

Melléklet

TANULMÁNY, GYÓGYÁSZATIILAG HASZNOSITHATÓ PEPTIDEK SZISZTEMATIKUS FELKUTATÁSÁNAK LEHETŐSÉGEIRŐL

Készítette Dr. Furka Árpád egyetemi
tanár, Budapest, 1982 május 29.

Többek között az eddig felfedezett peptidhormonok példája is tanusítja, hogy a hosszabb-rövidebb peptidek az élő szervezetben számos fontos funkciót láthatnak el. Feltehető, hogy ezeknek a biológiailag aktív, és potenciális terápiás hatással bíró peptideknek eddig csak egy kis töredékét ismerjük. Ez indokolja, hogy ezen a területen világszerte és hazánkban is, intenzív kutatómunka folyik.

Az újfajta biológiai hatással rendelkező peptidek felkutatására kétféle elvi lehetőség kínálkozik:

1. A peptidek izolálása az élő szervezetből, előzetesen felismert biológiai hatásuk alapján.

2. A peptidek szintetikus előállítása és biológiai hatásaik utólagos felderítése.

Eddig az izolálásos módszer bizonyult járhatóbbnak annak ellenére, hogy ez is igen munkaigényes. Ennek az a magyarázata, hogy az adott tagszámú peptidek lehetséges száma olyan gyorsan nő a tagszámmal, hogy már a tetrapeptidek teljes számban /160 ezer/ történő előállítása is a gyakorlatban megoldhatatlan feladatnak tűnik. Ha a 20 féle fehérjealkotó aminosavat vesszük alapul, a lehetséges peptidek számát N_n a következő egyenlet fejezi ki a tagszám n függvényében:

$$N_n = 20^n$$

Ha a peptideket lépésenként és egymástól függetlenül állítjuk elő, az n tagú peptidek esetében az ehhez szük-

- 2 -

séges szintetikus lépések száma S_n a következőképpen adódik:

$$S_n = n - 1 / 20^n$$

Megjegyzendő, hogy szintetikus lépésen most és az elkövetkezőkben egy teljes kapcsolási ciklust értünk, tehát a kapcsolási lépésen kívül beleértjük a védőcsoportokkal kapcsolatos műveleteket is.

Jó szervezéssel, azaz szisztematikus módon eljárva a szintézisnél, a szintetikus lépések száma csökkenthető. A minimálisan szükséges lépésszám:

$$S_n = \sum_{i=2}^n 20^i$$

Az előállított peptideket hatásvizsgálatnak kell alávetni, ami szintén reménytelenül nagy szám, hiszen minden peptidet többféle hatásra kell tesztelni. Ha a tesztek fajtáinak számát t -vel jelöljük, a szükséges hatásvizsgálatok teljes számát T_n a következő egyenlet fejezi ki:

$$T_n = t20^n$$

Az 1. táblázatban találjuk felsorolva a különböző tagszámú peptidok lehetséges számát, a szintézisükhöz szükséges lépésszámot, valamint a tesztek számát, 10 különböző hatásvizsgálattal számolva $t=10$. A feltüntetett értékek kerekítve vannak. A felsorolt számokból világosan kitűnik, hogy szinte reménytelen dolog már a tripeptidek teljes számban történő előállítására és tesztelésére vállalkozni.

- 3 -

1. Táblázat

Különböző tagszámú n peptidek lehetséges száma N_n , lépésenkénti előállításukhoz szükséges szintetikus lépések száma S_n , szisztematikus szintézistervezés mellett, továbbá a hatásvizsgálatok száma T_n , 10 különböző hatásvizsgálattal számolva $t=10$

/A feltüntetett értékek kerekítve vannak/

n	N_n	S_n	T_n
2	4 száz	4 száz	4 ezer
3	8 ezer	8 ezer	80 ezer
4	160 ezer	168 ezer	2 millió
5	3 millió	3 millió	30 millió
6	64 millió	67 millió	640 millió
7	1 milliárd	1 milliárd	13 milliárd
8	25 milliárd	26 milliárd	256 milliárd
9	512 milliárd	537 milliárd	5 billió
10	10 billió	10 billió	102 billió

A lehetséges peptidek nagy száma miatt lépésenkénti szintézissel már a kis tagszámú peptidek teljes számban történő előállítására sem lehet vállalkozni. További problémát jelent a szükséges ~~teszt~~vizsgálatok nagy száma a biológiailag hatásos molekulák kiválasztása céljából. Ezen a szinte reménytelen helyzetben próbál valamelyest javítani a következőkben előterjesztendő javaslat.

- 4 -

Kistagszámu, biológiaiilag aktiv peptidek szisztematikus felkutatása, peptidkeverékek szintézise és aktivitásvizsgálata utján

Az alább ismertetendő javaslat tulajdonképpen egy kutatási terv, amelynek révén az eddigieknél sokkal nagyobb eséllyel lehet vállalkozni biológiaiilag aktiv peptidek felkutatására. Amikor ezt a tervet leírom, teljesen tudatában vagyok annak, hogy e terv végrehajtásának milyen nagy ipari jelentősége lehet. Az is világos, hogy megvalósítása csak több intézmény együttműködése esetén lehetséges. Elsősorban a gyógyszeripar közreműködése kívánatos, hiszen a befektetések a gyógyszeripar révén térülhetnek meg.

A kutatási javaslat lényege az, hogy a peptideket nem egyenként szintetizáljuk, hanem olyan peptidkeverékeket állítunk elő amelyek több száz, vagy több ezer peptidet tartalmaznak nagyjából azonos mólarányban és ezeket a keverékeket vetjük alá hatásvizsgálatnak. Amint látni fogjuk ezen az uton a szintéziseknél és a hatásvizsgálatoknál egyaránt nagyon sok munkát lehet megtakarítani. Az első közelítés során csak annak megállapítása a cél, hogy a kérdéses peptidkeverék mutat-e valamilyen biológiai hatást. Ha észlelünk hatást, természetesen el kell tudni dönteni azt, hogy az aktivitás a peptidkeverék mely komponensétől /vagy mely komponenseitől/ származik.

A peptidkeverék szintézisének módszere.

Mint hogy nem egyes peptidek hanem peptidkeverékek szintéziséről van szó, ezek utólagos tisztítása a melléktermékektől, nem jöhet szóba. Ezért a klasszikus módszer sem jöhet szóba. A peptidkeverékek szintézisének a szilárdfázisú módszerrel kell alkalmazni. Már itt megemlítjük, hogy a szintézisnél nem szükségképpen 20 fajta aminosavat használunk. Esetenként használhatunk többet, ha például nem szokványos aminosavrészeket is be akarunk építeni, de kevesebbet is, mert például a dekapptidek esetében már nem törekedhetünk a

- 5 -

teljességre, hanem meg kell elégednünk bizonyos típusu aminosavak szerepeltetésével. Az i -edik pozícióban variálható aminosavrészek számát k_1 -vel jelöljük. A C-terminálison variálható aminosavrészek száma: k_1 ; az N-terminálison pedig: k_n .

A szintézis kivitelezése céljából a felhasználandó gyantát k_1 számú egyenlő mennyiséget tartalmazó adagra osztjuk /annyi adagra ahány fajta aminosavrészt akarunk szerepeltetni a peptidek C-terminálisán/. Ezután mindegyik adag gyantára külön-külön felkapcsoljuk a k_1 fajta aminosav valamelyikét, majd eltávolítjuk mindegyik mintáról az amino-védőcsoportot. Mindegyik mintából félreteszünk egy-egy kis adagot később történő felhasználás céljára, majd a mintákat alaposan összekeverjük. Ezután az aminoacil-gyantakeveréket k_2 egyenlő adagra osztjuk és mindegyikre rákapcsoljuk a k_2 féle védett aminosav valamelyikét, majd külön-külön ismét eltávolítjuk az amino-védőcsoportokat. Összekeverés előtt most is félreteszünk egy-egy kis adagot későbbi felhasználás céljára. Az összekevert gyanta kis részletéről lehesítjük a dipeptidek keverékét biológiai tesztvizsgálatok számára, majd a többit k_3 egyenlő részre osztjuk és felkapcsoljuk a harmadik pozícióban szereplő aminosavakat. Így haladunk tovább a szintézissel míg az n tagot tartalmazó peptidek keverékéhez jutunk el.

Az elmondottakhoz néhány megjegyzést kell fűzni. Mint a közönséges szilárdfázisu szintézisnél itt is törekedni kell arra, hogy a reagens feleslegének alkalmazásával jó konverziót érjünk el. Szerencsére azonban, a nem 100%-os konverzió, vagy kismértékű, nem kívánt hasadásos reakciók nem okoznak olyan komoly problémát mint a közönséges szintéziseknél. A szintézis munkaigényét jelentősen csökkentené, ha az acilezési reakciókat a megfelelően védett és aktivált aminosavak keverékével lehetne végrehajtani. Ez azonban nem látszik járható utnak az aktivált aminosavak eltérő reaktivitása miatt, ami odavezetne, hogy a peptid-

- 6 -

keverék komponensei nagyon eltérő koncentrációban keletkeznének ami nem engedhető meg mert zavart okozna a hatásvizsgálatoknál. A peptidok azonos koncentrációban történő képződését a minták mechanikus összekeverésével és egyenlő adagokra történő szétmérésével lehet biztosítani. Így módja van minden egyes aminosavösszetevőnek teljes mértékben elreagálni. Mérlegelendő a továbbiakban, néhány, azonos reaktivitású aminosavszármazék keverékével történő acilezés lehetősége is. Kisebb reaktivitásbeli eltéréseket kompenzálni lehetne a keveréket alkotó aminosavszármazékok molarányának alkalmas megválasztásával. A következő számításoknál azonban, a több aminosavszármazék keverékével történő acilezés lehetőségét figyelmen kívül hagyjuk.

A szintézissel előállított peptidok számát, vagyis a peptidkeverék komponenseinek számát, általános esetben a következő képlet szerint számíthatjuk ki:

$$N_n = k_1 \cdot k_2 \cdot \dots \cdot k_{n-1} \cdot k_n$$

Ha mindegyik pozícióban ugyanannyi aminosavrészt /k/ variálunk,

$$N_n = k^n$$

Az n tagszámú, és N_n komponenst tartalmazó peptidkeverék előállításához szükséges szintetikus lépések száma /különlépésnek tekintve az első aminosavrész gyantára történő felkapcsolását is/:

$$S_n = k_1 + k_2 + \dots + k_{n-1} + k_n$$

Ha mindegyik pozícióban k fajta aminosavrészt variálunk,

$$S_n = nk$$

- 7 -

A fenti képletekből kitűnik a peptidkeverékek szintézisének előnye: a szintetikus lépések számát a variálható aminosavrészek számának összeadásával kapjuk, az előállított peptidok számát pedig a variálható aminosavrészek összeszorozásával.

Egy példa: A 20 féle aminosavréssz variálásával előállítható 160000 tetrapeptid keverékének szintéziséhez mindössze 80 szintetikus lépés szükséges!

Megjegyzendő, hogy menetközben előállítottuk az összes négyenél kisebb tagszámú peptidet is, tehát a 400 dipeptidet és a 8000 tripeptidet. Ennyi peptid hagyományos módon való előállításához összesen 168400 szintetikus lépésre lenne szükség. Más összehasonlítás: 80 szintetikus lépéssel, hagyományos módszerrel, mintegy 30 tetrapeptid szintetizálható.

A peptidkeverék hatásvizsgálata.

A peptidkeverékek szintézise első közelítésben annak megállapítása céljából történik, hogy van-e közöttük biológiaiilag aktív. Abból a feltételezésből indulhatunk ki, amit azonban még kísérletileg alá kell támasztani, hogy a hatásvizsgálat történhet keverék formájában is. Ez nagy nyereséggel jár a hagyományos módszerhez viszonyítva, hiszen annyiszor kevesebb tesztet kell végrehajtani, ahány peptidet tartalmaz a keverék. Így például a 8000 tripeptidet tartalmazó keverék elvileg egyetlen teszt-sorozattal vizsgálható meg. Ha van közöttük aktív peptid, akkor a t számú teszt valamelyike pozitív eredményt ad. Több aktív peptid jelenléte esetén természetesen több fajta teszt is adhat pozitív eredményt. Az n tagszámú peptidok keverékének szintézisének célszerű a kisebb tagszámú peptidok keverékét is tesztelni. A szintézis már úgy van tervezve, hogy erre lehetőség nyíljon. Ezt a követelményt is figyelembe véve, t számú különböző tesztvizsgálat esetén az összes tesztvizsgálatok száma:

- 8 -

$$T_n = t/n - 1/$$

A fenti összefüggés ugyan elvileg kétségtelenül helytálló, azonban a gyakorlati megvalósíthatóságához néhány megjegyzést kell tenni. Minden kétséget kizáróan létezik egy határ, amelyen túl a hatásvizsgálatnak alávetendő peptidkeverék komponenseinek száma nem növelhető. Hogy ez mennyi, azt kísérletek nélkül nehéz megítélni. A keverék valószínűleg több ezer komponenst tartalmazhat ugyan, mégis úgy ítéltető meg jelenleg, hogy a fentebb kifejtett módszer alkalmazhatóságának inkább a tesztelhetőség szab határt, semmint a szükséges szintetikus lépések száma. Ha a komponensek száma túl nagy a keverékben, túl nagy adagban kell a keveréket alkalmazni, hogy az egyes komponensek hatása észrevehető legyen. Megjegyzendő, hogy a keverék feltehetően számos olyan analógot tartalmaz amely többé-kevésbé hatásos, és ezek hatása feltehetően összegeződik. Mindemellett létezik olyan határ, amelyet a komponensek számát illetően nem lehet túllépni. Ezért bizonyos esetekben célravezető lehet, hogy az n tagszámú /tehát az előállítani kívánt legnagyobb tagszámú/ peptidok hatásosságát "összekeverés" nélkül vizsgáljuk. Más esetekben a szintézist úgy célszerű tervezni, hogy a komponensek száma ne haladja meg az optimális határt.

Az aktív peptid "visszakeresése".

Ha kimutattuk, hogy a peptidkeverék tartalmaz aktív komponenst, vagyis a keverék mutat ujszerű biológiai hatást, a következő feladat az aktív peptid izolálása és szerkezetének megállapítása, amit a szintézis követ. Ha már egyszer a kezünkben van az aktív komponenst, vagy komponenseket tartalmazó keverék, az izolálás történhet a már jól bevált elválasztási módszerekkel is, hiszen ezek lehetővé teszik az aktív vegyület kiválasztását akár több ezer inaktív összetevő közül is. Lehetséges azonban másféle módszert is követni. Ennek ismertetésére kerül most

- 9 -

sor. Ez a közelítés az aktiv peptid megismerésére feltehetően kevesebb fáradsággal jár mint az izolációs módszer, amellelt több információval is szolgál, ugyanis egyuttal a szerkezet-hatás összefüggést is feltérképezi. Alkalmazhatóságának feltétele az, hogy legyen módszerünk az aktivitás kvantitatív meghatározására. Egyszerűség kedvéért tételezzük most fel hogy a keverékünk egyetlen hatásos komponenszt tartalmaz /hozzávéve ehhez az ugyanolyan hatást mutató kisebb aktivitású analógokat/.

Az 1. visszakeresési lépés.

Elővesszük azt a k_n db. mintát amit félretettünk az n -tagu peptidek szintézisének, mielőtt azokat összekevertük volna. Mindegyikről lehasítjuk az n -tagu peptidek keverékét. Ezek a peptidkeverékek abban fognak eltérni egymástól, hogy más az n -edik /azaz az N-terminális/ aminosavrész. Mindegyik mintát kvantitatív hatásvizsgálatnak vetjük alá. Ez megmutatja, hogy hogyan függ az aktivitás az N-terminális aminosavrészről, vagyis megállapíthatjuk az aktiv peptid N-terminális aminosavrészét, és egyuttal azt is hogy ennek helyettesítése más aminosavrészszel, hogyan befolyásolja az aktivitást. Tételezzük fel például, hogy a legnagyobb aktivitást mutató minta /és így aktiv peptidünk/ aminoterminálisa: Phe /fenilalanin/. Mindjárt megjegyezzük, hogy ha több egyformán nagy aktivitással járó terminális aminosavat találunk /és ez vonatkozik a további visszakeresési lépésekre is/ célszerű, ha az aktiv peptid N-terminálisaként a legolcsóbb, vagy a szintézisének a legkisebb problémát okozó aminosavat választjuk.

A 2. visszakeresési lépés.

Ezután elővesszük az $n-1$ tagu peptidek szintézisének félretett k_{n-1} db. mintát és mindegyikhez külön-külön hozzákapcsoljuk az előzőleg megállapított aminosavat, példánkban a Phe-t. Ha mindegyik mintáról lehasítjuk a peptideket ezáltal k_{n-1} számú különböző peptidkeverékhez jutunk. Ezek-

- 10 -

ben az a közös, hogy mindegyik peptidnek Phe az N-terminális. Ha a peptidkeverékeket kvantitatív hatásvizsgálatnak vetjük alá megállapíthatjuk az aktív peptid $n-1$ -edik /vagyis az N-terminális előtti/ aminosavrészét. Egyúttal megtudhatjuk, hogy ennek felcserélése más aminosavrészekkel milyen hatást gyakorol az aktivitásra. Tétélezzük fel, hogy aktív peptidünk utolsó előtti aminosavrésze Arg /arginin/. Megjegyzendő, hogy ennél a visszakeresési lépésnél a Phe-t k_{n-1} számú mintához kellett hozzákapcsolni, és ugyanennyi /tehát k_{n-1} / esetben kellett hatásvizsgálatot végezni. Ilyenkor már nem végezzük el mind a t féle tesztet, hanem csak azt, amelyik a kérdéses hatásra vonatkozik. Tehát a szintetikus lépések száma és a tesztvizsgálatok száma egyaránt: k_{n-1} . Azt is megjegyezzük, hogy az ezt megelőző visszakeresési lépésnél csak hatásvizsgálat van / k_n számú/ szintetikus lépésekre nincs szükség.

A 3. visszakeresési lépés.

Ennél és az ezt követő visszakeresési lépéseknél kétféle utat követhetünk. A szintézis közben félretett mintákban a peptidok tagszámát ki kell egészítenünk n -re, úgy, hogy azok N-terminális szakaszukon az aktivitást biztosító aminosavrészeket tartalmazzák. Ezt kétféle módon valósíthatjuk meg. Vagy az aminosavak /példánkban a védett Arg majd Phe/ egyenkénti hozzákapcsolásával, vagy az előzetesen megszintetizált, megfelelő szekvenciájú oligopeptid /példánkban a védett Phe.Arg/ egy lépésben történő hozzákapcsolásával. A szükséges szintetikus lépések száma a kétféle megoldásnál erősen eltér. Az elvégzendő hatásvizsgálatok száma viszont mindkét esetben azonos. Térjünk rá most a 3. visszakeresési lépésre.

A lépésenkénti toldás. Vegyük elő az $n-2$ tagszámú peptidok szintézisének félretett k_{n-2} számú mintát. Mindegyikhez hozzákapcsoljuk külön-külön előbb a védett Arg-ot, majd a védett Phe-t. A peptidok lehasítása után mind a k_{n-2} db. peptidkeveréket aktivitásvizsgálatnak vetjük alá és megállapítjuk

- 11 -

az aktív peptidnek az N-terminálistól számított ^(harmadik) aminosav-részt. Az elvégzendő hatásvizsgálatok száma: k_{n-2} . A szükséges szintetikus lépések száma: $2k_{n-2}$. A k_1 előtti szorzószám annál nagyobb, minél rövidebb peptideket toldunk. A szorzószám értéke megegyezik a hozzátoldandó aminosavrészek számával.

Toldás oligopeptiddel. Előre megszintetizáljuk a védett dipeptidet /példánkban Phe.Arg/ és ezt kapcsoljuk a félretett k_{n-2} számú minta mindegyikéhez. Ezután úgy járunk el mint fent. A hatásvizsgálatok száma itt is: k_{n-2} . A szintetikus lépések száma /az oligopeptid szintézisét nem számítva/ k_{n-2} . Ez a megoldás gazdaságosabbnak tűnik. Gyakorlatilag azt jelenti ez, hogy a visszakereséssel párhuzamosan szintetizáljuk az aktív peptidet is, mégpedig klasszikus módszerrel az aminoterminális felől indulva. A növekvő peptid egy-egy kis részletét feláldozzuk az egyes visszakeresési lépéseknél. Nagy előnye ennek a visszakeresési módszernek, azon túl, hogy kevesebb szintetikus lépésből áll, hogy amire a visszakeresési folyamat végére érünk, készen áll a szintetikus aktív peptid is.

Több aktív peptid visszakeresése. A szintetikus peptidkeverékekben több, eltérő hatású aktív peptid is előfordulhat. Ilyenkor a visszakeresési lépések száma annyiszor lesz nagyobb, ahány aktív peptidről van szó. Vagyis, ha "a" számú peptidről van szó, a szintetikus lépések és a hatásvizsgálatok számát úgy kapjuk ha a fentebb levezetett értékeket szorozzuk a-val. Megjegyzendő, hogy több eltérő hatású peptid egyidejű jelenléte a keverékben komplikálhatja a visszakeresést, különösen ha ellentétes hatású peptid is előfordulnak. Ennek részletezésétől azonban eltekintünk.

A visszakeresési folyamat akkor zárul, amikor vagy az oligopeptides, vagy a lépésenkénti toldás alkalmazásával valamennyi aktív peptid szekvenciáját megállapítottuk.

- 12 -

A szintetikus lépések számának és a hatásvizsgálatok számának összegezése a teljes szintetikus és visszakeresési folyamatra

A szintetikus lépések száma oligopeptides toldásnál.

Szintézisnél:
$$S_n = \sum_{i=1}^n k_i$$

Visszakeresésnél:
$$S_n = a \sum_{i=1}^{n-1} k_i$$

Szintézisnél és visszakeresésnél összesen:

$$S_n = (a + 1) \sum_{i=1}^{n-1} k_i + k_n$$

Ha mindegyik lépésben k aminosavat variálunk:

$$S_n = [n(a + 1) - a]k$$

A szintetikus lépések száma lépésenkénti toldásnál.

Szintézisnél:
$$S_n = \sum_{i=1}^n k_i$$

Visszakeresésnél:
$$S_n = a \sum_{i=1}^{n-1} (n - i)k_i$$

Szintézisnél és visszakeresésnél összesen:

$$S_n = \sum_{i=1}^n k_i + a \sum_{i=1}^{n-1} (n - i)k_i$$

Ha mindegyik lépésben k aminosavat variálunk:

$$S_n = nk + ak \sum_{i=1}^{n-1} i$$

- 13 -

A hatásvizsgálatok száma, mely egyaránt érvényes oligopeptides és lépésenkénti toldásnál.

Szintézisnél: $T_n = t(n - 1)$

Visszakeresésnél: $T_n = a \sum_{i=1}^n k_i$

Szintézisnél és visszakeresésnél összesen:

$$T_n = t(n - 1) + a \sum_{i=1}^n k_i$$

Ha mindegyik lépésnél k aminosavat variálunk:

$$T_n = t(n - 1) + ank$$

Egy példa: az összes pentapeptid előállítás és hatásvizsgálata.

$$N_5 = 3200000, \quad n=5 \quad k=20 \quad t=10 \\ a=1$$

Szintetikus lépések száma összesen

Oligopeptides toldással	180
Lépésenkénti toldással	300
<u>Tesztek száma</u>	140

- 14 -

A módszer kiterjesztése más vegyülettípusokra

Az előzőekben kifejtett módszer nemcsak aktiv peptidok szisztematikus felkutatására alkalmas. Ugyanaz az elv minden más szekvenciális felépítésű vegyülettípusra érvényes, vagyis amikor a vegyülettípushoz tartozó vegyületek az egymáshoz sorban kapcsolódó építőelemek minőségében és sorrendjében különböznek egymástól. Ezek között lehetnek természetes vegyületek mint például az oligoszacharidok, vagy oligonukleotidok, de elképzelhetőek mesterséges vegyületek is. Ez utóbbiak esetében szekvenciális kopolimer típusú vegyületek jöhetnek szóba, vagy szekvenciális polikondenzátumok.

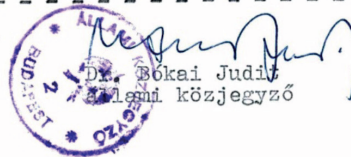


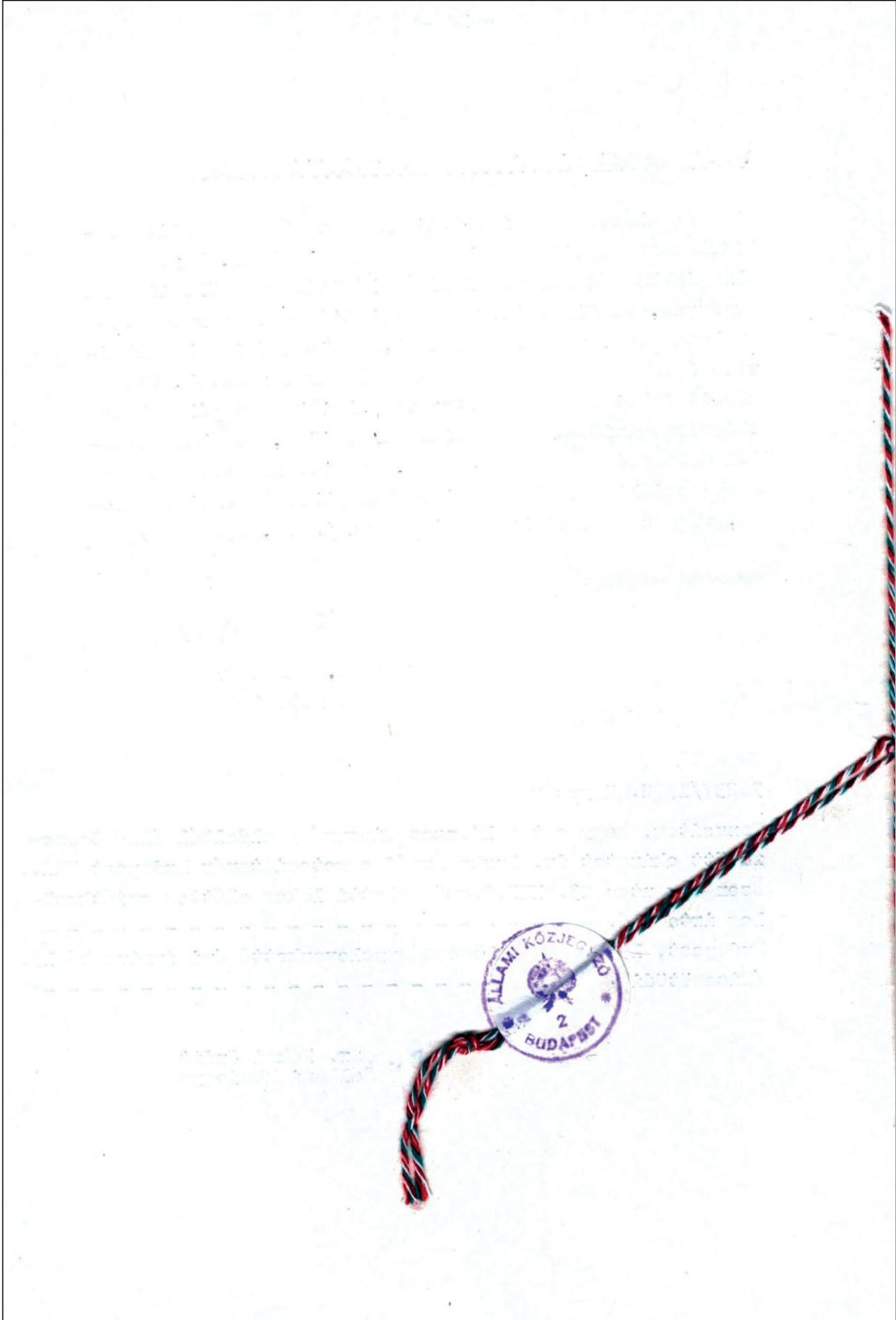
Furka Árpád

Dr. Furka Árpád
egyetemi tanár

36237/1982. ügyszám

Tanusítom, hogy ezt a 14.szaz Tizennégy oldalból álló össze-
fűzött okiratot dr. Furka Árpád egyetemi tanár Budapest VII.,
Csengery utca 23. III.2.szám alatti lakos előttem sajátkezü-
leg írta alá. -----
Budapest, 1982. Ezerkilencszáznolcvankettő évi június hó 15.
/Tizenötödik/ napján. -----





The first description of the principles of combinatorial chemistry

Author: Professor Árpád Furka

Eötvös Loránd University Budapest Hungary

English version

The original Hungarian document was notarized in June 15, 1982

**STUDY ON POSSIBILITIES OF SYSTEMATIC SEARCHING FOR
PHARMACEUTICALLY USEFUL PEPTIDES**

Written by Dr. Árpád Furka, university professor

Budapest, May 29, 1982

As exemplified, among others, by the peptide hormones discovered so far, the shorter-lengthier peptides take part in a number of important functions in the living organism. It can be supposed, that only a small fraction is known of these biologically active peptides having potential therapeutic effect. This fact motivates the intensive international and domestic research activity in this field.

Two, in principle different, approaches offer themselves for searching for peptides bearing new biological effects:

1. Isolation of peptides from living organisms based on their previously known biological effects.
2. Preparation of peptides by synthesis with post determination of their biological effects.

Until now the isolation procedure proved to be more effective in spite of the fact that this method is also very laborious. This may be explained by the fact that the number of possible peptides grows rapidly with the number of residues so even the synthesis of all tetrapeptides (160 thousands) seems to be a hopeless task. If we consider the 20 natural

amino acids the dependence of the number (N_n) of possible peptides on the number of residues (n) is expressed by the following formula:

$$N_n = 20^n$$

If the n -residue peptides are synthesized stepwisely and independently, the number of the required synthetic steps (S_n) can be calculated as follows:

$$S_n = (n-1) 20n$$

It is noted, that a synthetic step means a complete coupling cycle, that is, in addition to the coupling step itself incorporates the operations connected with the protecting groups, too.

With good organization, that is, choosing a systematic synthesis route the number of synthetic steps can be reduced. The minimum number of synthetic steps is:

$$S_n = \sum_{i=2}^n 20^i$$

The synthesized peptides are supposed to be submitted to screening tests. Since several tests have to be done on each peptide, the total number of the required screening tests is hopelessly large. If the number of kinds of screening tests is denoted by t , the total number of screening tests is expressed by the following equation:

$$T_n = t 20^n$$

Table 1 shows the the possible number of peptides depending on the number of residues, the number of synthetic steps required for their synthesis, and number of the screening tests, calculating with 10 different tests ($t=10$). The figures - which are rounded - clearly show, that even the synthesis and testing of all tripeptides would be an almost

Table 1

Possible number of peptides (N_n) containing different number of residues (n), the number of synthetic steps required for their synthesis (S_n) in an optimized process, furthermore the number of screening experiments (T_n) calculating with 10 different screening tests ($t=10$) (the figures are rounded)

n	N_n	S_n	T_n
2	4 hundred	4 hundred	4 thousand
3	8 thousand	8 thousand	80 thousand
4	160 thousand	168 thousand	2 million
5	3 million	3 million	30 million
6	64 million	67 million	640 million
7	1 billion	1 billion	13 billion
8	25 billion	26 billion	256 billion
9	512 billion	537 billion	5 trillion
10	10 trillion	10 trillion	102 trillion

Because of the very large number of possible peptides, the stepwise synthesis of all peptides - even in the case of small ones - is an unrealizable task. The large number of the screening experiments constitutes a further problem. The proposal to be outlined on the next pages will try to somewhat improve this almost hopeless situation.

Systematic search for biologically active small peptides through synthesis and screening of peptide mixtures

The proposal to be outlined here constitutes a research project which makes possible to search for biologically active peptides with much greater chance than before. When I write down this project I'm fully aware of its potential importance in industry. It is also clear, that its realization is possible only through cooperation of different institutions. Primarily the participation of the pharmaceutical industry is desirable since the investments can be recovered through pharmaceutical industry.

The essence of the proposal is that instead of one by one synthesis of peptides, peptide mixtures should be prepared containing several hundred or several thousand peptides in approximately 1 to 1 molar ratio, and these peptide mixtures should be submitted to screening tests. It will be shown that on this way much labor can be saved both in the synthetic work and in the screening experiments. In the first stage one has to determine whether or not the mixture shows any biological effect. If biological effect is observed, of course, it has to be determined which component (or which components) are responsible for the activity.

Method for synthesis of peptide mixtures

Since not single peptides but rather mixtures of peptides are synthesized, post synthetic purification and removal of by-products are out of question. Because of this, the classical method of synthesis (in solution) can not be used either. In the synthesis of peptide mixtures the solid phase method have to be applied. It is noted here, that in the syntheses not necessarily the 20 amino acids are used. In some cases more than 20 amino acids may be used, for example if - in addition - non-common amino acids are intended to be used as building blocks. Less than 20 amino acids may be used, for example, in decapeptides, since the synthesis of all peptides seems to be unrealistic and have to compromise with the use of fewer kinds of amino acids. Let denote by k the number of the amino acids intended to vary in the i position. The numbers of amino acids varied in the C-terminal and N-terminal position are k_1 and k_n , respectively.

Realization of the synthesis

The resin is divided into k_1 equal portions (that is to as many portions as many amino acids are intended to vary at the C-terminal of peptides). Then each portion of resin is coupled with one of the k_1 kinds of amino acids then the amino-protecting group is removed from every sample. A small quantity is removed from every sample and they are taken aside for later use, then the samples are thoroughly mixed. Then the mixture of aminoacyl resins is divided into k_2 equal portions and each of them is coupled with one of the k_2 kinds of protected amino acids then the amino-protecting groups are removed from each sample. Before mixing, again small samples are removed and taken aside. The mixture of dipeptides is cleaved from a small portion of the mixed resin to use it in biological tests. The rest of the mixed resin is divided into k_3 equal parts and the amino acids intended to occupy the third position are coupled to them. Then the synthesis is likewise continued until the mixture of n -residue

peptides is reached.

It is worthwhile to add some notes. As in an ordinary solid phase synthesis, one has to make an effort to achieve good conversion by applying the reagents in excess. Fortunately, however, conversions lower than 100%, or minor unwanted splitting reactions do not cause so serious problems like in ordinary syntheses. The labour requirement could be significantly reduced by using mixtures of properly protected amino acids in acylation reactions. This, however, does not seem to be an acceptable solution because of the differences in the reactivity of the activated amino acids which would lead to the formation of peptides in significantly different concentrations thus causing problems in the screening experiments. Formation of peptides in equal concentrations can only be assured by mechanical mixing of samples followed by dividing into equal portions. This makes possible a complete conversion for every amino acid component. Possibility of acylations with mixtures of several amino acids of identical reactivity might be a matter of further considerations. Smaller differences in reactivities could be compensated by properly selected molar ratios of the amino acid derivatives of the mixture. In the following calculations the possibility of acylations with the mixtures of amino acid derivatives will be left out of considerations.

The number of peptides formed in the synthesis, that is, the number of components in the peptide mixtures - in a general case - can be calculated by the following formula:

$$N_n = k_1 \cdot k_2 \cdot \dots \cdot k_{n-1} \cdot k_n$$

If the same number (k) of amino acids are varied in every position

$$N_n = k^n$$

The number of synthetic steps in the synthesis of a peptide mixture containing N_n peptides (considering the attachment of the first amino acid to the resin as separate step) is:

$$S_n = k_1 + k_2 + \dots + k_{n-1} + k_n$$

If the same number (k) of amino acids are varied in each position,

$$S_n = nk$$

The formulae show the advantage of the synthesis of peptide mixtures: the number of the synthetic steps can be calculated by summing the numbers of the varied amino acids, while the number peptides is given by the product of the numbers of the varied amino acids.

One example: synthesis of the mixture of tetrapeptides prepared by varying the 20 kinds of amino acids, needs only 80 synthetic steps! It is noted, that in the same run all shorter peptides - that is the 400 dipeptides and the 8000 tripeptides - are formed, too. The traditional synthesis of these peptides would need 168 400 synthetic steps. A different comparison: in the traditional method with 80 steps only about 30 tetrapeptides can be synthesized.

Screening of peptide mixtures

Peptides mixtures - in the first approximation - are synthesized to determine whether or not they contain biologically active component. It is supposed - although it needs experimental verification - that screening experiments can be made with mixtures, too. This offers great advantage over the traditional method since the number of screening tests is reduced by a factor equal to the number of components of the mixture. For example, the mixture of the 8000 tripeptides can be examined by a single series of tests. If there is active peptide among them, one of the executable t tests gives positive result. If the number of active peptides is more than one, then, of course, more tests may give positive result. In the synthesis of the mixture of n -residue peptides it is worthwhile to test the shorter peptides, too. The synthesis is

so designed to allow for this. Taking this requirement into account, and the number of kinds of tests being t , the total number of the executable tests is:

$$T_n = t(n-1)$$

Although this equation certainly holds, its realizability in practice deserves some notes. There is - without any doubt - an upper limit in the number of components of the peptide mixtures to be submitted to screening tests. It is difficult to estimate this number without experiments. The mixtures may probably contain many thousands of components, and as it can be judged today, the method outlined above is rather limited by possibilities of screening tests than by the number of the required synthetic steps. If there are too many components in the mixture, too large samples have to be applied in the screening experiments to achieve observable effect for a single component. The mixture supposedly contains a number of more or less active analogs and their effect is probably summarized. Nevertheless, an unsurpassable limit in the number of components certainly exists. Therefore in certain cases may prove useful to examine the effect of the n -residue mixtures without final mixing. In other cases the synthesis should be designed so not to surpass the optimal number of components.

"Backsearching" for the active peptide

If the peptide mixture is detected to contain active component, that is, if the mixture shows a new type biological effect, then the further task is the isolation and structure determination of the active peptide followed by its synthesis. Once the mixture containing the active component or components is in our hand the isolation can be carried out using the effective separation methods, since these make possible to separate the active compound even from thousands of inactive components. It is possible, however, to follow a different method, too. This will be outlined here. This approach to the identification of the active peptides is supposed to be less tedious than the isolation method, moreover it supplies additional

information concerning the structure-effect relationship. Applicability of the method requires a procedure for quantitative determination of activity. For the sake of simplicity let's suppose that the mixture contains a single effective component (besides analogs having the same kind of effect but smaller activity).

Backsearching step No. 1

The experiments are started with the k_n samples taken aside in the synthesis of the n -residue peptides before final mixing. The mixtures of n -residue peptides are cleaved from each resin sample. The mixtures of peptides differ from each other in the n -th (that is the N-terminal) residue of their component peptides. Each peptide mixture is submitted to a quantitative activity determination. This shows how the activity depends on the terminal amino acid residue, that is, this way we can determine the N-terminal residue of the active peptide, and in addition it will show the effect of its replacement by other amino acid residues. Let's suppose, for example, that the N-terminal residue in the sample showing the highest activity (as well as in the active peptide) is Phe (phenylalanine). It is noted here that if there are several samples showing equally high activity it is practical to choose as the N-terminal residue of the active peptide the cheapest or the synthetically less problematical amino acid. This note holds for the subsequent backsearching steps, too.

Backsearching step No. 2

The experiment is continued with the k_{n-1} samples taken aside in the synthetic stage of the $(n-1)$ -residue peptides. The amino acid determined before, that is Phe in our example, is coupled to each sample. Cleavage of the peptides from the support gives k_{n-1} different peptide mixtures. Their common feature is that every peptide has Phe in the N-terminal position. By submitting the peptide mixtures to quantitative screening experiments one can determine the amino acid residue occupying position $n-1$ (that is, the pre-N-terminal position) in the active

peptide. This experiment also shows the effect on activity of substitution of this amino acids with other ones. Let's suppose that the pre-aminoterminal amino acid is Arg (arginine). It should be noted that in this backsearching step the Phe is coupled to k_{n-1} samples and the same number (k_{n-1}) of screening experiments have to be done. Not all of the t kinds of tests are required, only the one proved before to be positive. Consequently the number of the synthetic steps and the number of screening experiments are the same: k_{n-1} . It is also noted that in the previous backsearching step only screening test are done (their number is k_n) synthetic steps are not needed.

Backsearching step No. 3

This, and the subsequent backsearching steps may be realized using two different approaches. The peptides in samples taken aside during the synthesis have to be elongated to contain n residues, in such way, to carry on their N-terminal section the amino acid residues assuring activity. This can be realized on two ways. Either by stepwise coupling with amino acids (in our example with protected Arg then Phe) or by coupling in a single step with a previously synthesized oligopeptide having the required sequence (in our example Phe.Arg). The required synthetic steps in the two approaches significantly differ. The number of the screening experiments, however, are the same in both cases. Let's turn now to the No. 3. backsearching step.

Stepwise elongation

Let's take the k_{n-2} samples taken aside in the synthesis of $(n-2)$ -residue peptides. Each sample is coupled first with protected Arg then with protected Phe. After cleaving the peptides from the support each of the k_{n-2} peptide mixtures are submitted to activity tests to determine the amino acid residue occupying the third position counting from the N-terminal end. The number of screening tests to be executed is k_{n-2} . The number of the required synthetic steps is:

$2k_{n-2}$. The multiplying factor preceding k is the bigger the shorter are the peptides to be elongated. The numerical value of the factor is equal to the number of amino acids to be coupled with in the elongation process.

Elongation with oligopeptide

A previously synthesized dipeptide (in our example Phe.Arg) is coupled to each of the k_{n-2} samples taken aside, then the process is continued as described above. The number of screening test is also k_{n-2} . The number of synthetic steps (leaving out of consideration the synthesis of the oligopeptide) is also k_{n-2} . This procedure seems to be more economical. In practice it means that the active peptide is synthesized in parallel with the screening tests using the classical method started from the N-terminus. Small fractions of the growing peptide are sacrificed in the backsearching steps. This backsearching method has the great advantage (in addition to the fact that it needs less synthetic steps) that when the backsearching procedure is finished the active peptide is synthesized, too.

Backsearching of more than one active peptide

In the synthetic peptide mixtures several active peptides may be present, showing different effects. In these cases the number of backsearching steps will be bigger by a factor equal to the number of the differing active peptides. That is, if the number of the active peptides is " a " the values deduced above are multiplied by a . It is noted that the presence in the mixture of peptides having different effects may complicate the backsearching process especially in the case of peptides with opposing effects. This, however, is not treated in details.

The backsearching process ends when the sequences of all active peptides are determined by applying either the oligopeptide or the stepwise elongation method.

**Total number of synthetic steps and screening tests summarized
for the whole synthetic backsearching process**

Number of synthetic steps using oligopeptide elongation

In synthesis:
$$S_n = \sum_{i=1}^n k_i$$

In backsearching:
$$S_n = a \sum_{i=1}^{n-1} k_i$$

Total in synthesis and backsearching:
$$S_n = (a + 1) \sum_{i=1}^{n-1} k_i + k_n$$

If k amino acids are varied in each step:
$$S_n = [n(a + 1) - a]k$$

Number of synthetic steps using stepwise elongation.

In synthesis:
$$S_n = \sum_{i=1}^n k_i$$

In backsearching:
$$S_n = a \sum_{i=1}^{n-1} (n-1)k_i$$

Total in synthesis and backsearching:
$$S_n = \sum_{i=1}^n k_i + a \sum_{i=1}^{n-1} (n-1)k_i$$

If k amino acids are varied in each step:
$$S_n = nk + ak \sum_{i=1}^{n-1} i$$

Number of screening tests equally valid using the oligopeptide and stepwise elongation

In synthesis: $T_n = t(n-1)$

In backsearching: $T_n = a \sum_{i=1}^n k_i$

Total in synthesis and backsearching: $T_n = t(n-1) + a \sum_{i=1}^n k_i$

If k amino acids are varied in each step: $T_n = t(n-1) + ank$

An example: preparation and screening of all pentapeptides

$$N_5 = 320000, \quad n=5 \quad k=20 \quad t=10 \quad a=1$$

Total number of synthetic steps

Ogopeptide elongation	180
Stepwise elongation	300
Number of tests	140

Extension of the method to other types of compounds

Applicability of the method outlined before is not restricted for only the systematic searching for active peptides. The same principle applies to all other sequential types of compounds,

that is, when the compounds belonging to this type of compounds differ from each other only in their building blocks or the sequences of these building blocks. Among them may occur natural compounds like oligosaccharides or oligonucleotides but synthetic products may be taken into account, too. Among these later ones one may think about sequential copolymers or sequential polycondensates.

Dr. Arpad Furka

university professor

File number 36237/1982

I certify this stitched document comprising 14, that is, fourteen pages was subscribed in my presence by Dr. Arpad Furka, university professor, with his own hands.

Budapest, 1982. Nineteen hundred and eighty two, June 15, (fifteen).

Dr. Judit Bokai

state notary public