

# Kémiai térképek felvételi módszerei és elemzése kemometriai módszerekkel

VAJNA Balázs,<sup>a,\*</sup> SZEPESVÁRY Pál,<sup>b</sup> KEGLEVICH György<sup>a</sup> és MAROSI György<sup>a</sup>

<sup>a</sup>BME Szerves Kémia és Technológia Tanszék, Budafoki út 8, 1111 Budapest, Magyarország

<sup>b</sup>ELTE Fizikai Kémia Tanszék, Pázmány Péter sétány 1/A, 1117 Budapest, Magyarország

## 1. Bevezetés

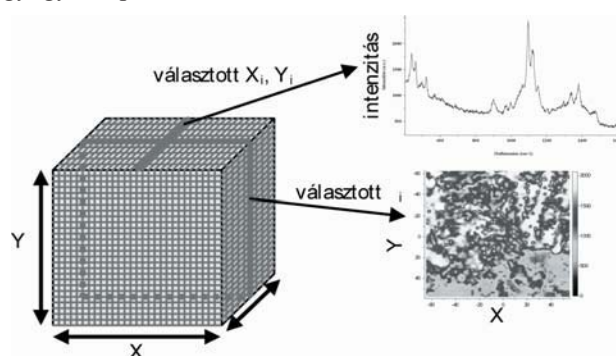
A rezgési spektrometria számos mérési módszert felölel, melyek közül legkiterjedtebben a közeli infravörös (near infrared, NIR), a közepes infravörös (mid-infrared, MIR) és a Raman-spektrometria használatos. A készülékek folyamatos fejlődése lehetővé tette egyre komplikáltabb vizsgálatok elvégzését, illetve lehetőségessé vált szilárd fázisú minták vizsgálata. Közleményünkben egy új rezgési spektrometriai módszert, a kémiai térképezést mutatjuk be, mellyel a vizsgált minták alkotóelemeinek térbeli koncentráció-eloszlása vizsgálható és vizualizálható.

Barer és Cole 1949-ben a Nature-ben közölte az első olyan vizsgálatot, ahol a spektrométert egy mikroszkóppal összekötve néhány mikrométer átmérőjű pontról vettek fel spektrumot.<sup>1</sup> Eredményeik utat nyitottak mikroszkopikus minták megbízható spektrometriai vizsgálatára. Mintegy négy évtizeddel később került sor az első „kémiai térkép” felvételére egy MIR spektrométer, egy mikroszkóp és egy mozgatható tárgyasztal segítségével.<sup>2</sup> A mintáról így egyszerre szerezhető térbeli és spektrális információ: lehetségessé válik különböző vegyületek detektálása és azonosítása a minta eltérő pontjain, illetve képet alkothatunk egy minta térbeli szerkezetéről.

Mára a kémiai térképezés (angol nevén „chemical imaging”, rövidítve: CI) a rezgési spektrometria egyik legdinamikusabban fejlődő területévé vált. A tudomány és az ipar számos ágában nyer egyre komolyabb jelentőséget ez az ágazat, a félvezetőipartól<sup>3</sup> kezdve az orvostudományon,<sup>4</sup> papíriparon,<sup>5</sup> élelmiszeriparon<sup>6</sup> és műanyagiparon<sup>7</sup> át egészen a gyógyszerkészítmény-technológiáig.<sup>8,9</sup>

A kémiai térképek kihívást jelentenek nemcsak a vizsgáló berendezés felépítése, hanem az adatelemzés tekintetében is. Felépítésük az 1. ábrán látható: a vizsgált felület mindegyik pontjához egy rezgési spektrum rendelhető. Ezáltal egy háromdimenziós szerkezet jön létre, melyet az irodalomban *hiperspektrális adatkockának* (*hyperspectral data cube*)<sup>8</sup> is neveznek. A struktúra másik lehetséges értelmezése (ill. felvételi módja, ld. később), hogy az adatkocka minden lapját egy meghatározott hullámhosszhoz tartozó „fényképnek” feleltetünk meg. Az adatrendszer ennek megfelelően óriási méretű, és fejlett (többváltozós, ún. *kemometriai*) technikákat igényel a kiértékeléshez. (Az itt bemutatásra kerülő kemometriai algoritmusok jól alkalmazhatók más spektrometriai módszerek, pl. tömegspektrometria,<sup>10</sup> NMR spektrometria<sup>11</sup> esetében is.)

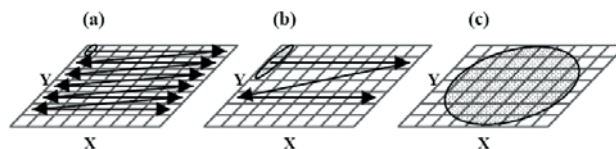
E tanulmányban a szerzők célja bemutatni a kémiai térképezés lehetséges felvételi módszereit és az adatelemzés lehetőségeit, illetve egy, a kutatócsoportunkban végzett vizsgálat példáján illusztrálni a módszer alkalmazhatóságát gyógyszeripari területen.



1. Ábra. Kémiai térképek adatszerkezete.

## 2. Kémiai térképezés alapjai

A kémiai térképezés megvalósításához spektrométerre és megfelelő optikára van szükség. Noha elméletileg számos spektrometriás módszer térképezési metodikává fejleszthető, a „kémiai térképezés” kifejezés alatt elsősorban rezgési (infravörös és Raman) spektrometriai technikákon alapuló módszereket értünk.<sup>9</sup> E módszerek közül is legelterjedtebben a közeli infravörös (NIR; 13000-4000  $\text{cm}^{-1}$  ill. 0,76-2,6  $\mu\text{m}$ ), a közepes infravörös (MIR; 4000-400  $\text{cm}^{-1}$  vagy 2,6-26  $\mu\text{m}$ ) valamint a Raman (kialakítástól függően változó hullámhossztartomány, általában  $\sim 4000\text{-}150\text{ cm}^{-1}$ ) technikákat alkalmazzák.



2. Ábra. térképek felvételének lehetőségei: (a) pontontkénti térképezés (point mapping); (b) vonalas térképezés (line mapping); (c) teljes felületi képalkotás (global imaging).

A kémiai térképek háromdimenziós szerkezetűek ugyan, azonban jelenleg az összes spektrum egyidejű felvétele egyszerre mindegyik mérési pontban nem lehetséges.

\* Telefon: (1) 463-5918 ill. (30) 310-2209; fax: 463-3648; e-mail: balazs.vajna@gmail.com

Amint a 2. ábra mutatja, háromféle mérési elrendezés áll rendelkezésre. Az egyik lehetőség (a) a teljes hullámszámtartományt fellelő spektrumok felvétele pontról pontra haladva egy mozgatható tárgyasztal segítségével, mely technika a továbbiakban „térképezés” (*mapping*) néven szerepel a közleményben. A térképezés másik (b) formája, amikor egy rezgő tükör segítségével nem egyetlen pontról, hanem egy vonal mentén egyszerre több pontról veszünk föl spektrumot. A harmadik módszer (c) másféle berendezés-kialakítást igényel; alkalmazásakor a minta elemezni kívánt felületének egészéről készül kép egyetlen megadott hullámhosszon („képkalkotás”, angolul *global imaging*). A hullámhossz változtatásával nagy számú kép készíthető, és ezek „egymásra helyezésével” jön létre az adathalmaz háromdimenziós szerkezete. A térképezéses módszer (*mapping*) elsősorban a Raman spektrometriában elterjedt, a teljes felületi képkalkotás (*global imaging*) pedig elsősorban a közeli infravörös spektrometriában vált szokásossá, de számos példa található a másik két kombinációra is.

### 3. Kémiai térképezést szolgáló berendezések felépítése

Míg a pontonkénti és a vonalas térképezés legtöbbször ugyanazzal az eszközzel, kis módosításokkal megoldható, a felületi képkalkotás teljesen más alkatrészeket és összeállítást igényel. Mindazonáltal minden CI rendszer sematikusán hasonló felépítést mutat. A rendszer egy megvilágító fényforrásból (infravörös fény vagy infravörös-látható lézer), képkalkotó optikából, monokromátorból (a hullámhossz kiválasztására vagy hullámhossz szerinti felbontásra), egy detektorból és egy vezérlő egységből áll, mely utóbbi általában egy megfelelő szoftverrel ellátott asztali számítógép.<sup>12</sup>

#### 3.1. Pontonkénti és vonalas térképezés

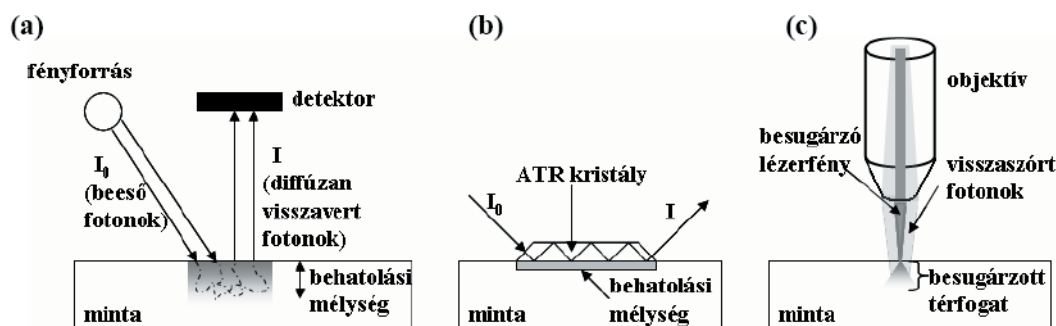
A térképezéses módszerek már említett közös jellemzője, hogy egy mozgatható tárgyasztalra van szükség a minta kontrollált mozgatásához. Ezeket a berendezéseket általában egy spektrométer és egy mikroszkóp alkotja<sup>12</sup>. A működés elve a következő: egy adott pontról megtörténik a

spektrum felvétele, majd a tárgyasztal odébb lép egy előre meghatározott távolsággal, és egy újabb spektrumot vesz fel a készülék. Ez addig ismétlődik, amíg a kívánt felület egésze pásztázásra nem kerül.

Gyógyszerek IR spektrometriás vizsgálatakor a pontbeli térképezést jellemzően két technika, a diffúz reflexió (DR) és a gyengített (csillapított) teljes reflexió (ATR) segítségével valósítják meg.

A beeső fény kölcsönhatásba lép a mintával azáltal, hogy szóródik a mintát alkotó részecskéken, illetve ez a szórt fény tovább szóródhat a szomszédos részecskéken (3.(a) ábra). Ezt a jelenséget nevezik diffúz reflexiónak, amiatt, hogy a szóródás mértéke és mikéntje fotononként eltérő. A szórt fény egy része visszajut az optikára és onnan a detektorra (természetesen a fénynek főként olyan hullámhosszú komponensei, amelyek nem nyelődtek el visszaverődés közben). A közepes infravörös (MIR) tartományban az elnyelés olyan erős, hogy a vizsgálandó mintát inert anyaggal hígítani kell, ennélfogva a módszer nem alkalmas gyógyszerkészítmények vizsgálatára.<sup>8</sup> A közeli infravörös (NIR) hullámhosszú fény esetében azonban a gyenge elnyelés miatt ez a technika kiválóan alkalmazható például tabletták analizésére.<sup>13,14</sup>

A másik elterjedt módszer a gyengített teljes reflexió technika, melyet a közepes infravörös (MIR) spektrometriai vizsgálatokhoz fejlesztettek ki, hogy ne legyen szükség a minták inert anyaggal (pl. KBr) történő hígítására. A jelenség azon alapszik, hogy a fény megfelelően kicsi szögben érkezik két anyag közötti határfelületre teljes visszaverődést szenved. (E szög nagysága a két közeg törésmutatója közötti eltérés nagyságától függ.) A fotonok azonban a hullámhosszukkal arányos mértékben, pár mikrométer távolságig behatolnak a másik közegbe (a gyakorlatban tehát a vizsgálandó mintába), ahol egy részük – a hagyományos spektrometriában megszokott módon – elnyelődhet. A módszer alkalmazható porok és tabletták vizsgálatára,<sup>15,16</sup> illetve akár a tablettapréselés in situ nyomon követésére is.<sup>17</sup>



3. Ábra. Pontbeli térképezés lehetséges módszerei: (a) diffúz reflexió (MIR, NIR); (b) gyengített teljes reflexió (MIR, NIR); (c) optikai fókuszálás (Raman).

A Raman spektrometrián alapuló térképezés jóval egyszerűbb. Sokszor egyáltalán nincs szükség mintaelőkészítésre, a mintát egyszerűen a tárgyasztalra kell helyezni, és a mikroszkóphoz kapcsolt, optikai képet szolgáltató kamera segítségével a fénysugarat a minta

felületére kell fókuszálni. A fókuszálást nagy nagyítású objektívek alkalmazásakor automatikus fókuszáló rendszer segítheti. A látható és közeli infravörös hullámhosszú lézertörések lehetővé teszik a minták üveg, műanyag felületén és mérőcellákban történő vizsgálatát.

A detektorok fejlődésével lehetségessé vált a pásztázási idő rövidítése. Ha a diódasoros detektor nem csak egy, hanem több sornyi diódát tartalmaz, akkor egy vonal megvilágításával annak több pontjáról egyszerre vehető fel spektrum, lehetővé téve a vonalas térképezés megvalósítását.<sup>18</sup> Ekkor a megvilágított vonalról visszazórt fotonok a mikroszkóp optikáján keresztül a monokromátorra jutnak. A monokromátor vonalas térképezésre is alkalmas berendezések esetén általában egy optikai rács. A rácsra érkező vonal felbomlik különböző hullámhosszúságú komponenseire, egy két dimenzióban szétterjedő fénysugarat alkotva. Detektorként ennek megfelelően többsoros diódasort kell alkalmazni, amelynek minden sora megfelel a vonal egy pontjának. Diódasoros detektor helyett általában CCD detektort használnak (*charge coupled device*, „töltéscsatolt eszköz”). Fourier-transzformációs mikrospektrométereknél nagyobb hullámhosszakra érzékeny ún. *focal plane array* (FPA) detektort alkalmaznak. A számítógép mérés után együtt tárolja a spektrumokat a vizsgált pontok (vonalak) térbeli helyzetét jellemző információval együtt, így ezekből létrehozható a bevezetésben ismertetett adatszerkezet.

### 3.2. Teljes felületi képalkotás

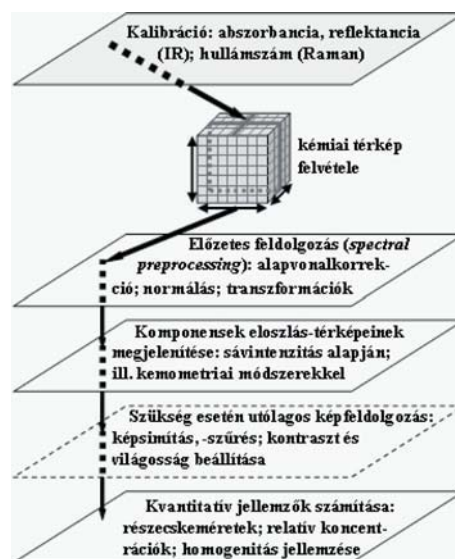
E technika fő előnye a pontonkénti térképezéshez képest az, hogy nem tartalmaz mozgó alkatrészeket. Ebből fakadó hátránya természetesen az, hogy nehezebben kontrollálható a vizsgált felület mérete, és érdes felületek vizsgálatakor értelemszerűen fellép a felület heterogén fókuszáltságából eredő zavarás.

A mintát diffúz NIR fényvel (NIR-CI), vagy kitágított lézer sugárforrással (Raman-CI) világítják meg. A mintáról visszazórt fény egy szűrőn halad keresztül, amely monokromátorként funkcionál, és egyszerre csak egy szűk hullámhossz-tartományú fényt enged keresztül. A régebbi CI eszközöknél egy sorozat sáváteresztő szűrőt építettek be a készülékbe, gyakran egy forgatható kör mentén az egyikről másira kapcsolás egyszerűsítése végett. Az új kialakítású készülékek hangolható szűrőket tartalmaznak, amelyek számítógéppel – azaz elektromos jelek segítségével – vezérelhetők, így gyorsabb váltást és sokszorta több hullámhossz kiválasztását teszik lehetővé. Kétféle típusuk közül az akusztó-optikai hangolható szűrő (AOTF) a NIR rendszerek esetében terjedt el,<sup>19</sup> a folyadék-kristályos hangolható szűrő (LCTF) pedig a Raman képek szűrésében tűnik ígéretesnek.<sup>20</sup> Az innen továbbjutó fény egyenesen a detektorra érkezik, amely a térképező készülékekhez hasonlóan NIR esetében FPA, Raman esetében pedig CCD detektor. A különböző hullámhosszú-kon felvett képek egymásra helyezésével jön létre az adatok háromdimenziós elrendezése.

## 4. Kémiai térképek elemzése

A kémiai térképek gyakran óriási mennyiségű adatot tartalmaznak, amelyeknek nyers formában történő, illetve minden fontos részletre kitérő vizuális értékelése szinte teljességgel lehetetlen. Ennek megfelelően szükség van egyrészt előzetes feldolgozásra, amely a zavaró tényezőket eltávolítja, másrészt pedig olyan matematikai módszerekre, amelyek hasznos vizuális vagy éppen kvantitatív információ

válogatnak ki az adathalmazból. A kémiai térképek feldolgozásának jellemző lépéseit a 4. ábra mutatja.



4. Ábra. Kémiai térképek feldolgozása és kiértékelése.

### 4.1. Kalibráció

A válaszjelek intenzitásmérésének kalibrációjára mind az IR, mind a Raman mérések esetén szükség van. A reflektancia kalibrációjakor kémiai térkép készül egy nagy fényvisszaverő képességű anyagról (100%), illetve készül egy térkép fény nélkül is, a visszavert fényt összegyűjtő lencsét eltakarva (0%). Emellett léteznek többpontos, megbízhatóbb kalibrációs módszerek is.

A Raman spektrometria során jóval egyszerűbb a helyzet, mert a válaszjel egyszerűen a beérkező fotonok számával arányos, amelyet nem szükséges (bár lehetséges) kalibrálni. Ezeknél a rendszereknél a hullámhossz-skála pontosságát kell bizonyos időközönként ellenőrizni egy állandó, jól meghatározott sávval rendelkező standard (pl. szilícium egykristály) segítségével.

### 4.2. Előzetes adatfeldolgozás

A kémiai térképek előzetes feldolgozása lényegében a megfelelő spektrometriai módszerekben megszokott algoritmusokkal történik. A művelet célja a különböző zavarásokból eredő jelek (pl. felületi egyenetlenség hatása, fluoreszcencia, detektor zaj) szűrése és eltávolítása a térképből. Raman spektrumok esetén jellemzően alapvonal-korrekciót és normálást kell alkalmazni, esetenként a spektrumok deriválása is hasznos lehet. A NIR spektrumok leggyakoribb módosítása a *standardizálás*<sup>20</sup> (SNV) és az ún. *multiplicative scatter correction*<sup>21</sup> (MSC). Különböző tömörítési eljárásokat is alkalmaznak a térképezéskor keletkező adatfájlok méretének csökkentése és a feldolgozó algoritmusok gyorsítása céljából.<sup>22</sup>

### 4.3. Komponensek eloszlás-térképeinek megjelenítése

Az előző fejezetekben ismertetett lépések mindegyike azt a célt szolgálja, hogy megjeleníthető legyen a komponensek térbeli eloszlása a vizsgált felületen. Mivel a kémiai

térképek óriási mennyiségű adatot tartalmaznak, vagy egy jellemző részletre fókuszáló egyváltozós, vagy a változók sokaságát elemző sokváltozós statisztikai módszereket kell alkalmazni a fontos paraméterek azonosításához vagy meghatározásához.<sup>23</sup>

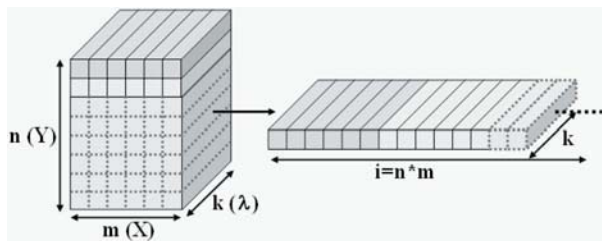
### 4.3.1. Egyváltozós kiértékelési módszerek

A legegyszerűbb módszer a komponensek térbeli eloszlásának meghatározására, ha valamelyik komponensre jellemző sáv intenzitását ábrázoljuk a hely függvényében. Az így kapott kép elvileg arányos a vizsgált komponens koncentráció-eloszlásával. Sávintenzitás helyett ábrázolhatjuk két hullámhossz (hullámszám) közötti sáv(ok) alatti területet, akár két sávintenzitás/sávterület hányadosának értékét is. Ezeket és a hasonló módszereket hívjuk a térképezésben alkalmazott egyváltozós eljárásoknak.<sup>8</sup>

E módszerek azonban csak bizonyos esetekben vezetnek jó eredményre. Legfontosabb követelmény, hogy az elemezni kívánt komponens rendelkezzen intenzív rezgési sávval. Többkomponensű rendszerek esetén – különösen NIR spektrumok esetében – további probléma, hogy az egyes komponensek rezgési sávjai átlapolnak egymással: igen nehéz olyan sávot kiválasztani, amely kizárólag egyetlen komponensre jellemző.

### 4.3.2. Sokváltozós kiértékelési módszerek

A problémára megoldást jelentenek a többváltozós adatelemzési (kemometriai) módszerek, melyeknek számos fajtája alkalmazható a kémiai térképek kiértékelésére. Ezek a módszerek többnyire kétdimenziós adatmátrixok feldolgozására lettek kifejlesztve, így a kiértékelés előtt a három-dimenziós adatszerkezet kétdimenzióssá alakítása, ún. *feltárása* (*unfolding*) szükséges. Ezt az egyik térbeli dimenzió másik mellé rendezésével érjük el, ahogy az az 5. ábrán látható. A kemometriai algoritmusok lefuttatása után az eredményt tartalmazó adathalmazt visszarendezik három-dimenzióssá, amennyiben ez szükséges. (Mára már olyan adatfeldolgozó módszerek is léteznek, melyek három- vagy többdimenziós adatszerkezeteket is kezelni tudnak, de a kémiai térképezés területén nem bizonyultak olyan hatékonyak, mint a tradicionális mátrixkezelő algoritmusok<sup>8</sup>).



5. Ábra. A kémiai térkép „feltárása”, a háromdimenziós adatszerkezet kétdimenziós mátrixszá konvertálása.

Az algoritmusok alapvetően két típusra oszthatók. Az egyik típust az *a priori* ismereteket igénylő módszerek alkotják, amelyeknél felhasználunk a mintával kapcsolatos előzetes információkat (mintát alkotó komponensek spektrumai, kalibráló sorozatok ismert koncentrációkkal, stb.). Az algoritmusok másik csoportjába az *a priori* ismereteket

nem igénylő módszerek tartoznak, melyek alkalmazásakor semmilyen előzetes tudás nem szükséges a vizsgált mintáról, minden információ a kémiai térkép által szolgáltatott adatrendszerből származik. Így anélkül vizualizálható a legfontosabb információhordozók (komponensek) térbeli elhelyezkedése, hogy ismernénk magukat az alkotóelemeket. E módszerek előnye, hogy megbízhatóságuk számítható, illetve képesek kvantitatív információ szolgáltatására (pl. komponensek koncentrációinak meghatározására a térkép pontjaiban).

#### 4.3.2.1. *A priori* ismereteket felhasználó módszerek

Számos esetben a kémiai térképet alkotó spektrumok felírhatók az alkotó komponensek spektrumainak lineáris kombinációiként. Ez a közelítés jól ismert a klasszikus rezgési spektrometriában. Amennyiben  $p$  komponens van jelen egy pontban, melyről a spektrum készül:

$$\mathbf{X} = \mathbf{C}\mathbf{S}^T + \mathbf{E} \quad (1)$$

ahol  $\mathbf{S}^T (p \times k)$  a tiszta komponensek referencia-spektrumaiból képzett mátrix. Minden spektrum  $k$  intenzitásértékből áll, és egy sort alkot ebben a mátrixban.  $\mathbf{X} (i \times k)$  a feltárás (ld. 5. ábra) után a Raman-térképből kapott mátrix, mely az összesen mért  $i$  darab spektrumot tartalmazza sorok formájában. A  $\mathbf{C} (i \times p)$  mátrix sorvektorok formájában tartalmazza az egyes mérési pontok koncentrációprofilját az összes  $(p)$  komponensre. Az  $\mathbf{E}$  mátrix a reziduális zajt reprezentálja. Az egyenletnek egy igen szemléletes illusztrációja megtalálható Gendrin és társai összefoglalójában.<sup>9</sup>

Amennyiben ismerjük a referencia-spektrumokat, a mérési pontok koncentrációprofiljai igen egyszerűen számíthatók az egyenlet átrendezésével (a pszeudoinvert mátrix elemeinek kiszámításával).

$$\mathbf{C} = \mathbf{X}\mathbf{S}(\mathbf{S}^T\mathbf{S})^{-1} \quad (2)$$

Természetesen az alkotó komponensek spektruma is meghatározható, amennyiben az egyes mérési pontok koncentrációprofilját ismerjük.

$$\mathbf{S}^T = (\mathbf{C}^T\mathbf{C})^{-1}\mathbf{C}^T\mathbf{X} \quad (3)$$

Ez a legtriviálisabb és fizikailag legkorrektebb módszer a komponensek térbeli koncentráció-eloszlásának meghatározására, és a *klasszikus legkisebb négyzetek módszere* (*classical least squares*, CLS) nevet viseli. Számos publikáció található, melyben ily módon nyertek információt a komponensek eloszlásáról.<sup>24-27</sup>

Ugyan fizikailag korrekt az előzőekben bemutatott okfejtés, matematikailag (és gyakorlatilag) mégsem tekinthető a legjobbnak. Amennyiben a komponensek spektrumai nagyon hasonlítanak egymáshoz (pl. egy hatóanyag különböző polimorf módosulatai), és/vagy rossz a jel/zaj viszony, a kapott megoldás esetenként igen pontatlanná válhat. Ilyen esetekben célszerű használni az ún. *parciális legkisebb négyzetek módszerét* (*partial least squares*, PLS), melynek a matematikai alapját az eredeti változók lineáris kombinációinak, ún. *látens változóknak* regressziója képezi<sup>10</sup> (ld. még a faktoranalízist a 4.3.2.2. fejezetben). Ekkor egy ún. tanítóhalmazon (training set) felállítjuk a lineárisnak tekintett összefüggést a spektrumok és a koncentrációk

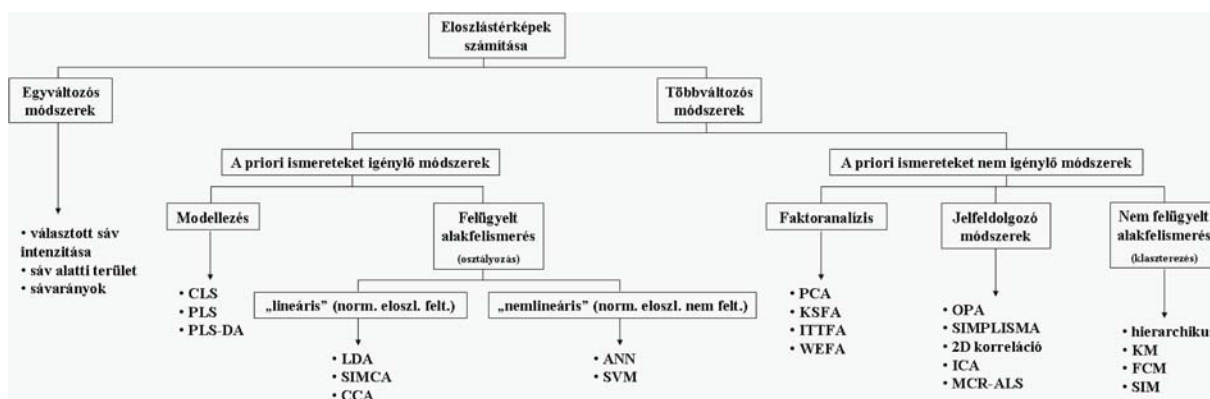
között, majd validálás (test/validation set) után ezt a modellt alkalmazzuk a kémiai térképet alkotó ismeretlen összetételű spektrumokra. A módszer igen elterjedt az analitikai gyakorlatban, ennek megfelelően a kémiai térképezésben is megszokottá vált az alkalmazása.<sup>28-31</sup>

Egy PLS módszeren alapuló technika (*partial least squares - discriminant analysis*, PLS-DA) a tiszta komponensek referencia-spektrumait tekinti tanítóhalmaznak, és ezek alapján becsüli a keverékspektrumok koncentráció-profilját.<sup>32</sup> Érdekes fejlemény, hogy bár mindenhol külön módszerként tárgyalják ezt az algoritmust, 2009-ben bebizonyították róla, hogy a kémiai térképezés esetében matematikailag – így persze eredményét tekintve is – teljesen ekvivalens a közönséges legkisebb négyzetek módszerével.<sup>33</sup>

A spektrumok osztályozhatók is, ha definiálunk megfelelő spektrum-osztályokat, melyekkel összevethetők a mért spektrumok. Az osztályozási (és az ún. klaszterezési, ld. 4.3.2.2. fejezet) módszerek vizuálisan jól interpretálhatók. Minden spektrumot egy vektornak tekintünk a  $k$  dimenziós

térben (hiszen egy spektrum  $k$  számértéket tartalmaz). Az osztályozási módszereknél az *a priori* információt arra használjuk, hogy régiókat különítsünk el ebben a térben. Ezután az ismeretlen mintáról felvett spektrumokat az alapján értékeljük, hogy melyik előre definiált régióba esnek a  $k$  dimenziós térben.

A kémiai térképezésben alkalmazott tradicionális osztályozási módszerek az LDA (*linear discriminant analysis*) és a SIMCA (*soft independent modeling of class analogies*). Az LDA esetén a spektrális tér egyes régióit elkülönítő „falak” helye oda kerül, ahol az osztályok közötti és az osztályokon belüli variancia aránya a legnagyobb (azaz a különböző osztályokba tartozó spektrumok legjobban elválnak egymástól). A módszert leginkább biológiai szövetek IR térképeire alkalmazták.<sup>34,35</sup> A SIMCA módszernél a régiók elkülönítése nem az eredeti spektrális térben, hanem annak egy transzformált alterében (az ún. főkomponens-térben, ld. még 4.3.2.2. fejezet) történik.<sup>36</sup> Ezt a módszert agydagatok metasztázisainak IR térképezés vizsgálatában használták.<sup>37</sup>



6. Ábra. Kémiai térképek feldolgozására használt kemometriai módszerek. A lineárisnak nevezett módszerek az objektumok normális eloszlását feltételezik.

Az osztályozást aszerint is elvégezhetjük, hogy a térkép-spektrumokat korreláltatjuk a referencia-spektrumokhoz egy egyszerű korrelációs együttható számításával. Amelyik komponens spektrumával leginkább korrelál az adott pontban mért spektrum, ahhoz a komponenshez rendeljük a vizsgált felület adott pontját. Ez a *cosine correlation analysis* (CCA), melyet gyógyszer-tabletták komponens-eloszlásának felderítésére is lehet alkalmazni.<sup>38</sup>

Az utóbbi évtizedben egyre komolyabb jelentőséget kapnak a nemlineáris osztályozási módszerek, melyeknek ugyan nehéz működés közben nyomonkövetni a lépéseit, de olyankor is kiváló osztályozást végeznek, amikor a többi (lineáris) módszer csődöt mond. *Mesterséges ideghálózatokat* (*artificial neural network*, ANN) alkalmaztak műanyagok NIR térképezésére<sup>39</sup> és szövetek FTIR mikroszkópiás<sup>40</sup> vizsgálataiban. Más szerzők tápszerek<sup>41</sup> illetve faanyagok<sup>42</sup> osztályozást végeztek el kémiai térképezéssel és ún. SVM (*support vector machine*) osztályozással.

#### 4.3.2.2. *A priori* ismereteket nem igénylő módszerek

Az előző módszerek mindegyike akkor működőképes, ha ismerjük a mintát alkotó komponensek referencia-spektrumait. Ez a gyakorlatban igen gyakran teljesül (pl.

egy fejlesztés alatt álló gyógyszer esetében), azonban bizonyos esetekben ez egyáltalán nincs így (pl. ismeretlen szennyezők eloszlásának meghatározása ill. hamisítványok vizsgálata). Ekkor lehet szükség nem felügyelt módszerek alkalmazására, melyek pusztán a felvett térkép által szolgáltatott adatokból dolgoznak.

A kemometriai adatelemzés egyik általánosan használt módszere a *faktoranalízis*. E módszerekkel kiszűrhetők az egymással erős korrelációban álló változók, és az adathalmaz (pl. kémiai térképek) fő információhordozói különválaszthatók a nem informatív tényezőktől (pl. spektrális zaj, polarizációs effektusok).

A faktoranalízis egyik leggyakrabban alkalmazott típusa a *főkomponens-analízis* (*principal component analysis*, PCA).<sup>10</sup> Az adatokat az eredeti változótér helyett egy olyan ortogonális alterre vetítjük, ahol az új változók a lehető legjobban leírják az adathalmazban jelen lévő varianciát. Másképp megfogalmazva, az első néhány új változó tartalmazza az adathalmazban jelen lévő információ túlnyomó részét, és ezeket elválasztjuk a többi változótól, melyek főként zajt és egyéb zavaró effektusokat tartalmaznak. Ezt a módszert igen kiterjedten alkalmazzák komponensek térbeli koncentráció-eloszlásának meghatározására.

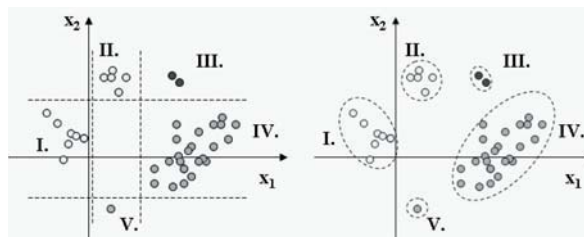
zására NIR<sup>14</sup> és Raman<sup>43,44</sup> térképekből. A módszer hátránya, hogy bár az új változók jól leírják a varianciát, fizikailag nem értelmezhetők (nem tekinthetők spektrumoknak, és nem is rendelkeznek ugyanazokkal a tulajdonságokkal). Léteznek egyéb faktoranalízises módszerek is, melyek hasonló működési elven az alkotó komponensek spektrumainak becslését segítik elő (*key-set factor analysis* – KSFA,<sup>45</sup> *iterative target transformation factor analysis* – ITTFA,<sup>46</sup> *window-evolving factor analysis* – WEFA<sup>47,48</sup>). A BTEM (*band-target entropy minimization*) módszer szintén faktoranalízis alapú, a főkomponensek transzformációjával (forgatásával) és egy entrópia jellegűnek definiált érték minimalizálásával egyenként határozza meg az alkotóelemek spektrumait. (A módszer alapja az a tény, hogy a tiszta anyagok entrópiája mindig kisebb, mint a keverékéké.) Igen kis koncentrációban jelen lévő komponens Raman spektruma is meglepően jól becsülhetőnek bizonyult egy tablettáról felvett Raman-térkép BTEM-mel történő feldolgozása során.<sup>49</sup>

A 4.3.2.1. fejezet elején ismertetett CLS módszer szerinti modellezés akkor is lehetséges, ha nem rendelkezünk referencia-spektrumokkal és egyéb *a priori* információval a vizsgált rendszerről. A tiszta komponensek spektrumai a keverékek alapján számos módszerrel becsülhetők. A modellezési metodikák a nyers  $X$  mátrixszal dolgoznak transzformációk alkalmazása nélkül. A SIMPLISMA (*simple-to-use interactive self-modeling mixture analysis*) algoritmus az  $X$  mátrix oszlopainak statisztikai elemzésével megbecsüli, hogy melyek a legkevésbé átlapoló sávok (mely sávok tartoznak valószínűleg csak egyetlen komponenshez), majd e sávok intenzitásértékeit helyettesíti be a  $C$  mátrixba.<sup>50</sup> Ezután a CLS megfelelő egyenletével ((3) egyenlet) becsüli a tiszta komponensek spektrumait. Az OPA (*orthogonal projection approach*) algoritmus kiválasztja az egymástól leginkább különböző spektrumokat.<sup>51</sup> Egy harmadik módszer (sample-sample 2D correlation spectroscopy) az  $X$  mátrix kovariancia-mátrixát hozza létre és elemzi: segítségével kiválaszthatók a térkép spektrumai közül a „legtisztábbak”, melyek leginkább csak egy komponens sávjait tartalmazzák.<sup>52</sup> Ezt a megközelítést több esetben is alkalmazták gyógyszerekről készített kémiai térképek vizsgálatára.<sup>53,54</sup> Az ICA módszer (*independent component analysis*, ICA) szintén a korrelációt figyeli, és az egymással legkevésbé korreláló spektrumokat választja ki.<sup>55</sup>

Az *a priori* ismereteket nem igénylő módszerek egy további típusát alkotják az iterációs algoritmusok. Közös tulajdonságuk, hogy bemenő adatként szükségük van egy (akár igen rosszul becsült) kiindulási spektrum- vagy koncentráció-profilra, és ezt különböző fizikai korlátozások érvényesítésével finomítják addig, míg elegendően jó minőségű „tiszta” spektrumokat, illetve elegendően pontos koncentráció-profilt nem kapunk. Korlátozások lehetnek például, hogy a spektrumok intenzitásértékei és a komponensek koncentrációi nem lehetnek negatívak, és/vagy a számolt koncentrációk összegének konstansnak kell lenniük. A legjobban működő algoritmusok az MCR-ALS<sup>56</sup> (*multivariate curve resolution – alternate least squares*), az NMF<sup>57</sup> (*non-negative matrix factorization*), és ennek egy módosított válfaja, a PMF<sup>58</sup> (*positive matrix factorization*). Általánosságban elmondható, hogy a fenti-ekben ismertetett nem felügyelt módszerek (SIMPLISMA, OPA stb.) által

durván becsült spektrumok ezen iteratív módszerekkel történő finomításával meglepően jó eredményt lehet elérni. Ezt jól bizonyítják műanyagok,<sup>59</sup> kőzetek,<sup>60</sup> biológiai minták<sup>60</sup> és gyógyszerek<sup>61</sup> térképezéses vizsgálatairól készült közlemények is. A különböző iterációs módszerek összehasonlítása is megtalálható mind NIR,<sup>62</sup> mind Raman<sup>63</sup> térképezés esetében.

A 4.3.2.1. fejezetben ismertetett osztályozási módszerek nem felügyelt rokonai az ún. *klaszterezési* eljárások. Itt nem régiókat különítünk el a spektrális térben egy tanítóhalmaz segítségével, hanem a hasonló tulajdonságú, azaz a spektrális térben csoportokat alkotó spektrumokat keressük meg.



7. Ábra. LDA osztályozási (felügyelt) és klaszterezési (nem felügyelt) eljárások működésének értelmezése és illusztrációja két dimenzióban.

A klaszterezési módszerek az objektumok (ebben az esetben spektrumok) közötti távolságok alapján végzik a csoportosítást. A legrégebbi technika a *hierarchikus klaszterezés*, mellyel mi állíthatjuk be, hogy az objektumokat mekkora távolság esetén tekintjük különböző klaszterbe tartozónak. A másik legelterjedtebben alkalmazott metódus a *K-közép* (*K-means*, KM) módszer, mely szintén minden objektumot egy-egy klaszterbe sorol be, de a klaszterek számát előre meg kell adni. Ez utóbbi rokona a *fuzzy-C-means* (FCM) algoritmus, mely egy objektumot nem csak egy csoportba sorol be, hanem meghatározza, hogy melyik csoportba milyen mértékben sorolható be az adott objektum. Mindhárom módszert összehasonlították egy biológiai szövetek kémiai térképeinek elemzésével foglalkozó tanulmányban.<sup>64</sup> Bár ez esetben a hierarchikus klaszterezés bizonyult a leghatékonyabbnak, a legtöbb közleményben mégis FCM módszerrel<sup>65-67</sup> vagy annak továbbfejlesztett változataival<sup>68</sup> dolgozták fel a kémiai térképeket.

A klaszterezési módszerek egy része kifejezetten a kémiai térképek feldolgozására lett továbbfejlesztve. A SIM (spectral identity mapping) módszer a felügyelt módszereknél tárgyalt CCA rokona,<sup>69</sup> míg más módszerek a spektrális információ mellett a térbeli információt is felhasználják.<sup>70,71</sup> Ez utóbbi megközelítés lehetővé tette hatóanyagok különböző, de hasonló rezgési spektrummal rendelkező polimorf módosulatainak pontos elkülönítését egy tablettáról készített Raman-térképen.<sup>71</sup>

## 5. Kémiai térképezés gyakorlati alkalmazásai

A kémiai térképezés útján elérhető nagy információmennyiségnek köszönhetően a módszer egyre nagyobb jelentőségre tesz szert számos tudományos és ipari területen, melyek közül hazánkban az egyik legjelentősebb ágazat a gyógyszeripar. Jelen közleményünkben ennek megfelelően gyógyszertabletták esetében mutatjuk be a Raman-térképe-

zés és az adatelemző módszerek alkalmazhatóságát. (A gyógyszeripari alkalmazások igen részletes összefoglalója magyar nyelven is megtalálható a szakirodalomban<sup>61</sup>).

### 5.1. Vizsgálati módszer ismertetése

A vizsgálatokat a HORIBA Jobin Yvon cég által gyártott LabRam típusú Raman-mikroszkóppal végeztük. A 785 nm hullámhosszú gerjesztő fényt egy félvezető diódalézer sugárforrás biztosította. A berendezésben a lézerguság egy megfelelő optikai tükör- és szűrőrendszeren keresztül egy (mikrométeres felbontást biztosító) Olympus BXC-40 típusú optikai mikroszkópba jut. A sugár a mikroszkóp cserélhető objektívain keresztül kerül a mintára. Vizsgálataink alkalmával 10x-os nagyítást alkalmaztunk, mellyel a lézer fókuszpontjának átmérője kb. 3,0  $\mu\text{m}$ . A visszazórt fényt ugyanez az objektív gyűjti össze, majd egy tükörrendszeren, köztük egy, a Rayleigh-szórást kiszűrő féligáteresztő tükrön keresztül egy optikai rácstra (1800 vonás/mm) bocsátja. A hullámhossz szerint felbontott fény egy CCD detektorra kerül.

### 5.2. Gyógyszertabletták analízise

Jelen tanulmányban két tablettá példáján bemutatjuk, hogy a kémiai térképezés és a korszerű adatelemzési (kemometriai) módszerek ötvözésével előzetes információ nélkül is felderíthető a minták belső szerkezete. Két gyógyszertárban kapható, Na-ibandronát hatóanyagot tartalmazó készítményt (*Bondronat*<sup>®</sup> és *Bonviva*<sup>®</sup>, F. Hoffmann-La Roche Ltd., Svájc) hasonlítottunk össze úgy, hogy az elemzéshez semmilyen a priori információt nem használtunk fel. Az ellenőrzéshez felhasználtuk a tablettá kvalitatív összetételét, mely a betegájékoztatóban megtalálható. A tabletták korpuszának (magjának) összetevői: Na-ibandronát, laktózmohidrá, mikrokristályos cellulóz, polivinil-pirrolidon (PVP), keresztkötött PVP, sztearinsav, koloid  $\text{SiO}_2$ . Célunk annak felderítése volt, hogy az elvileg azonos összetételű, de mégis eltérő gyártmányú tabletták megkülönböztethetők-e egymástól. Az eredménynek a gyógyszer-hamisítványok vizsgálatában kiemelt jelentősége van, különösen a magas színvonalon végzett hamisítás esetében.

#### 5.2.1. Vizsgálati körülmények és kiértékelés módja

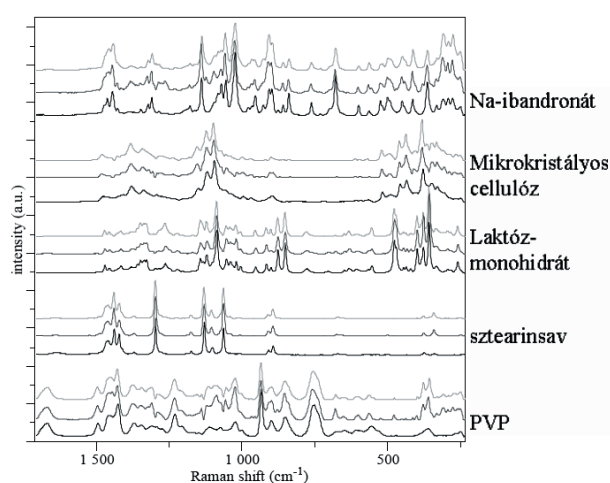
A vizsgálandó tablettákat félbevágtuk és vágott felülettel fölfelé a tárgyasztalra helyeztük. A vágott felületről 400  $\mu\text{m}$   $\times$  400  $\mu\text{m}$ -es területen Raman-térképet készítettünk 10  $\mu\text{m}$ -es lépésközzel (térbeli felbontással). A spektrumok felvételét 200-2000  $\text{cm}^{-1}$  tartományban végeztük 28 s felvételi idővel és pontonként 2 spektrum átlagolásával.

A térkép spektrumait először alapvonal-korrekciónal és normálással korrigáltuk. A mintát teljesen ismeretlennek tekintettük, és nem használtuk fel a betegájékoztatóban található, az alkotó komponensekre vonatkozó információt. A spektrumok statisztikus becslésére a 4.3.2.1. alfejezetben említett SIMPLISMA módszert alkalmaztunk, majd a becslés jóságát MCR-ALS iterációval (lásd ugyanott) finomítottuk. A számításokat MatLAB 7.6 szoftverrel (MathWorks, USA) és PLS\_Toolbox 5.5 kiegészítővel (Eigenvector Research, USA) végeztük. A térkép spektrumait ezután a klasszikus legkisebb négyzetek módszerével (CLS) modelleztük az így becsült „tisztá” spektrumok felhasználásával, ezáltal a

térkép minden pontjában meghatároztuk a komponensek koncentrációit. A koncentrációkat a hely függvényében vizuálisan ábrázoltuk a Labspec 4.02 szoftver segítségével. A fentiek után ellenőrzésképpen a kiindulási komponensek referencia-spektrumainak felhasználásával is elvégeztük a vizsgálatokat CLS módszert alkalmazva.

#### 5.2.2. Eredmények

A 8. ábrán látható, hogy mindkét tablettá esetében öt komponens azonosítható az elvileg hét komponensből. Ez az eredmény nem meglepő, hiszen a keresztkötött és a nem keresztkötött polivinil-pirrolidon (PVP) Raman-spektruma a legapróbb részletekig megegyezik, illetve a koloid szilícium-dioxid gyakorlatilag nem rendelkezik mérhető Raman-spektrummal. A becsült spektrumok elhanyagolható mértékben különböznek a tiszta komponensekről felvett Raman-spektrumoktól.



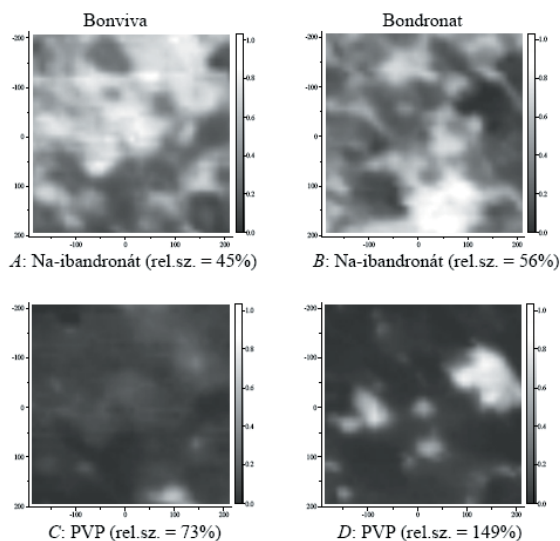
8. Ábra. Bonviva (világosszürke) és Bondronat (sötétszürke) tabletták alkotó komponenseinek SIMPLISMA/MCR-ALS módszerrel becsült Raman-spektrumai összehasonlítva a referencia-spektrumokkal (fekete).

A becsült spektrumok segítségével ábrázolható a komponensek koncentráció-eloszlása. Négy komponens eloszlásának jellege nagyon hasonló a két tablettában (pl. a hatóanyagé, ld. 9. ábra), ellenben a PVP eloszlásában jelentős különbség tapasztalható: a PVP elmosódó, homogén elrendeződést mutat a Bonvivában, míg jól elkülönülő önálló szemcséket alkot a Bondronatban.

A kép alapján szemmel látható különbséget számszerűen is alátámasztja a PVP egyes pontokban számolt koncentrációjának szórása (a Bonviva esetén 73%, míg a Bondronatnál 149%). Az eredmény valószínűleg azzal magyarázható, hogy a PVP-t oldat formában vitték be a Bonviva készítménybe egy granulálási művelet során, míg a Bondronat esetében a PVP-t szilárd formában adták a többi komponenshez. (A hatóanyag-koncentrációk szórásában is megfigyelhető némi különbség, de ez nem szignifikáns.)

E tanulmány szempontjából leglényegesebb szempont, hogy a priori ismeretek nélkül is megkülönböztethető két hasonló készítmény, melyek csak egy technológiai lépésben különböztek egymástól. Így valószínűleg olyan hamisítványok azonosítására is mód nyílik, melyek kémiai összetételükben nem, csak gyártási technológiájukban (így

az alkotók térbeli eloszlásában) térnek el a valódi tablettáktól. Másik lehetséges alkalmazás annak megállapítása, hogy két illegális készítmény (hamisítvány vagy kábítószer) azonos kiindulási anyagokból és azonos technológiával (azaz ugyanabban a laborban) készült-e.



9. Ábra. Hatóanyag (A,B) és PVP (C,D) koncentráció-eloszlása a Bonviva (A,C) és Bondronat (B,D) tablettákban. Zárójelben az adott komponens egyes pontokban számított koncentrációinak relatív szórása látható.

## 6. Összefoglalás

Közleményünkben összefoglaltuk a kémiai térképezés felvételi módszereit és az eddig alkalmazásra került matematikai adatelemzési módszereket. Számos példa bizonyítja a kémiai térképezés széleskörű alkalmazhatóságát a legkülönbözőbb területeken. Saját kísérleteink azt bizonyítják, hogy Raman-térképezés és *a priori* információt nem igénylő modellezési módszerek (SIMPLISMA és MCR-ALS) ötvöztetésével apró különbségek is kimutathatók hasonló készítmények közt. Ez előrevetíti a módszer alkalmazhatóságát a gyógyszer-iparban és a bűnügyi vizsgálatok területén.

## Köszönetnyilvánítás

A munkát a K76 346 számú OTKA projekt támogatta.

## Hivatkozások

- Barer, R.; Cole, A.R.H.; Thompson, H.W. *Nature* **1949**, *163*, 198–201.
- Harthcock, M.A.; Atkin, S.C. *Appl. Spectrosc.* **1988**, *42*, 449–455.
- Yang, X.M.; Tryk, D.A.; Hasimoto, K.; Fujishima, A. *Appl. Phys. Lett.* **1996**, *69*, 4020–4022.
- Ferris, D.; Lawhead, R.; Dickman, E.; Holtzappple, N.; Miller, J.; Grogan, S.; Bambot, S.; Agrawal, A.; Faupel, M. *J. Low. Genital Tract. Dis.* **2001**, *5*, 65–72.
- Tatzer, P.; Wolf, M.; Panner, T. *Real-Time Imaging* **2005**, *11*, 99–107.
- Robert, P.; Bertrand, D.; Devaux, M.F.; Sire, A. *Anal. Chem.* **1992**, *64*, 664–667.
- Markwort, L.; Kip, B.; Da Silva, E.; Roussel, B. *Appl. Spectrosc.* **1995**, *49*, 1411–1430.
- Gowen, A.A.; O'Donnell, C.P.; Cullen, P.J.; Bell, S.E.J. *Eur. J. Pharm. Biopharm.* **2008**, *69*, 10–22.
- Gendrin, C.; Roggo, Y.; Collet, C. *J. Pharm. Biomed. Anal.* **2008**, *48*, 533–553.
- Mark, H.; Workman, J. *Chemometrics in Spectroscopy*, Academic Press, San Diego, **2007**.
- Robinette, S.L.; Veselkov, K.A.; Bohus, E.; Coen, M.; Keun, H.C.; Ebbels, T.M.D.; Beckonert, O.; Holmes, E.C.; Lindon, J.C.; Nicholson, J.K. *Anal. Chem.* **2009**, *81*, 6581–6589.
- Pappas, D.; Smith, B.W.; Winefordner, J.D. *Appl. Spectrosc. Rev.* **2000**, *35*, 1–23.
- Dubois, J.; Wolff, J.-C.; Warrack, J.K.; Schoppelrei, J.; Lewis, E.N. *Spectroscopy* **2007**, *22*, 40–50.
- Clarke, F. *Vib. Spectrosc.* **2004**, *34*, 25–35.
- Reich, G. *Pharmind* **2002**, *64*, 870–874.
- Chan, K.L.A.; Kazarian, S.G. *Appl. Spectrosc.* **2003**, *57*, 381–389.
- Chan, K.L.A.; Hammond, S.V.; Kazarian, S.G. *Anal. Chem.* **2003**, *75*, 2140–2146.
- Šašić, S.; Clark, D.A. *Appl. Spectrosc.* **2006**, *60*, 494–502.
- Lewis, E.N.; Schoppelrei, J.; Lee, E. *Spectrosc. Mag.* **2004**, *04*, 26–36.
- Barnes, R.J.; Dhanoa, M.S.; Lister, S.J. *Appl. Spectrosc.* **1989**, *43*, 772–777.
- Geladi, P.; MacDougall, D.; Martens, H. *Appl. Spectrosc.* **1985**, *39*, 491–500.
- Vogt, F.; Cramer, J.; Booksh, K. *J. Chemometr.* **2005**, *19*, 510–520.
- Geladi, P.; Grahn, H. *Multivariate Image Analysis*, John Wiley & Sons, Chichester, UK, **1996**.
- Treado, P.J.; Nelson, M.P. in *Handbook of Vibrational Spectroscopy*; Chalmers, J.M., Griffiths, P.R., Eds.; John Wiley & Sons, London, **2002**; Vol. 2, pp. 1429–1459.
- Šašić, S.; Clark, D.A.; Mitchell, J.C.; Snowden, M.J. *Analyst* **2004**, *129*, 1001–1007.
- Vajna, B.; Farkas, I.; Szabó, A.; Zsolt, Zs.; Marosi, Gy. *J. Pharm. Biomed. Anal.* **2010**, *51*, 30–38.
- Clark, D.; Henson, M.; LaPlant, F.; Šašić, S.; Zhang, L. in *Applications of Vibrational Spectroscopy in Pharmaceutical Research and Development*; Pivonka, D.E., Chalmers, J.M., Griffiths, P.R., Eds.; John Wiley & Sons, London, **2007**, pp. 309–335.
- Romero-Torres, S.; Pérez-Ramos, J.D.; Morris, K.R.; Granta, E.R. *J. Pharm. Biomed. Anal.* **2005**, *38*, 270–274.
- Gendrin, C.; Roggo, Y.; Collet, C. *Talanta* **2007**, *73*, 733–741.
- Burger, J.; Geladi, P. *Analyst* **2006**, *131*, 1152–1160.
- Burger, J.; Geladi, P. *J. Chemometr.* **2006**, *20*, 106–119.
- Jovanovic, N.; Gerich, A.; Bouchard, A.; Jiskoot, W. *Pharm. Res.* **2006**, *23*, 2002–2013.
- Amigo, J.M.; Ravn, C.; Gallagher, N.B.; Bro, R. *Int. J. Pharm.* **2009**, *373*, 179–182.
- Beleites, C.; Steiner, G.; Sowa, M.G.; Baumgartner, R.; Sobottka, S.; Schackert, G.; Salzer, R. *Vib. Spectrosc.* **2005**, *38*, 143–149.
- Mansfield, J.R.; McIntosh, L.M.; Crowson, A.N.; Mantsch, H.H.; Jackson, M. *Appl. Spectrosc.* **1999**, *53*, 1323–1330.
- Massart, D.L.; Vandeginste, B.G.M.; Deming, S.N.; Michotte, Y.; Kaufman, L. *Chemometrics: A Textbook*, Elsevier Science, Amsterdam, **1988**.
- Krafft, C.; Shapoval, L.; Sobottka, S.B.; Schackert, G.; Salzer, R. *Biochim. Biophys. Acta (BBA)-Biomembr.* **2006**, *1758*, 883–891.
- Zugates, C.T.; Treado, P.J. *Int. J. Vib. Spectrosc.* **1999**, Volume 2, Edition 4 39. van den Broek, W.H.A.M.; Wienke, D.; Melssen, W.J.; Buydens, L.M.C. *Anal. Chim. Acta* **1998**, *361*, 161–176.
- Lasch, P.; Diem, M.; Hansch, W.; Naumann, D. *J. Chemometr.* **2006**, *20*, 209–220.
- Fernandez Pierna, J.A.; Baeten, V.; Dardenne, P.; *Chemometr. Intell. Lab. Syst.* **2006**, *84*, 114–118.
- Liu, J.J.; Bharati, M.H.; Dunn, K.G.; MacGregor, J.F. *Chemometr. Intell. Lab. Syst.* **2005**, *79*, 42–54.

43. Šašić, S. *Appl. Spectrosc.* **2007**, *61*, 239–250.
44. Šašić, S.; Clark, D.A. *Appl. Spectrosc.* **2006**, *60*, 494–502.
45. Malinowski, E.R. *Anal. Chim. Acta* **1982**, *134*, 129–137.
46. Vandeginste, B.G.M.; Derks, W.; Kateman, G. *Anal. Chim. Acta* **1985**, *173*, 253–264.
47. De Juan, A.; Maeder, M.; Hancewicz, T.; Tauler, R. *Chemometr. Intell. Lab. Syst.* **2005**, *77*, 64–74.
48. De Juan, A.; Tauler, R. *Trends Anal. Chem.-TrAC* **2004**, *23*, 70–79.
49. Widjaja, E.; Kim, R.; Seah, H. *J. Pharm. Biomed. Anal.* **2008**, *46*, 274–281.
50. Windig, W.; Guilment, J. *Anal. Chem.* **1991**, *63*, 1425–1432.
51. Sanchez, F.C.; Toft, J.; Van den Bogaert, B.; Massart, D.L. *Anal. Chem.* **1996**, *68*, 79–85.
52. Šašić, S.; Morimoto, M.; Otsuka, M.; Ozaki, Y. *Vib. Spectrosc.* **2005**, *37*, 217–224.
53. Šašić, S.; Clark, D.A.; Mitchell, J.C.; Snowden, M.J. *Appl. Spectrosc.* **2005**, *59*, 630–638.
54. Shinzawa, H.; Awa, K.; Okumura, T.; Morita, S.; Otsuka, M.; Ozaki, Y.; Sato, H. *Vibr. Spectr.* **2009**, *51*, 125–131.
55. Hyvarinen, A.; Karhunen, J.; Oja, E. *Independent Component Analysis*, first edition, John Wiley & Sons, Chichester, **2001**.
56. Tauler, R. *Chemometr. Intell. Lab. Syst.* **1995**, *30*, 133–146.
57. Lee, D.D.; Seung, H.S. *Nature* **1999**, *401*, 788–791.
58. Paatero, P.; *Chemometr. Intell. Lab. Syst.* **1997**, *37*, 23–35.
59. Duponchel, L.; Elmi-Rayaleh, W.; Ruckebusch, C.; Huvenne, J.P. *J. Chem. Inform. Comp. Sci.* **2003**, *43*, 2057–2067.
60. Sajda, P.; Du, S.; Parra, L. *SPIE Wavelets X*, San Diego, **2003**.
61. Vajna, B.; Nagy, Zs.; Patyi, G.; Zsolt, Zs.; Antal, I.; Marosi, Gy. *Acta Pharmaceutica Hungarica* **2009**, *79*, 104–116.
62. Gendrin, C.; Roggo, Y.; Collet, C. *J. Near Infrared Spectrosc.* **2008**, *16*, 151–157.
63. Andrew, J.J.; Hancewicz, T.M. *Appl. Spectrosc.* **1998**, *52*, 797–807.
64. Lasch, P.; Haensch, W.; Naumann, D.; Diem, M. *Biochim. Biophys. Acta (BBA)-Mol. Basis Dis.* **2004**, *1688*, 176–186.
65. Mansfield, J.R.; Sowa, M.G.; Scarth, G.B.; Somorjai, R.L.; Mantsch, H.H. *Anal. Chem.* **1997**, *69*, 3370–3374.
66. Mansfield, J.R.; Sowa, M.G.; Payette, J.R.; Abdulrauf, B.; Stranc, M.F.; Manisch, H.H. *IEEE Trans. Med. Imag.* **1998**, *17*, 1011–1018.
67. Wang, Y.-P.; Wang, Y.; Spencer, P. *Appl. Spectrosc.* **2006**, *60*, 826–832.
68. Noordam, J.C.; van den Broek, W.H.A.M.; Buydens, L.M.C. *Chemometr. Intell. Lab. Syst.* **2002**, *64*, 65–78.
69. Turner, J.F.; Zhang, J.; O'Connor, A. *Appl. Spectrosc.* **2004**, *58*, 1308–1317.
70. Krooshof, P.W.T.; Postma, G.J.; Melssen, W.J.; Buydens, L.M.C. Tran, T.N. *Trends Anal. Chem.* **2006**, *25*, 1067–1080.
71. Lin, W.-Q.; Jiang, J.-H.; Yang, H.-F.; Ozaki, Y.; Shen, G.-L.; Yu, R.-Q. *Anal. Chem.* **2006**, *78*, 6003–6011.

### Methods of vibrational chemical imaging and their evaluation with chemometric methods

Chemical imaging is a novel analytical tool based primarily on vibrational spectrometric methods. The aim of this study is to review the acquisition and evaluation methods of chemical images and to illustrate the applicability in counterfeit drug analysis.

A chemical image has a three-dimensional data structure with two spatial and one spectral dimension (Fig. 1). One way to obtain imaging data is *point* or *line mapping*, where a motorized stage is moving the sample step-by step and a spectrum is taken from an area of 1–10  $\mu\text{m}^2$  after each step, until the whole surface is scanned (see Fig.2). Another way is to perform *global imaging*, where a large surface is illuminated, and the reflected light is collected through an adjustable wavelength filter, acquiring one image at a certain wavelength at a time. Multiple images at several wavelengths form the 3D *hyperspectral image*.

A chemical imaging experiment requires a spectrometer with an appropriate excitation source, imaging optics and a computer. A microscope is needed to perform point/line mapping. Advanced objectives are required for MIR/NIR mapping (diffuse reflection or attenuated total reflection, Fig. 3). The excitation source is diluted in the case of global imaging, and the reflected light is directly collected on a couple-charged device (CCD) or focal plane array (FPA) detector.

The main aim during sample characterization is to determine the spatial distribution of certain or all of the components. Univariate modeling (based on peak intensities, areas or ratios) is possible, but usually leads to inaccurate results due to the overlapping of bands of different components. Alternatively, several multivariate approaches can be found in the literature (Fig. 6). Since most of them were developed to evaluate two-dimensional data, the data cube has to be unfolded into a two-dimensional matrix by unfolding one spatial dimension (Fig. 5). Supervised methods can be applied if there is *a priori* information about the sample (e.g. the reference spectra of the pure components are available). Classical least squares (CLS) modeling is a direct method to estimate component concentrations with Equation (2) or pure spectra with Eq. (3). Partial least squares (PLS) is a factor-analysis

based method providing more stable prediction even if the pure component spectra are highly overlapped. The spectra can be grouped with supervised classification methods.

Unsupervised techniques are used when information about the analyte is sparse. The distribution of the components can be directly estimated with factor analysis methods such as principal component analysis (PCA), or the pure component spectra can be estimated with self-modeling curve resolution methods (SIMPLISMA, sample-sample 2D correlation analysis, MCR-ALS). Grouping of spectra can be carried out in an unsupervised way with several clustering methods (Fig. 7).

In the present study the authors tested the applicability of Raman mapping and unsupervised chemometric methods in comparing two pharmaceutical tablets (with the same ingredients) without *a priori* information. The aim was to model the case of high quality counterfeits in which the ingredients are the same, but the manufacturing technology is different from the original (and its administration might lead to unforeseen consequences).

Raman mapping spectra were collected using a Horiba Jobin-Yvon LabRAM system coupled with an external 785 nm diode laser source and an Olympus BX-40 optical microscope. An objective of 10x magnification was used for optical imaging and spectrum acquisition. The spectral range was 200–2000  $\text{cm}^{-1}$ .

Two different products from the same manufacturer (Bonviva<sup>®</sup> and Bondronat<sup>®</sup>, F. Hoffmann-La Roche, Switzerland) were compared. Their composition is shown in Table 1. The tablets were milled and cores were scanned on an area of 400 mm x 400 m with 10 mm spatial resolution. Raman maps were evaluated with MCR-ALS method (initial guesses were obtained with SIMPLISMA algorithm).

Fig. 8 shows that the spectrum estimation with MCR-ALS is very accurate, thus, so is the estimated distribution of spectral “concentration” (compared to CLS). As Fig. 9 shows, the two products can be distinguished from each other based on the distribution of PVP. In conclusion, it was shown that Raman mapping and chemometric algorithms provide a good combination for the analysis of unknown pharmaceutical products.