mają jednakową zdolność tworzenia wiązania peptydowego. Najreaktywniejsze z nich byłyby preferowane w syntezie, co wyklucza osiągnięcie maksymalnego zróżnicowania sekwencji peptydów (**biblioteka kompletna**) (9).

Podejście to stosuje się praktycznie wyłącznie w syntezie bibliotek oligonukleotydowych, ponieważ zaktywowane nukleotydy mają bardzo podobne reaktywności.

**Synteza dzielona** (ang. *split-synthesis*) polega na zdefiniowaniu takiej liczby osobnych „kanałów” syntezy z ilu monomerów buduje się elementy biblioteki (9,10). W każdym z nich zachodzi przyłączanie tylko jednego, określonego monomeru. Synteza przeprowadzana jest na podłożu stałym, np. kul-

## addycja:



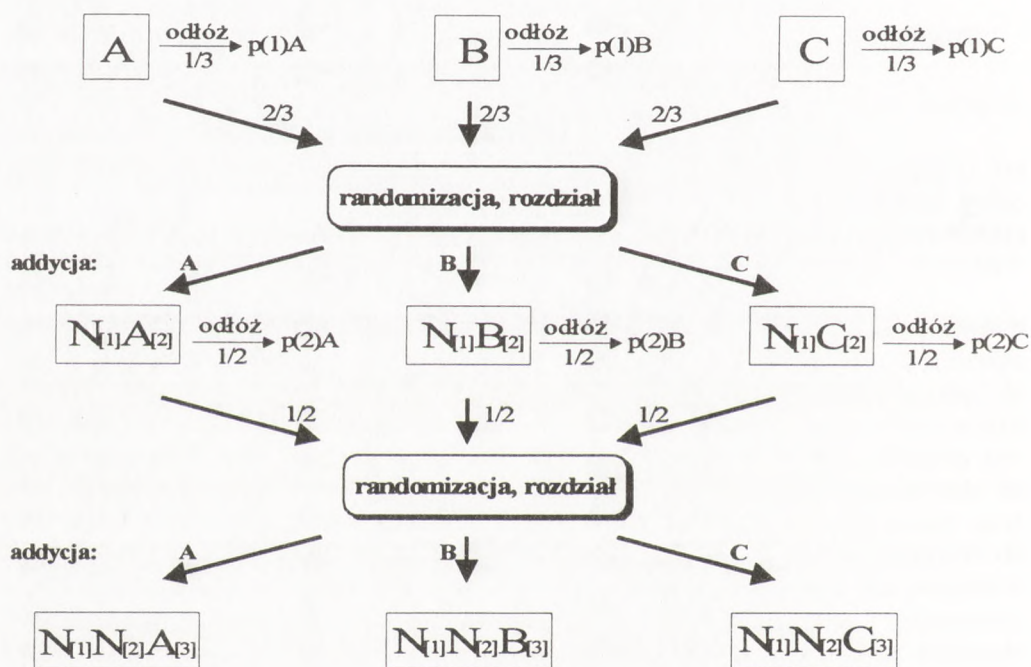
Rys. 1. Przykład zastosowania metody syntezy dzielonej przy otrzymywaniu kombinatorycznej biblioteki tripeptydów z użyciem trzech aminokwasów: alaniny (Ala), glicyny (Gly) i waliny (Val).



kach polistyrenowych. Rozważmy syntezę np. biblioteki trójpeptydów chemicznie syntetyzowanych z alaniny, glicyny i waliny (rys. 1). Mamy do czynienia z trzema monomerami składników biblioteki, wymagane zatem będą trzy kanały syntezy; podłoże dzieli się na trzy równe części do każdej z nich przyłącza się jeden z aminokwasów. W kolejnym kroku podłoże łączy się, dokładnie miesza i ponownie rozdziela. Następnie dokonuje się addycji kolejnego aminokwasu. Powtarzając opisaną procedurę  $n$  razy i mając do dyspozycji  $m$  monomerów otrzymuje się bibliotekę złożoną z maksimum  $m^n$  różnych cząsteczek. W opisanym przypadku, po trzech cyklach syntezy z użyciem trzech aminokwasów otrzymujemy  $3^3 = 27$  peptydów o różnej sekwencji.

Na każdej kulce podłoża wchodzącego w skład otrzymanej w ten sposób biblioteki znajduje się tylko jeden rodzaj peptydu. Bibliotekę można poddać selekcji odpowiednio wyznakowanym (np. fluorescencyjnie) **związkiem — akceptorem**, a peptyd znajdujący się na wyselekcjonowanej w ten sposób kulce podłoża, po oddzieleniu, poddać sekwencjonowaniu (10).

Metoda syntezy dzielonej pozwala na osiągnięcie wysokiego stopnia zróżnicowania elementów biblioteki kombinatorycznej. Osiąga się to dzięki oddzieleniu etapów addycji poszczególnych monomerów podczas syntezy przez co różnice ich reaktywności nie zmniejszają różnorodności biblioteki. Łatwość otrzymania biblioteki kombinatorycznej na drodze syntezy dzielonej przyczy-



Rys. 2. Ogólny schemat syntezy biblioteki kombinatorycznej i oznaczenia subbibliotek przy zastosowaniu procedury analizy zwrotnej.



niła się do powszechnego stosowania kombinatorycznych **bibliotek rozproszonych**.

Houghten i wsp. (11), na drodze syntezy dzielonej, otrzymali zestawy **subbibliotek**, w których w określonej pozycji zdefiniowano dwa aminokwasy (ang. *dual definition*). Tak zatem z 18 aminokwasów białkowych (bez Cys i Trp) otrzymali 324 subbiblioteki ( $18^2$ ) heksapeptydów. Dwa pierwsze aminokwasy były zdefiniowane, po czym w wyniku czterech cykli syntezy dzielonej dołączono cztery dalsze. W każdej subbibliotece było  $18^4$  różnych peptydów o wzorze ogólnym  $X_1X_2(N)_4$ . Cała biblioteka składała się na tym etapie z ok. 34 milionów różnych heksapeptydów. Peptydy z każdej z 324 subbibliotek zostały odblokowane i uwolnione z podłoża. Następnie były testowane pod kątem wyznaczenia minimalnego stężenia  $IC_{50}$  dla reakcji wiązania specyficznego przeciwciała do wybranego jako cząsteczka docelowa 13-peptydu. Jeśli dla jednej z subbibliotek zaobserwowano pozytywne wyniki selekcji, powiedzmy w bibliotece  $A_1A_2(N)_4$ , syntetyzowano następnie 19 nowych subbibliotek (na tym etapie włączono już do syntezy tryptofan) —  $A_1A_2X_3(N)_3$ , po jednej dla każdego aminokwasu zdefiniowanego w pozycji trzeciej i testowano w celu znalezienia  $X_3$ . Proces powtarzano tak długo aż oznaczono ostatnią resztę aminokwasową, a co za tym idzie odnaleziono sekwencję aktywnego peptydu.

Cechą charakterystyczną tego podejścia jest **zastąpienie sekwencjonowania** związków biblioteki wielokrotnymi **syntezami i selekcjami**. Jest to istotna zaleta szczególnie w przypadku bibliotek związków, których metody sekwencjonowania nie są opracowane. Ponieważ zmniejsza się różnorodność badanych sekwencji peptydowych wraz z identyfikacją każdej reszty aminokwasowej następuje zwielokrotnienie ilości aktywnego składnika w bibliotece. Ułatwia to jego wykrycie. Za pomocą tej metody udało się otrzymać specyficzne peptydy wykazujące aktywność skierowaną przeciwko niektórym bakteriom gramujemnym i gramodatnim (11).

Jeszcze inny wariant tego podejścia został opisany przez Janda'ę i wsp. jako tzw. **procedura analizy zwrotnej** (ang. *recursive deconvolution strategy*) (9). Podczas syntezy biblioteki kombinatorycznej odkłada się tzw. **biblioteki cząstkowe**. Załóżmy, że tworzymy bibliotekę zbudowaną z trzech monomerów: A, B i C, a zatem stopnia trzeciego (rys. 2). Także i w tym przypadku wykorzystuje się metodę syntezy dzielonej. Na początku syntezy na trzech oddzielnych ścieżkach syntezy do podłoża stałego przyłącza się odpowiednio A, B lub C. Część tej biblioteki odkłada się i oznakowuje jako bibliotekę cząstkową  $p(1)$ , przy czym wielkość katalogowanej frakcji jest odwrotnością **stopnia biblioteki** w pierwszym cyklu. Stąd, przy założeniu, że otrzymuje się bibliotekę trzeciego stopnia,  $1/3$  każdej puli jest odkładana i oznaczana — w dalszych etapach syntezy wielkość magazynowanej frakcji równa jest odwrotności stopnia biblioteki minus jeden. Pozostały materiał jest łączony, mieszany, rozdzielany i ponownie przyłączane są monomery A, B oraz C. Znowu część tej biblioteki zostaje zachowana jako biblioteka cząstkowa  $p(2)$  — złożona jest ona z puli  $N_1A_2$ ,  $N_1B_2$  i  $N_1C_2$ . Pozostałość ponownie łączy się i miesza, po czym dokonuje się trzecia addycja, otrzymując bibliotekę złożoną z puli



$N_1N_2A_3$ ,  $N_1N_2B_3$  i  $N_1N_2C_3$ .

Selekcji aktywnych składników biblioteki dokonuje się najpierw poprzez inkubację wszystkich subbibliotek z cząsteczką docelową, a następnie wraca się do odłożonych bibliotek cząstkowych. Przykładowo, poddając selekcji trzy pule  $N_1N_2X_3$ , gdzie X oznacza A, B lub C, mamy 9 składników w każdej puli, czyli 27 we wszystkich trzech. Założmy, że składnik  $N_1N_2B_3$  pozytywnie oddziaływał z cząsteczką docelową. Wykorzystujemy teraz bibliotekę cząstkową  $p(2)$ , dzielimy ją na trzy części i do każdej z nich przyłączamy monomer B, otrzymując w końcu trzy pule  $p(2)A$ ,  $p(2)B$  i  $p(2)C$  o wzorze ogólnym  $N_1X_2B_3$ . Te trzy subbiblioteki zawierają już tylko dziewięć różnych składników, zróżnicowanie zostało zatem zmniejszone trzykrotnie. Po ponownej inkubacji z cząsteczką docelową składnik  $N_1A_2B_3$  wykazał pożądane cechy. Wykorzystuje się następnie bibliotekę cząstkową  $p(1)$  przyłączając kolejno monomery A i B, otrzymuje się trzy nowe pule o wzorze ogólnym  $X_1A_2B_3$ . Testuje się je ponownie w celu zdefiniowania  $X_1$ .

Metody selekcji używane przy wyszukiwaniu aktywnych składników bibliotek niekodowanych mają jedną wspólną cechę: badane są bezpośrednio elementy biblioteki.

Bardzo czułą metodą selekcji jest test ELISA, który dobrze nadaje się również do wyznaczania względnego powinowactwa wybranych składników biblioteki do cząsteczki akceptorowej. Jeżeli dysponuje się specyficznym przeciwciałem do cząsteczki docelowej, wyznaczenie powinowactwa polega na określeniu minimalnego  $IC_{50}$  wyselekcjonowanego elementu biblioteki dla reakcji wiązania przeciwciała do cząsteczki akceptorowej (10).

Zastosowanie znajdują również reakcje histochemiczne. Lam i wsp. (10) zsyntetyzowali bibliotekę pentapeptydów na podłożu w postaci kulek polistyrenowych stosując metodę syntezy dzielonej. Cząsteczka docelowa została sprzężona z alkaliczną fosfatazą. Po inkubacji tak wyznakowanej cząsteczki z biblioteką peptydów, po zadaniu mieszaniny inkubacyjnej substratem dla alkalicznej fosfatazy (np.  $\alpha$ -naftylofosforanem sodowym; po hydrolizie uwolniony  $\alpha$ -naftol daje w obecności soli dwuazoniowej *Fast Blue* intensywne niebieskie zabarwienie) możliwe było wyodrębnienie za pomocą mikromanipulatora pojedynczej, na niebiesko zabarwionej kulki podłoża (o średnicy ok. 200  $\mu\text{m}$ ) z aktywnym peptydem spośród pozostałych, bezbarwnych kulek podłoża. Na każdej kulce podłoża peptydy występowały w ilości około 200 pikomoli, czyli wystarczającej do przeprowadzenia sekwencjonowania pentapeptydów. Cząsteczkę akceptorową można wyznakować za pomocą związku o właściwościach fluorescencyjnych — jeżeli w bibliotece znajdzie się składnik wiążący się z cząsteczką związku akceptorowego, element biblioteki (podłoża) zawierający go będzie wykazywał fluorescencję. Konieczny jest tu rozdział mechaniczny, np. za pomocą fluorescencyjnego, automatycznego segregatora komórkowego (FACS).

Druga klasa bibliotek kombinatorycznych, **biblioteki kodowane**, różnią się od **niekodowanych** wyposażeniem każdego składnika biblioteki w unikatową, przyporządkowaną jemu i tylko jemu cząsteczkę reporterową — **etykietę kodującą**. Wyselekcjonowany, aktywny element biblioteki identyfikuje



się poprzez poznanie struktury przypisanej mu cząsteczki reporterowej.

Jedną z najbardziej obiecujących metod kodowania syntetycznych, kombinatorycznych bibliotek są metody wykorzystujące układy biologiczne.

Opisano wykorzystanie fagów włóknikowatych z rodziny *Inoviridae* (M13, fd, f1) w syntezie bibliotek peptydowych (12,13,14). Peptydy biblioteki były zintegrowane z białkami otoczki wirionu. Każdy fag był nosicielem peptydu o unikatowej sekwencji, który był eksponowany na zewnątrz faga, a jednocześnie posiadał w swoim genomie kodujący go odcinek DNA.

Specyfika sekwencji syntetycznego DNA zapewnia różnorodność sekwencji peptydów biblioteki. Ogólnie można ją zapisać wzorem  $(NNS)_x$  (12,13,15), gdzie NNS oznacza tryplet kodujący aminokwas, a  $x$  długość łańcucha peptydowego. N oznacza jeden z czterech nukleotydów (A, C, G lub T), natomiast S — G lub C. Osiągnięto w ten sposób zmniejszenie degeneracji kodu genetycznego; kodowane są wszystkie aminokwasy białkowe, ale tylko jeden kodon stop — amber (ATC → UAG).

Syntetyczny DNA kodujący peptydy biblioteki wprowadzono do genomu faga tuż za genem III kodującym białko płaszczka *pIII*, znajdujące się na szczycie wirionu. Otrzymanym produktem transformowano *E. coli*, a następnie z użyciem faga pomocniczego otrzymano fagi potomne ze związanymi z białkiem *pIII* peptydami biblioteki eksponowanymi w stronę środowiska zewnętrznego otaczającego faga.

Selekcja takiej biblioteki polega na adsorpcji fagów z aktywnymi peptydami na podłożu zawierającym unieruchomioną cząsteczkę akceptorową. Związane z nią fagi eluuje się, infekuje nimi *E. coli*, powtarzając następnie całą procedurę selekcji. Każdy cykl izolacji zwiększa udział fagów z aktywnymi peptydami w stosunku do pozostałych fagów. W celu poznania sekwencji interesującego nas peptydu izoluje się DNA z wyselekcjonowanego klonu faga i sekwencjonuje się rejon zawierający sekwencję  $(NNS)_x$  kodującą peptyd.

We współczesnej biologii i medycynie molekularnej bardzo szerokie zastosowanie znajdują przeciwciała o określonej specyficzności. Czasochłonny proces immunizacji i tradycyjne metody otrzymywania przeciwciał monoklonalnych (16) mogą zostać zastąpione przez kombinatoryczną syntezę bibliotek przeciwciał *in vitro* z użyciem fagów włóknikowatych (15,17). Warunkiem sukcesu na tym polu jest zręczne naśladowanie procesów zachodzących w naturze.

Odpowiedź limfocytów B na obecność antygeny dzieli się na dwa etapy. W odpowiedzi pierwotnej wytwarzane są przeciwciała o niskim powinowactwie do antygeny, głównie klasy *IgM*, pochodzące z ogólnie dostępnej puli limfocytów B istniejącej w momencie immunizacji. Proces ten może być naśladowany *in vitro* poprzez amplifikację w reakcji PCR sekwencji kodujących łańcuchy *H* i *L* na matrycy DNA pochodzącej z biblioteki *cDNA*, otrzymanej ze szpiku kostnego (17). Wektor selekcyjny *pComb8* (18), dzięki któremu przeciwciała biblioteki obecne są w wielu kopiach na powierzchni fagów pozwala na selekcję przeciwciał o niskim powinowactwie do antygeny (17,19). Stąd, specyficzne pary łańcuchów  $V_H$  i  $V_L$  mogą być wyizolowane z pierwotnego, bardzo różnorodnego repertuaru biblioteki.



Drugi etap odpowiedzi immunologicznej, odpowiedź wtórna, polega na dojrzewaniu powinowactwa specyficznych przeciwciał. Biorą w tym udział mutacje somatyczne rejonów zmiennych V w połączeniu z selekcją komórek syntetyzujących przeciwciała o największym powinowactwie do antygeny. Dobrą metodą naśladowania mutacji somatycznych *in vitro* jest wprowadzenie mutacji do genów przeciwciał w reakcji PCR z udziałem zdegenerowanych sekwencji starterowych na matrycy wyselekcjonowanych genów przeciwciał o niskim powinowactwie do antygeny. Można też przeprowadzić ekspresję tych genów w szczepie *mutD E. coli*, w którym częstość mutacji spontanicznych jest  $10^3$  —  $10^5$  wyższa niż w szczepie dzikim (17). Ekspresja fragmentów Fab przeciwciał zachodzi przy udziale wektora selekcyjnego *pComb3* (19); fagi posiadają na swojej powierzchni tylko jedną kopię fragmentu *Fab* przeciwciała (20). Pozwala to na selekcję *Fab* o największym powinowactwie do antygeny (18), co jest procesem analogicznym do selekcji limfocytów B w odpowiedzi wtórnej.

Barbas i wsp. (13) otrzymali **pólsyntetyczną bibliotekę przeciwciał** o różnorodności równej co najmniej tej, jaką otrzymałoby się w opisanym procesie naturalnym. Wykorzystali oni gen przeciwciała *TT7E*, wiążącego **toksoid** anty-tężcowy. W trzech kolejnych reakcjach PCR na matrycy plazmidu *pC3TT7E* (zawierającego sekwencję kodującą fragment *Fd* przeciwciała *TT7E*) otrzymali gen łańcucha ciężkiego przeciwciała randomizowany w rejonie zmiennym *CDR3* (13). Randomizacja polegała na umieszczeniu w obrębie sekwencji kodującej *CDR3* zdegenerowanego oligonukleotydu o długości 48 par zasad —  $(NNS)_{16}$ . Zrekombinowany gen wklonowano do wektora selekcyjnego *pComb3* z wprowadzonym genem łańcucha lekkiego przeciwciała *TT7E*. Po transformacji *E. coli* selekcjonowano klony produkujące fagi z przeciwciałami wiążącymi się z cząsteczką akceptorową. Zmodyfikowany *CDR3* może dać teoretycznie około  $10^{20}$  różnych sekwencji, rząd ewentualnego zróżnicowania zbliżony jest zatem do możliwego w procesie naturalnej immunizacji. Natomiast zakres analizowanej różnorodności zależy od kompletności biblioteki i ilości badanych rekombinantów.

Po wyselekcjonowaniu aktywnych klonów, specyficzne przeciwciała można otrzymać w formie rozpuszczalnej. Dokonuje się tego poprzez wycięcie z *pComb3* sekwencji kodującej białko pIII, z którym wiążą się *Fab* (po ligacji następuje zmiana funkcji *pComb3* z wektora selekcyjnego na wektor ekspresyjny). Syntetyzowane do dalszych badań w komórkach gospodarza fragmenty *Fab* przeciwciał uwalniane są do roztworu (19).

Pomimo że procesy mutagenyzy i selekcji *in vitro* są znacznie prostsze i szybsze od zachodzących podczas naturalnej immunizacji, posiadają one szereg cech wspólnych. Po pierwsze, przeciwciała z biblioteki o różnorodności odpowiadającej repertuarowi przeciwciał odpowiedzi pierwotnej posiadają powinowactwo podobnego rzędu wielkości jak ich naturalne odpowiedniki. Po drugie, pomimo że mechanizmy mutagenyzy *in vivo* i *in vitro* są różne, mutacje mające szczególny wpływ na powinowactwo przeciwciał zachodzą w podobnych „gorących” regionach genów (ang. *mutational hot-spots*). Po trzecie,



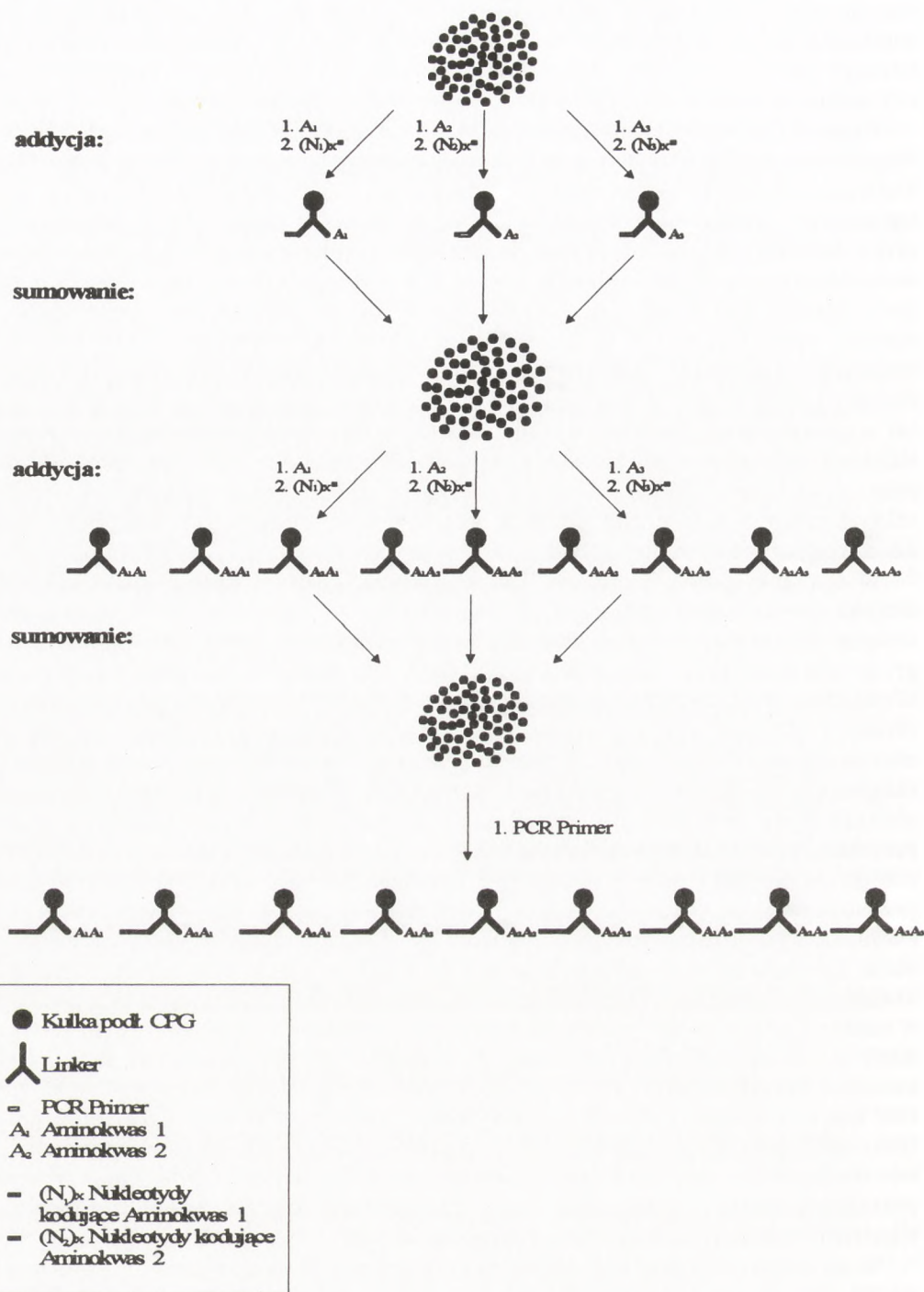
wzrost powinowactwa po mutagenезie i selekcji *in vitro* ma podobny rząd wielkości jak przy przejściu z odpowiedzi pierwotnej w odpowiedź wtórną. Istnieje poza tym szansa otrzymania kombinacji przeciwciał niemożliwych do uzyskania *in vivo*.

Fagowe biblioteki kombinatoryczne syntetyzują się szybko i są one bardzo bogate pod względem różnorodności składników, łatwa jest amplifikacja i sekwencjonowanie DNA kodującego składniki biblioteki. Istnieje jednak pewne ograniczenie osiąganego różnorodności ze względu na możliwość wykorzystania przez układy biologiczne tylko białkowych aminokwasów, a cząsteczkami wchodzącymi w skład biblioteki mogą być tylko peptydy i białka. Połączenie chemicznych metod syntezy bibliotek kombinatorycznych z genetycznymi sposobami identyfikacji ich składników umożliwiłoby wykorzystanie do budowy bibliotek cząsteczek o bardzo zróżnicowanej naturze chemicznej, osiągnięcie wysokiej różnorodności ich składników oraz łatwą ich analizę. Brenner i wsp. (4) zaproponowali technikę syntezy kombinatorycznej, w której przyłączenie każdego monomeru sprzężone jest z przyłączeniem ściśle określonej, tylko jemu przypisanej sekwencji oligonukleotydy — etykiety kodującej. Biblioteki otrzymane tą techniką nazwał **genetycznie kodowanymi bibliotekami kombinatorycznymi** (rys. 3).

W opisanej przez Brennera i wsp. syntezie **genetycznie kodowanej biblioteki peptydowej** wykorzystuje się metodę syntezy dzielonej na podłożu stałym. Wydajną syntezę równoległą zapewnia zastosowanie odpowiednich grup ochronnych: dimetoksytrytylu (*DMT*) jako grupy ochronnej dla oligonukleotydów oraz 9-fluorenylmetoksykarbonylu (*Fmoc*) dla peptydów. Grupy *Fmoc* i *DMT* zapewniają kompatybilną syntezę obu rodzajów cząsteczek dzięki czemu oligonukleotydy nie ulegają degradacji w specyficznych warunkach wymaganych do syntezy peptydów i odwrotnie. Peptydy i ich reporterowe oligonukleotydy połączone są ze sobą za pomocą specjalnego, trójfunkcyjnego łącznika, którego budowa umożliwia syntezę równoległą i zarazem przyłączenie go do podłoża stałego. Łącznik powinien posiadać odpowiednią budowę przestrzenną umożliwiającą łatwą deprotekcję syntetyzowanych związków i wysoką wydajność syntezy. Nielsen i wsp. (21) opisali cztery odrębne podłoża stałe wyposażone w różne łączniki do zastosowania w syntezie bibliotek kombinatorycznych. Najwydajniejszą syntezę uzyskano w przypadku podłoża, w którym grupy funkcyjne przyłączające aminokwasy i nukleotydy są najbardziej od siebie oddalone. Natomiast, podczas selekcji biblioteki za pomocą związku akceptorowego zauważono, że powinowactwo jej aktywnych elementów zależy od tego, czy do łącznika dołączona jest jedna czy dwie kopie peptydu biblioteki. Peptyd asocjujący z przeciwciałem 3-E7, wiążącym  $\beta$ -endorfinę, na łączniku wiążącym pojedynczą cząsteczkę peptydu, miał  $K_D$  wiązania ponad trzykrotnie większą niż wtedy gdy był związany z łącznikiem niosącym jego dwie cząsteczki ( $K_D$  równe odpowiednio 24 nM i 7nM).

W omawianym podejściu bibliotekę przeprowadza się w formę rozpuszczalną po zakończonej syntezie (biblioteka składa się zatem z pojedynczych cząsteczek). Aktywne elementy biblioteki są selekcjonowane, a następnie kodu-





Rys. 3. Ogólny schemat syntezy genetycznie kodowanej biblioteki kombinatorycznej.



jące je odcinki DNA amplifikuje się w reakcji PCR (4,9). Jest to możliwe dzięki specjalnemu zaprojektowaniu etykiety kodującej, w której sekwencja kodująca element biblioteki umieszczona jest między regionami flankującymi — jeden z nich jest komplementarny, drugi zaś identyczny z sekwencjami starterowymi reakcji PCR (starter F i starter R). W ostatnim etapie selekcji następuje identyfikacja aktywnych elementów biblioteki poprzez klonowanie i sekwencjonowanie produktów reakcji PCR. Needles i wsp. (7) zastosowali analogiczną budowę etykiety kodującej z tym, że wyposażyli ją w dodatkowy odcinek DNA komplementarny do sekwencji starterowej dla reakcji sekwencjonowania stosując metodę dideoksy.

## 5. Zakończenie

Biblioteki kombinatoryczne należą do nowych technologii i bardzo obiecujących. Zaproponowano wiele metod ich syntezy i selekcji, tak że ich wybór zależy przede wszystkim od potrzeb i wyposażenia laboratorium badacza. Biblioteki kodowane, lub wytwarzane z użyciem procedury analizy zwrotnej, przypuszczalnie są w stanie zawierać największą różnorodność składników. Biblioteki **kodowane genetycznie** (z wykorzystaniem etykiet oligonukleotydowych) natomiast są wygodne w poszukiwaniach ligandów cząsteczek akceptorowych rzadkich lub trudnych do wyizolowania w większych ilościach. Ponieważ obecnie większość kombinatorycznych bibliotek jest pochodzenia peptydowego lub oligonukleotydowego, biblioteki kodowane dają możliwość użycia do ich syntezy monomerów innych niż aminokwasy i nukleotydy. Automatycznie rozszerzyłyby to obszar różnorodności badanych związków chemicznych. Wymaga to jednak opanowania metod syntezy w warunkach kompatybilnych z syntezą oligodeoksynukleotydów (ortogonalne systemy grup ochronnych)(22).

W wyniku przeprowadzania dalszych badań okaże się czy biblioteki kombinatoryczne mogą rzeczywiście być użytecznym narzędziem w poszukiwaniach nowych leków. Jednak, naszym zdaniem, syntetyczne biblioteki kombinatoryczne z pewnością stanowią nowe, praktycznie niewyczerpane źródło różnorodności chemicznej i w najbliższym czasie staną się jednym z ważniejszych nowych podejść metodycznych zarówno w badaniach podstawowych jak i stosowanych.

Autorzy dziękują Komitetowi Badań Naukowych za finansowanie w ramach projektu grantowego nr 2 P303 008 04.

## Literatura

1. Lam K. T., Hruby V. J., Lebl M., Knapp R. J., Kaźmierski W. M., Hersh E. M., Salmon S. E., (1993), *Bioorganic & Medicinal Chemistry Letters*, 3 (3), 419.
2. Fodor S. P. A., Read J. L., Pirrung M. C., Stryer L., Lu A. T., Solas D., (1991), *Science*, 251, 767.



3. Maskos U., Southern E. M., (1993), *Nucl. Acid Research*, 21, 4663.
4. Brenner S., Lerner R. A., (1992), *Proc. Natl. Acad. Sci. USA*, 89, 5381.
5. Bock L. C., Griffin L. C., Latham J. A., Vermaas E. H., Toole J. J., (1992), *Nature*, 355, 564.
6. Markiewicz W. T., Adrych-Rożek K., Markiewicz M., Żebrowska A., Astriab A., (1994), *Innovations in Solid Phase Synthesis, 1994: Biological and Biomedical Applications*, Ed. R. Epton, Mayflower Worldwide, 339.
7. Needles M. C., Jones D. G., Tate E. H., Heinkel G. L., Kochersperger L. M., Dower W. J., Barret R. W., Gallop M. A., (1993), *Proc. Natl. Acad. Sci. USA*, 90, 10700.
8. Pavia M. R., Sawyer T. K., Moss W. H., (1993), *Bioorganic & Medical Chemistry Letters*, 3, 387; (1994), artykuły w: *Innovations in Solid Phase Synthesis, 1994: Biological and Biomedical Applications*, Ed. R. Epton, Mayflower Worldwide; von Jung G., Beck-Sickingher A. G., (1992), *Angew. Chem.*, 104, 375.
9. Janda K. D., (1994), *Proc. Natl. Acad. Sci. USA*, 90, 10779.
10. Lam K. S., Salmon S. E., Hersh E. M., Hruby V. J., Kazmierski W. M., Knapp R. J., (1991), *Nature*, 354, 82.
11. Houghten R. A., Pinilla C., Blondelle S. E., Appel J. R., Dooley C. T., Cuervo J. H., (1991), *Nature*, 354, 84.
12. Smith G. P., (1985), *Science*, 228, 1315.
13. Devlin J. J., Panganiban L. C., Devlin P. E., (1990), *Science*, 243, 404.
14. Scott J. K., Smith G. P., (1990), *Science*, 249, 386.
15. Barbas III C. F., Bain J. D., Hoekstra D. M., Lerner R. A., (1992), *Proc. Natl. Acad. Sci. USA*, 89, 4457.
16. Jakóbiśiak M. (red.), (1993), *Immunologia*, PWN, Warszawa.
17. Gram H., Marconi L.-A., Barbas III C. F., Collet T. A., Lerner R. A., Kang A. S., (1992), *Proc. Natl. Acad. Sci. USA*, 89, 3576.
18. Kang A. S., Barbas III C. F., Janda K. D., Bencovic S. J., Lerner R. A., (1991), *Proc. Natl. Acad. Sci. USA*, 88, 4363.
19. Barbas III C. F., Kang A. S., Lerner R. A., Bencovic S. J., (1991), *Proc. Natl. Acad. Sci. USA*, 88, 7978.
20. Winter G., Griffiths A., Hawkins R., Hoogenboom H. R., (1994), *Annu. Immunol.*, 12, 433.
21. Nielsen J., Brenner S., Janda K. D., (1993), *J. Am. Chem. Soc.*, 115 (21), 9812.
22. Markiewicz W. T., (1995), *Wiadomości Chemicz.*, 11/12.

## Chemical Combinatorial Libraries — Their Synthesis and Applications

### Summary

The paper reviews the new methodologies of synthesis and screening of combinatorial libraries of chemical compounds (CCL). The original classification of chemical combinatorial libraries is presented. The combinatorial libraries of chemical compounds are the collections of low (peptides, their analogues, oligonucleotides etc.) or high molecular compounds (fragment antibodies, chemically modified phages etc.). They can be obtained via chemical or chemical/biological routes and screened with labelled acceptor (substrate) molecules in order to reveal those of interacting library entities which best meet the selection criteria (conditions of selection). The potential applications of CCLs in basic research and technology are indicated.

### key words:

combinatorial library, oligonucleotide library, peptide library, antibody library, chemoselection, phagemid, Fab.



Notatka dodana po korekcie

Ostatnio opisano jeszcze jeden wariant bibliotek rozproszonych, w którego syntezie wykorzystuje się zarówno syntezę łączoną jak i dzieloną — **biblioteki bibliotek**. Są to takie biblioteki rozproszone, w których kulki podłoża same są bibliotekami. Ich selekcja pozwala wykryć motywy strukturalne (ang. „*pharmacophore*” *motis*) warunkujące występowanie określonych właściwości związków. (Sepetov N. F., Krchnak V., Stankova M., Wade S., Lam K. S., Lebl M., (1995), Proc. Natl. Acad. Sci. USA, 92, 5426).

*Adres do korespondencji:*

Wojciech T. Markiewicz, Instytut Chemii Bioorganicznej PAN,  
ul. Noskowskiego 12, 61-704 Poznań; fax: (061) 52 05 32,  
e-mail: markwt@ibch.poznan.pl