

Szerkezetkutatás NMR-spektroszkópiával a gyógyszeriparban: hatékonyságnövelés nem egyenletes mintavételezés segítségével

SZIGETVÁRI Áron és Ifj. SZÁNTAY Csaba*

Richter Gedeon Nyrt, Szerkezetkutatási osztály, Gyömrői út 19-21, 1103 Budapest, Magyarország

1. Bevezetés

A szerkezeti képletek ugyanolyan módon és mértékben hatják át a kémiai szakirodalmat, és annak ugyanolyan természetes közlési eszközei, mint amilyen magától értetődő természetességgel alkalmazzuk és értelmezzük a szavakat a mindennapi nyelvhasználatban. Ha valaki a mindennapokban egy tárgyról (például egy számítógépről) beszél, akkor többnyire nem kételkedünk abban, hogy az a tárgy valóban az, aminek állítja (számítógép, nem pedig kávéfőző). Hasonlóan, ha egy kémiai szakkönyvben vagy egy folyóiratcikkben találkozunk szerkezeti képletekkel, többnyire az köti le a figyelmünket, hogy a mű szerzői mit mondanak az adott vegyületről, annak tulajdonságairól, kémiai reakcióiról – azon viszont ritkán merengünk el, vajon magát a szerkezetet a szerzők honnan ismerik vagy hogy mennyire lehetünk biztosak annak helyességében. Azonban egy vegyület kémiai szerkezetét nem lehet úgy „látni” [1], mint egy valódi tárgyat. Az atomok kapcsolódási sorrendjét és térbeli elrendeződését indirekt módon, fizikai-kémiai mérési adatokból következtetjük ki. Ennek a szerkezet-meghatározási folyamatnak a végeredménye nagymértékben függ részint az adott molekuláról megszerezhető kísérleti adatok mennyiségétől és minőségétől, részint pedig az ezekből levont következtetések helyességétől [2]. Így tehát az, hogy valóban jó szerkezeti képletet társítunk-e ahhoz az anyaghoz, amiről beszélünk (vagyis a „számítógép” biztosan nem „kávéfőző”), közel sem olyan magától értetődő, mint ahogy azt sokan gondolják. Ezt mutatja, hogy a kémiai irodalomban közölt molekulaszerkezetek meglepően nagy arányban tévesek, amire már számos, a témát vizsgáló szerző rámutatott – ezek közül itt most csak négy jellemző példára hivatkozunk: [3-6].

Az, hogy nem szívesen kételkedünk a látott szerkezetek hitelenségében nem véletlen: ahhoz vagyunk szokva, hogy az ismeretek (beleértve a tudományos ismereteket) befogadása nagyrészt bizalmi alapon történik, és mentálisan roppant megterhelő lenne úgy olvasni a kémiai szakirodalmat, hogy közben folyamatosan azon gondolkodunk, vajon megbízhatunk-e a leírt szerkezetekben. Pedig – ahogy fentebb említettük – valójában egyáltalán nem mindegy, hogy milyen típusú és mélységű adatok és mennyire precíz okfejtés alapján állítják a szerzők egy anyagról, hogy annak mi a szerkezete. Még a „szerkezet” szó is értelmezési kérdéseket vet fel, például közel sem mindig nyilvánvaló, hogy „tér szerkezet” alatt a szintetikus vegyészek számára fontos abszolút konfigurációt értjük, vagy a molekula fontosabb konformációs tulajdonsá-

gait is ide soroljuk, mint például ahogy az egy 3D QSAR-modell (three-dimensional quantitative structure-activity relationship, azaz háromdimenziós mennyiségi összefüggés az atomok térbeli elrendeződése és a biológiai hatás között) felállításakor szükségessé válik. Tehát amikor egy anyag szerkezetéről beszélünk, valójában annak több szintjéről nyilatkozhatunk, és az egyes szintekhez eltérő bizonyossági tényezők tartozhatnak. A továbbiakban „szerkezet” alatt egy anyag konstitúciós képletét és relatív konfigurációját fogjuk érteni, „szerkezeti konfidencia” alatt pedig azt értjük, hogy e két tekintetben mennyire biztos állításokat tudunk tenni. Ahhoz, hogy megfelelő bizalommal fordulhassunk a kémiai szakirodalomban leírt képletekhez, arra van szükség, hogy minden kutató a leírt szerkezeti képleteket kellően magas szerkezeti konfidenciaszinten bizonyítsa, amihez be kell vetni a nagyműszeres szerkezetkutatás eszköztárát.

Cikkünk célja kettős. Egyrészt hangsúlyozni kívánjuk a szerkezetmeghatározás e lehetséges rejtett bizonytalanságait, másrészt rá kívánunk mutatni arra, hogy kellően tudatos alkalmazás mellett a legmodernebb NMR-es technikák milyen módon tudják növelni a szerkezetkutatás hatékonyságát és a közölt szerkezetek bizonyossági szintjét. E tekintetben különös hangsúlyt helyezünk a világszerte rohamosan terjedő nem egyenletes minta-vételezés (nonuniform sampling, NUS) kérdéskörére [7]. Igyekszünk – némileg didaktikus tárgyalást követve – rávilágítani a NUS mibenlétére, továbbá bemutatjuk a NUS gyógyszeripari alkalmazásával kapcsolatos saját módszertani kutatási eredményeinket [8,9]. Cikkünkben szemléletesen nagyban támaszkodunk a „Nagyműszeres szerkezetkutatás a Richter Gedeon Nyrt.-ben” című, szintén a Magyar Kémiai Folyóiratban megjelent cikkre [1].

2. Hétköznapi szerkezetkutatási esettanulmány arról, vajon kellően gondosan gondolkodunk-e (úgy általában)

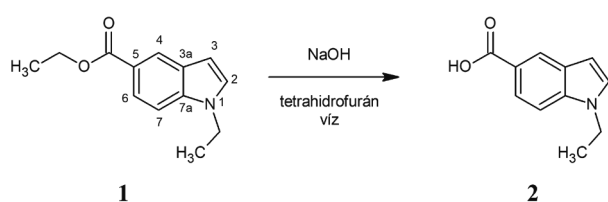
A következőkben tárgyalt példánk talán éppen azért tűnhet meglepőnek, mert rendkívül egyszerű és semmilyen dramaturgiai látványos végkifejletet nem tartalmaz: a spektroszkópiai alapon meghatározott molekulaszerkezet pontosan az, ami kémiai megfontolások alapján várható volt. Választásunk ebben a tekintetben tudatos. Ezzel a példával ugyanis éppen arra szeretnénk rámutatni, hogy annak ellenére, hogy a jelen esetben a *látszólag* gondos (de világszerte tipikusan gyakorolt) és a *valóban* gondos (a tudományos

* Tel.: +36 1 431 4952; fax: +36 1 432 6003; e-mail: cs.szantay@richter.hu

gondolkodás magas minőségi szintjét képviselő) dedukciós folyamat ugyanoda vezet, mégis milyen alapvető különbség van az eredmény bizonyítottóságának (vagyis konfidenciájának) a mértékében. Éppen ennek a bizonyítottósági szintnek a megléte vagy hiánya az, aminek többnyire nem vagyunk a tudatában a kémiai irodalom olvasásakor. Másrésztől a példánkkal azt is illusztrálni kívánjuk, hogy a nagy bizonyító erejű szerkezeti következtetések jóval több kísérleti adat megmérését és értelmezését is igénylik, ami gyakran vállalhatatlan többletidővel is járhat – hacsak nem vetünk be olyan korszerű módszereket, mint például a NUS!

2.1. Szerkezetigazolás ^1H NMR segítségével

Példánk az etil-1-etilindol-5-karboxilát (**1**) lúgos hidrolízisével (**1. ábra**) kapott termék (**2**) szerkezetvizsgálatáról szól.



1. Ábra. Egy egyszerű kémiai reakció, az indolgyűrű számozása és a várt termék (**2**) szerkezete.

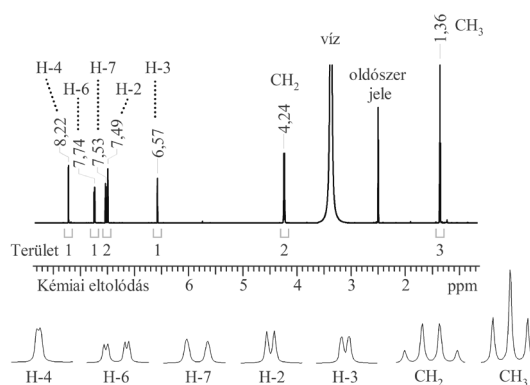
A manapság rutin módszernek számító kifelbontású tömegspektrometria (MS) segítségével a vizsgált anyag molekulatömege (M) 1 Da pontossággal adható meg, ez az érték jelen esetben az MS-mérés alapján $M=189$ -nek adódott, ami megfelel a várt terméknek.

A **2. ábrán** bemutatott ^1H NMR-spektrumból látható, hogy az aromás tartományban összesen ötszörös integrál (jel alatti terület) értékben vannak jeleink (egy-egy hidrogénatomtól származó jelek 6,57, 7,49, 7,53, 7,74 és 8,22 ppm-nél), ami megfelel az **1. ábrán** látható várt termék öt aromás hidrogénatomjának. Az egymással szomszédos H-2 és H-3 hidrogénatomokhoz tartoznak a 7,49 ppm és a 6,57 ppm kémiai eltolódású, egymással csatoló (azonos J_{HH} csatolási állandót mutató) és izolált spinrendszert adó dublett jelek. A többi aromás jel (H-6, 7,74 ppm; H-7, 7,53 ppm; H-4, 8,22 ppm) skaláris J_{HH} csatolási mintázata pedig diagnosztikus egy olyan triszubsztituált benzolgyűrűre nézve, amelyben két hidrogénatom egymáshoz képest *orto*, a harmadik pedig egyikükhöz képest *meta*, a másikukhoz képest *para* helyzetben van. Az ^1H NMR-spektrumban a további jelek (az 1,36 ppm kémiai eltolódásnál lévő háromszoros integrál intenzitású triplétt és a 4,24 ppm-nél jelentkező kétszeres integrál értékű kvartett) megfelel az NCH_2CH_3 egységnek. Minden eddigi információ tehát megfelel a várt szerkezetnek.

2.2 Az ^1H NMR-spektroszkópia korlátai a szerkezeti konfidencia tekintetében

Az előző alfejezet utolsó mondatában a kulcsszó a „konzisztencia” volt. Valóban, az esetek igen jelentős részében

azon az alapon tekintünk egy szerkezetet „bizonyítottnak”, hogy ilyen jellegű konzisztenciát találunk a (többnyire kémiai ismeretek alapján) várt szerkezet és a mért spektrális adatok között. Ennek a konzisztenciának a megléte azonban még nem tekinthető igazán alapos bizonyításnak mindaddig, amíg az összes olyan elképzelhető szerkezetet tudatosan ki nem zártuk, amire hasonló konzisztencia elvileg szintén igaz lehet. Így tehát ha egy pillanatra eltekintünk az **1. ábrán** bemutatott kémiai kontextustól (azaz a kémiai reakcióhoz tartozó háttér-információk és a reakcióval kapcsolatos általános ismeretek összességétől) és úgy tesszük fel a kérdést, hogy milyen szerkezetek rendelkezhetnek az ismert elemi összetételünkhöz, valamint ahhoz a spektrális adathalmazhoz, amit a **2. ábra** mutat, akkor az **1. ábrán** látható reakcióban várt termékén kívül egyéb szerkezetek is szóba jöhetnek (lásd alább).



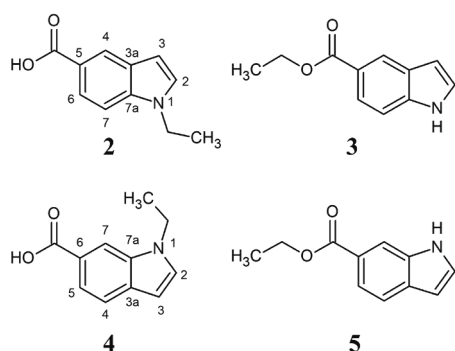
2. Ábra. A lúgos hidrolízis során kapott **2** termék ^1H NMR-spektruma.

Természetesen joggal vehető fel, hogy miért is tekintenénk el a kémiai kontextustól, hiszen az abból fakadó ismeretek, illetve feltételezések szerves részét képezik a szerkezet-meghatározási problémának. Ez valóban így van, és az esetek jó részében a kémiai kontextus ismeretének hiánya rendkívüli módon meg is nehezítené a szerkezet-meghatározást. A jelen példánk vonatkozásában például azt mondhatjuk, hogy az tény, hogy a metilénsoport két hidrogénatommagjának kémiai eltolódása 4,24 ppm önmagában még nem bizonyítja, hogy ez a metilénsoport egy nitrogénatomhoz kapcsolódik. Az NMR-szakkönyvekben lévő ^1H NMR kémiai eltolódás adatok áttekintését segítő táblázatokban (pl. [10]) ugyanis rendre az olvasható, hogy az oxigénatomhoz kapcsolódó metilénsoportok 4 ppm körüli kémiai eltolódásnál adnak jelet, míg a nitrogénatomhoz kapcsolódó metilénsoportok jele 3 ppm körül szóródik. Ezek az információk tehát bizonyítanálannak teszik, hogy az etilsoport valóban nitrogénatomhoz kapcsolódik-e, és felvetik annak a lehetőségét, hogy inkább *O*-etilsoport található a molekulában. Ez persze ellentmond a várt szerkezetnek és az általános szerves kémiai ismereteknek: egy lúgos hidrolízis reakcióban az észtercsoportnak kell bomlást szenvednie, ilyen körülmények között az aminfunkció nem szokott sérülni. Azonban azt is látni kell, hogy a kémiai megfontolásokon alapuló hipotézisalkotás olyan szerkezeti előítélethez vezethet, ami könnyen „megvezetheti” a szintetikus vegyészt és/vagy a szerkezetkutatót. Ezzel azért rendkívül fontos tisztában lenni, mert a saját gyakorlatunkban hihetetlen nagy szám-

ban talákoztunk olyan esetekkel, amikor a kémiai reakció nem a „tankönyvi” ismeretek alapján várt terméket adta, és ha pusztán a már említett konzisztencia megléte alapján állítottuk volna, hogy a szerkezet megfelel a vártnak, akkor biztosan „melléfogtunk” volna! Mindez különösen igaz a gyógyszeriparban igen gyakran vizsgált nyomszennyezők esetében, amik jellemzően éppen azért nyomszennyezők, mert a kémiai reakciók nagyon kis valószínűséggel megvalósuló mellékreakcióinak termékei – ezek a reakciók pedig néha rendkívül meglepőek és tipikusan nem kerülnek be a „tankönyvekbe”. Éppen ilyen okok miatt fontos a tudományos gondolkodásnak egy olyan fokú, és az adott szerkezetanalitikai probléma egyedi tulajdonságait olyan gondosan és kritikusan mérlegelő alkalmazása, ami a mindennapi gyakorlatban egyáltalán nem magától értetődő [2].

Ha ezen megfontolások alapján nézünk újra rá a **2.1 pontban** leírt gondolatmenetre, akkor a következő további megfontolásokat tehetjük. Az aromás ^1H NMR-jelek alakja alapján megerősítettük ugyan a benzolgyűrű szubsztitúciós mintázatát, de ebből még nem következik, hogy a karbonsav (vagy -észter) funkció tényleg az 5-ös pozícióban van (**vö. 1. ábra**). Hallgatólagosan elfogadtuk azt, hogy a kiindulási anyag szerkezete biztosan jó, de ezt nem szabad tényként kezelni, mert előfordulhat, hogy egy kémiai szintézis során egy sokkal korábbi lépésben a gyűrűzárás máshogy történt, mint ahogy reméltük.

A **3. ábrán** bemutatjuk azt, hogy milyen más lehetséges szerkezetek felelnek meg ugyanennek a ^1H NMR-spektrumnak, ha mélyebben vizsgáljuk meg a szerkezetkutatási problémát. Négy vegyületet (**2-5**) is fel tudunk rajzolni, amelyek összhangban állnak a mérési adatokkal. Természetesen a négy vegyület ^1H NMR-spektruma eltér egymástól, de a kémiai eltolódásbeli különbségek nem annyira jelentősek, hogy a már említett konzisztencia ne állna fenn.

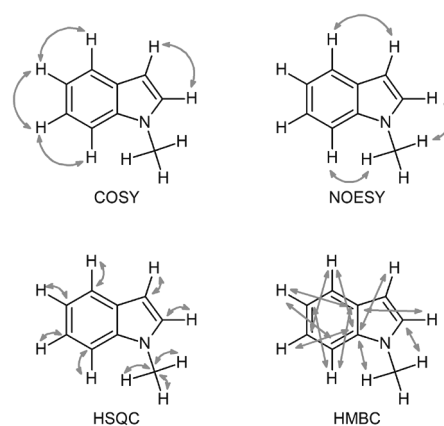


3. Ábra. A 2. ábrán bemutatott ^1H NMR-spektrum megfelel a **2-5** szerkezeteknek.

2.3.A szerkezeti konfidencia kétdimenziós NMR-spektrumok segítségével növelhető

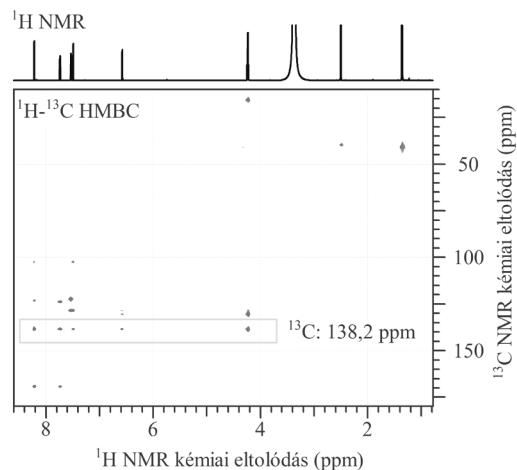
Mindezek után egy olyan, jóval gondosabb dedukciós folyamatban juthatunk el a végeredményhez, amelyben már a kémiai eltolódás korreláción alapuló kétdimenziós (2D) NMR-spektroszkópiát is igénybe vesszük. Az ilyen 2D NMR-módszerek lényege, hogy kapcsolatokat mérünk

ki az atommagok között. A mérni kívánt kapcsolat típusa alapján a négy legfontosabb kétdimenziós NMR-méréstípus a következő (**4. ábra**): ^1H - ^1H kapcsolatok kovalens kötések keresztül (correlation spectroscopy, COSY), ^1H - ^1H kapcsolatok téren keresztül (nuclear Overhauser effect spectroscopy, NOESY), ^1H - ^{13}C közvetlen kapcsolatok (heteronuclear single quantum correlation, HSQC), valamint ^1H - ^{13}C távolható kapcsolatok (heteronuclear multiple bond correlation, HMBC). Egy 2D NMR-spektrumban (pl. **5. ábra**) megjelenő spektrumcsúcs azt jelenti, hogy a mérni kívánt kapcsolattípust (korrelációt) sikerült kimutatni két atommag között. A csúcs koordinátái a kapcsolatban lévő két atommag kémiai eltolódását adják meg, amit a két tengelyen olvashatunk le. A csúcsmagasságot (jelintenzitást) kontúrvonalakkal jelöljük.



4. Ábra. A négy legfontosabb 2D NMR-kísélet, amelyek információ-tartalma kiegészíti egymást; a nyilak jelzik az adott kísérletben kimutatható kapcsolatokat.

Jelen példában a HMBC-spektrum segítségével (**5. ábra**) döntjük el, hogy a vizsgált anyag a **3. ábrán** látható szerkezeti képletek melyikének felel meg.



5. Ábra. A kémiai reakció termékéről felvett ^1H - ^{13}C HMBC-spektrum (az ábrán bekeretezett sornak jelentős szerepe lesz a későbbiekben).

Észrevehető, hogy a metilencsoport hidrogénatomjai (4,24 ppm) két aromás szénatomot (130,2 ppm és 138,2 ppm) „látnak” három kötési távolságra, ami bizonyítja, hogy az etilcsoport az indolgyűrűhöz kapcsolódik. Ha etilész-

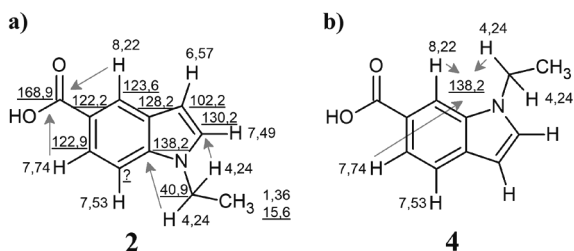
ter-csoport lenne a molekulában, ezek a hidrogénatomok a karbonil szénatomot „látnák” (168,9 ppm). Ezzel kizártuk a két etilészter-csoportot tartalmazó szerkezetet (**3** és **5**).

A **6a. ábrán** látható a várt 1-etilindol-5-karbonsav (**2**) szerkezet alátámasztása: mindegyik hidrogén- és szén-atomhoz tudunk olyan kémiai eltolódás értéket rendelni, ami összhangban van az adott szerkezeti elemmel és a HMBC-spektrumban fellelhető korrelációkkal is.

A **6b. ábra** mutatja az 1-etilindol-6-karbonsav (**4**) szerkezet cáfolatát. Az **5. ábrán** a HMBC-spektrumban a beke-retezett sorban látható, hogy a 138,2 ppm kémiai eltolódású szénatom a 8,22 ppm, 7,74 ppm és 4,24 ppm kémiai eltolódású hidrogénatomokkal ad nagy intenzitású korrelációt, emiatt a felsorolt hidrogénatomok legfeljebb három kötésre lehetnek a 138,2 ppm kémiai eltolódású szénatomtól. A **6b. ábra** alapján megállapítható, hogy semelyik szénatomhoz sem lehet hozzárendelni a 138,2 ppm kémiai eltolódást úgy, hogy a fenti feltételnek megfeleljen, így cáfoltuk ezt a szerkezetet is.

Ez a mélyen elemző megfontolás tehát az, ami a jelen példán keresztül a szerkezeti konfidencia illúziójával szemben a tudományos alaposággal alátámasztott valódi konfidenciát biztosítja. Egy ilyen elemzés birtokában a korábrinál jóval magasabb konfidenciaszinten állíthatjuk, hogy a vizsgált molekula az 1-etilindol-5-karbonsav.

A gyakorlatban az ilyen fajta dedukciós alaposág megvalósításának két fő korlátja van. Egyrészt a tudományos gondolkodás ennyire körültekintő mindennapi alkalmazása fokozott mentális erőfeszítést és kritikai attitűdöt igényel [2], másrészt pedig – ahogy fentebb is említettük – a 2D NMR-módszerek viszonylagos lassúsága miatt a szükséges mérési adatok megszerzésének időigénye gyakran túlságosan nagy. A kérdés tehát az, hogy hogyan lehet ez utóbbi problémát áthidalni.



6. Ábra. a) HMBC-spektrum alapján levezetett majdnem teljes ^1H és ^{13}C NMR kémiai eltolódás hozzárendelés (aláhúzás: ^{13}C), illetve a legfontosabb HMBC-korrelációk; b) az 1-etilindol-6-karbonsav szerkezetnek (**4**) ellentmondó a mérési adatok.

3. Gondolatok a 2D NMR-spektroszkópia időigényéről – mi is az a NUS

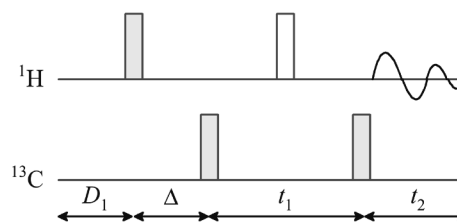
3.1. Néhány 2D NMR alapelv

A 2D NMR-mérések során olyan jelenségeket kell megfigyelnünk, amelyek a ^1H NMR-méréshez képest 10-1000-

szer kisebb jel/zaj-viszonnal mérhetőek. A kis jelintenzitás oka például a NOESY-mérésnél az, hogy a megfigyelt hidrogénatomok közötti távolság növekedésével a jelerősség rohamosan csökken (kétszeres távolság esetén $2^6 = 64$ -szer kisebb lesz a jel). A HSQC- és HMBC-mérések esetén a kis jel/zaj-viszony a ^{13}C izotóp túl kicsi, kb. 1,1%-os természetes előfordulása miatt van.

Az utóbbi évtizedek hardveres fejlesztéseinek (erősebb szupravezető mágnesek, fejlődés az elektronikában, hűtött mérőfejek elterjedése) köszönhetően ma már olyan nagy jel/zaj-viszonyt kaphatunk a ^1H NMR-mérés során ($>10^3:1$ egy 10 mg-nyi anyagot tartalmazó oldatról), hogy a fent említett jelvesztég ellenére is jól mérhetőek a 2D NMR-spektrumok egy közép kategóriás NMR-spektrométeren (pl. 400 MHz, forrásponti hőmérsékletű nitrogéngázzal hűtött mérőfej, $\geq 1050:1$ jel/zaj-viszony egy 0,1 tömeg%-os etilbenzol mintára gyártói specifikáció szerint).

A 2D NMR-mérések nagy időigénye a korszerű készülékeken egy másik problémára vezethető vissza, aminek megértéséhez leegyszerűsítve vázoljuk a ^1H - ^{13}C HMBC-mérés [**11**] (**7. ábra**) működését.



7. Ábra. A HMBC-mérés pulzusprogramja; a teli téglalapok 90° -os rádiófrekvenciás pulzust (rövid ideig tartó nagy teljesítményű besugárzást) jelképeznek, az üres téglalap pedig 180° -os pulzust.

Minden mérés előtt D_1 időtartamot (tipikusan kb. egy másodpercet) várakozunk azért, hogy a relaxációs folyamatok által kiépüljön a kellő mértékű longitudinális ^1H -mágneszettség, ami a mérés elindításához szükséges. Az első két pulzus és a közöttük lévő Δ időtartam azt szolgálja, hogy a ^1H -mágneszettség átalakuljon egy olyan állapotba, ami összeköttetést teremt a csatoló ^1H - és ^{13}C -atomok között. Az olyan molekulákat, amelyekben nincs ^{13}C -atom, nem kívánjuk detektálni (az ilyen molekulák zavaró hatásait ki kell szűrni), amit a pulzusprogramok megfelelő tervezésével lehet megvalósítani.

A t_1 időtartam alatt, azaz a ^{13}C -csatornán alkalmazott két pulzus között engedjük azt meg, hogy a ^{13}C -kémiai eltolódások befolyásolják a végső NMR-jelét. Az NMR-jelét a t_2 időben mintavételezzük a ^1H -csatornán; ez azt jelenti, hogy meghatározott időközönként leolvassuk az NMR-jel intenzitását és digitálisan eltároljuk azt (néhány ezer adatpont). Ezt az idődimenziót direkt dimenzióknak hívjuk, mert az adatsor pontjait közvetlenül a detektorból kapjuk.

A digitalizált jel (intenzitás- t_2 függvény) Fourier-transzformációjával (időtartománybeli jelből frekvencia-tartománybeli jellet számol) olyan ^1H NMR-spektrumhoz

jutunk, amelyben szelektíven a ^{13}C -atommagokkal csatoló ^1H -atommagok jelei jelennek meg. Ebben a spektrumban a ^1H -jelek fázisai különbözőek lesznek, a fázist befolyásolja az ^1H -atommaggal csatoló szénatom kémiai eltolódása, illetve a t_1 idő hossza.

A ^{13}C -kémiai eltolódások úgy mérhetőek meg, ha a ^1H -jelek váltakozó fázisát mintavételezzük a t_1 idő függvényében, majd a kapott jelet Fourier-transzformáljuk (8. ábra). A 2D NMR-mérések lassúságának oka arra vezethető vissza, hogy egy pulzusprogram-lefutással csak egyféle t_1 időt állíthatunk be, és ha a t_1 mentén szeretnénk mintavételezést végezni, akkor kénytelenek vagyunk a pulzusprogramot több százszor lefuttatni, és az egymást követő futások között inkrementálni (egységnyi mértékben megnövelni) a t_1 értékét. Ezt az idődimenziót indirekt dimenzióknak hívjuk, mert az egyes t_1 időpontokhoz tartozó információk külön mérésből származnak.

A kétdimenziós NMR-spektrum adottságait meghatározza mind a t_2 , mind a t_1 időkoordináta menti mintavételezési sűrűség és mintavételezési időtartam. A minimális mintavételezési sűrűséget (gyakoriságot) az határozza meg, hogy mekkora frekvencia (kémiai eltolódás) tartományban (spektrális ablakban) szeretnénk helyesen ábrázolni a spektrumot. Nagyobb frekvenciatartomány (gyorsabb oszcillációk) ábrázolásához sűrűbben kell mintavételezni, a mintavételezési időköz legfeljebb az ábrázolandó legnagyobb frekvencia kétszeresének reciproka lehet (Nyquist-kritérium [12]).

A mintavételezési időtartam pedig a spektrum digitális felbontását határozza meg. Minél közelebbi frekvenciákat (kémiai eltolódásokat) szeretnénk egymástól megkülönböztetni, annál hosszabb t_2 , illetve t_1 ideig kell venni a jelet (a hosszabb ideig megfigyelt jeltől pontosabban lehet meghatározni a frekvenciát). A jeldetektálás során a t_2 idő növelése nem befolyásolja jelentősen az összes mérésidőt, de a t_1 idő növelése azt jelenti, hogy az egész pulzusprogramot többször kell lefuttatni. A szükséges t_1 -inkrementumok számát a maximális t_1 -időhossz (ami annál nagyobb, minél jobb felbontást szeretnénk elérni az F_1 -frekvenciadimenzió

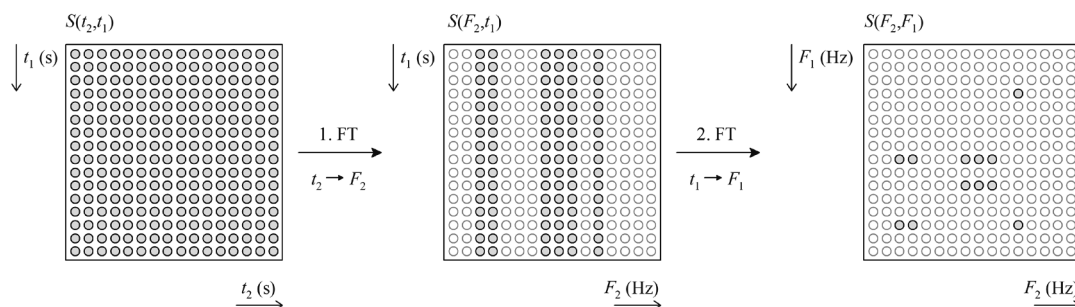
mentén) és a léptetési köz (ami annál kisebb, minél nagyobb tartományban szeretnénk helyesen ábrázolni az F_1 -frekvenciákat) hányadosa adja meg. Ha egy 200 ppm széles ^{13}C -kémiai eltolódás tartományban (F_1 -ablakszélesség) 1 ppm / sor felbontást szeretnénk elérni, akkor nagyságrendileg 200 t_1 -inkrementumra van szükség, tehát a t_1 -inkrementálás hosszadalmas művelete nem kerülhető ki a hagyományos 2D NMR-mérés során. Ezért egy 2D NMR-spektrum felvétele tipikusan több mint fél órát igényel.

A 2D NMR-spektrumok mérésidő-igényének csökkentésére többféle megközelítés ismert (pl. [7, 13-15]), de rutinszerű alkalmazásukat akadályozza, hogy a műszergyártók nem foglalkoztak a legtöbb módszer implementálásával. Kivételt jelentett ez alól a NUS [7], ami éppen ettől vált népszerűvé. Ezért fektetünk nagy hangsúlyt a NUS bemutatására.

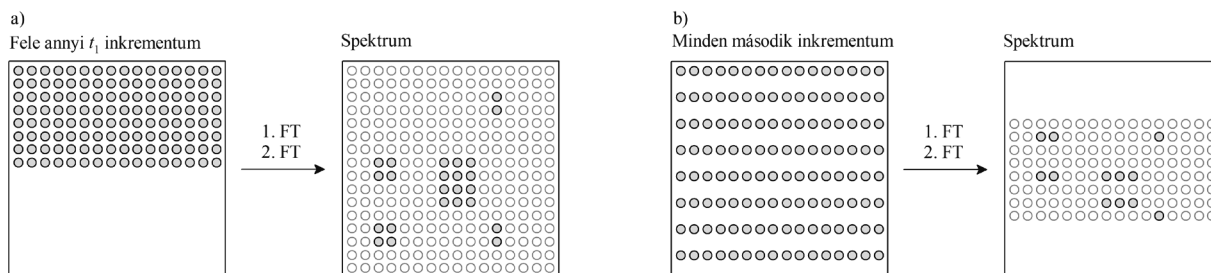
3.2. Az indirekt dimenzió menti mintavételezés módszertani újragondolása

A 2D NMR-spektrumok (5. ábra és 8. ábra) azt a benyomást kelthetik, hogy nagy részük üres, a hasznos jelek kis területekre koncentrálnak, tehát az a megérzésünk lehet, hogy sok mérésidőt veszítünk el az üres területek mérésére. Habár a Fourier-transzformáció sajátosságai miatt minden mérési pont értékes adatpont (mindegyik időtartománybeli pontot felhasználjuk mindegyik frekvenciakomponens kiszámításához), mégis érdemes átgondolni, hogy mi történik, ha bizonyos mérési adatpontokat kihagyunk, azaz, ha t_1 sorok kihagyásával szeretnénk gyorsítani a 2D NMR-méréseket.

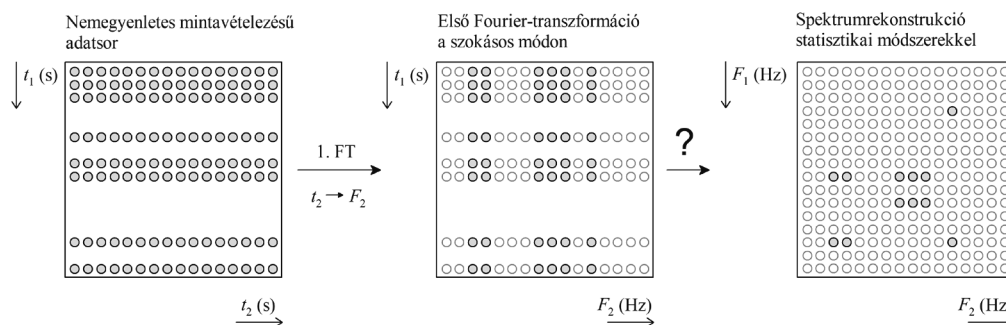
Tegyük fel, hogy csak fele annyi időnk van megmérni a spektrumot, mint amennyi valójában szükséges. Megtehetjük, hogy nem várjuk meg a mérés végét, az előre beállított inkrementumok felét vesszük csak fel (9a. ábra). Így csak rövidebb t_1 ideig tudjuk venni a jelet az indirekt dimenzióban, ami azt jelenti, hogy rosszabb lesz a felbontás, kétszer szélesebbek a jelek a függőleges dimenzióban az eredeti spektrumhoz képest. A spektrumban a jelek egy része összefolyik, a spektrumértelmezés nehézkessé, esetleg lehetetlenné válik.



8. Ábra. A mért adatpontok, az első Fourier-transzformáció sorozat (1. FT) után kapott köztes adatsor és a második FT után létrejövő 2D NMR-spektrum; az üres karikák nem hasznos (zajszintű) adatpontokat jelölnek



9. Ábra. a) Mérésidő-megtakarítás fele annyi inkrementum felvételével, illetve b) minden második inkrementum kihagyásával



10. Ábra. NUS mintavételezésű adatsor, ami hagyományos módszerekkel nem dolgozható fel

A mérési idő felét úgy is megspórolhatjuk, hogy kihagyunk minden második inkrementumot (9b. ábra). Mivel a spektrum (függőleges) felbontása a mintavételezett t_1 tartomány szélességétől függ, és ez a teljes mintavételezéshez (8. ábra) képest nem változik érdemben, ezért a 2D NMR-spektrum felbontása jó marad. A mintavételezés ritkításával azonban megfelezzük az ábrázolható frekvenciák tartományát (ezért van kifehéritve a 9b. ábrán a spektrum felső és alsó negyede), a leszűkült spektrális ablakból kilógó jelek visszahajlanak. A visszahajlás módszerével felvett 2D NMR-spektrumban tehát hamis frekvenciák megjelenésével fizetünk a megspórolt időért, de a hamis frekvenciák szabályosan keletkeznek és emiatt az eredeti spektrumot a legtöbb esetben vissza lehet fejteni.

Megtehetjük azt is, hogy a 9b. ábra mérési adatsorával szemben, nem szabályos rendben hagyunk ki inkrementumokat a mérés során. A 10. ábrán látható mérési módszer azért előnyös, mert egyrészt az adatsornak mind az elejéről, mind a végéről (nagy t_1 tartományban) veszünk jelet, ami a jó felbontást biztosítja, másrészt bizonyos t_1 szakaszokon sűrű mintavételezést végzünk, amivel el lehet kerülni a spektrális ablak beszűkülését és így a csúcsok visszahajlását. Viszont a NUS mintavételezéssel felvett adatsor feldolgozása komoly fejtörést okozott a jelfeldolgozással foglalkozó kutatóknak, mert a széles körben használt Fourier-transzformációs jelfeldolgozási algoritmus csak egyenletesen mintavételezett adatsorokra alkalmazható.

3.3.A nem egyenletes mintavételezéssel kapott adatsor feldolgozása

A NUS-mintavételezésű jeltől a spektrum előállítására rendkívül nehéz feladat. A NUS-os adat feldolgozásának alap-

jait Kazimierczuk és munkatársai cikke [16] alapján tárgyaljuk, de jelentős egyszerűsítésekkel élünk a könnyebb érthetőség érdekében.

A hagyományos (egyenletes) mintavételezés során a spektrumfeldolgozáskor a Fourier-transzformációt (FT) használjuk, ami egy intenzitás–idő bemeneti függvényre egy intenzitás–frekvencia kimeneti függvényt, azaz spektrumot ad. A FT során a bemeneti adatpontok száma megegyezik a kimeneti adatpontok számával.

Ezzel szemben a NUS-adatsor hiányos bemeneti függvényt jelent, a kitalálendő spektrális pontok száma nagyobb a bemenő adatpontok számánál. A sokismeretlenes, kevés egyenletből álló egyenletrendszerrel analóg módon, a NUS-adatsor feldolgozása során nincs egyértelmű kimenet. A különböző kutatócsoportok által kidolgozott algoritmusoknak más-más előnye és hátránya van.

A közös pont a spektrumrekonstrukciós algoritmusok egy nagy csoportjában az, hogy valamiképp kihasználják, hogy a két- és többdimenziós NMR-spektrumok nagy része üres (lásd 5. és 8. ábrák), amit a szakirodalom „sűrített érzékelésnek” – compressed sensingnek (CS) – hív. Nem baj, ha a spektrumnak egy nagy hányadát, a zajt nem tudjuk adatpontról adatpontra kitalálni. Elégséges az, ha csak a hasznos jeleket találjuk meg a spektrumokban.

Legyen egy olyan jelünk, amelynek két frekvencia-komponense van (az adatpontok a 10. ábrán lévő spektrumban a 12. oszlop információtartalmának feleltethetők meg). Ha megmérnénk az összes intenzitás–idő adatpontját a jelnek, akkor egy ritkább és egy sűrűbb oszcillációt látnánk (11. ábra teteje). Ehhez képest hagyjuk ki a mérési pontok

50%-át véletlenszerűen. A kihagyott adatpontok helyére pótlásként nullákat írunk be (1. lépés).

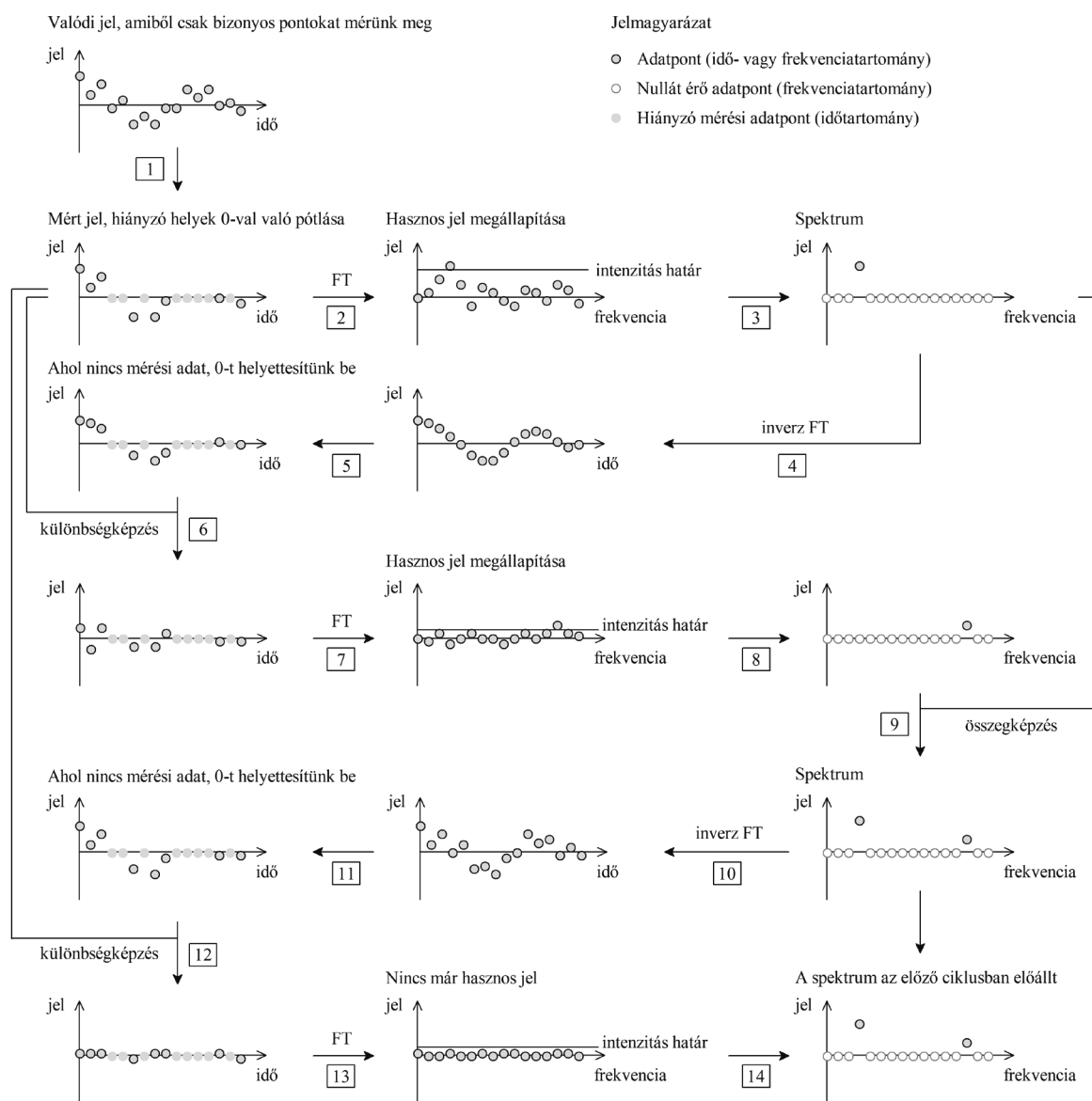
Ezután Fourier-transzformációt végzünk a nullákkal kiegészített adatsoron, ugyanúgy, mintha egy hagyományosan mintavételezett jelet akarnánk feldolgozni. A mesterségesen beírt nullák miatt hibás spektrumot kapunk a FT során. A hiba a teljes spektrumtartományra eloszlik, ezáltal a legnagyobb intenzitású, hasznos jelek láthatóak maradnak (2. lépés). Egy bizonyos jelintenzitás felett úgy tekintjük, hogy hasznos jeleket találtunk, a hasznos jelek pozícióját (frekvenciáját) és intenzitását eltároljuk (3. lépés) a készülő spektrumban.

A készülő spektrumot inverz-Fourier-transzformáljuk (4. lépés), az intenzitás–idő függvény alakjában is látszik, hogy most még csak egy frekvenciakomponensből áll a spektrum. Az időpontok egy részében nincs valódi mérési adat, ezért

azokat kitöröljük és 0-val helyettesítjük (5. lépés). Ezután kivonjuk a mérési adatsorból az 5. lépésben előállított függvényt, tehát megnézzük, hogy miként néz ki a mért jel, ha a legnagyobb intenzitású frekvencia-komponens kivesszük belőle (6. lépés).

Ezzel a különbség jellel kezdjük a következő elemzési ciklust (7–12. lépés). Csak akkor lépünk ki a ciklusból, ha az alkalmazott statisztikai módszer szerint nem találunk új hasznos jelet (13. lépés). Ekkor véget ér a spektrumrekonstrukció (14. lépés), kimeneti spektrumként elfogadjuk azt, ami az előző ciklusban előállt.

Az egyes CS-algoritmusok többek között abban térnek el egymástól, hogy milyen statisztikát alkalmaznak a hasznos jel megállapítására, a hasznosnak vélt jelet mekkora mértékben összegzik a készülő spektrumba, illetve milyen szabály szerint tekintik befejezettnek az iterációt.



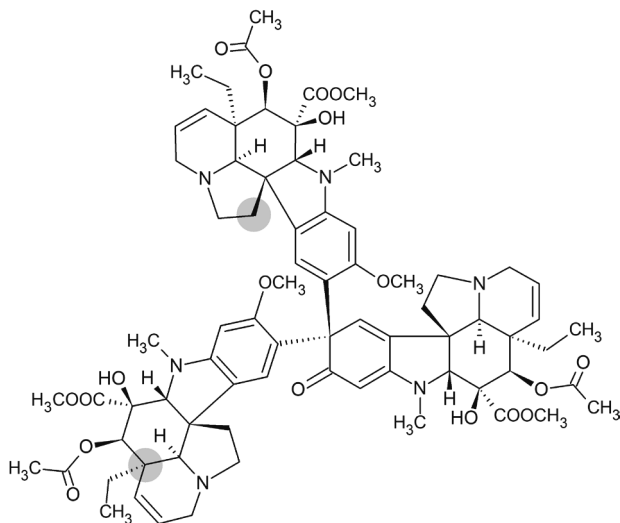
11. Ábra. A NUS-mintavételezésű adatsor feldolgozásának egyszerűsített folyamatábrája.

4. Néhány egzotikus szerkezetmeghatározási példa, amit NUS segítségével oldottunk meg

Azon túl, hogy a NUS-t napi szinten használjuk mérésünk hatékonyságnövelésére, néhány különleges szerkezetkutatási problémát szeretnénk részletesebben bemutatni, ami NUS nélkül nem lett volna gazdaságosan megoldható.

4.1. Vindolin-trimer szerkezetmeghatározása

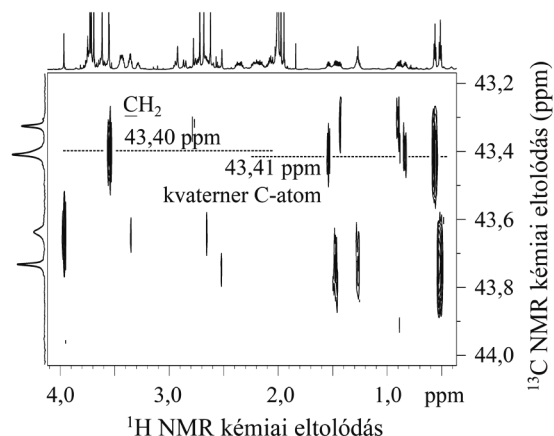
Egy 74 szénatomos trimer alkaloid (**12. ábra**) szerkezetmeghatározása [8] azért okozott számunkra nehézséget, mert a ^{13}C NMR-spektrum bizonyos tartományaiban olyan zsúfolt volt a spektrum, hogy a HMBC-spektrum megbízható értékeléséhez a ^{13}C -dimenzió mentén olyan nagy felbontóképességre volt szükség, amit csak sok ezer t_1 -inkrementummal lehetett volna elérni.



12. Ábra. A vindolin-trimer NMR és MS segítségével meghatározott szerkezeti képlete; szürke körrel jelöltük a 13. ábrán említett szénatomokat.

NUS-sal kevesebb mint egy óra alatt kiváló felbontású HMBC-spektrumot vettünk fel a mintáról. Segítségével képesek voltunk például a 43,40 ppm kémiai eltolódású ^{13}C -atomhoz (metilénsoport) tartozó jeleket megkülönböztetni a 43,41 ppm-es sorban lévő jelektől (kvaterner szénatom). Ez a két ^{13}C -rezonancia még a ^{13}C NMR-spektrumban is átfedett, ami ritka jelenség és jól érzékelteti a vizsgált molekula bonyolultságát. Hagyományos mintavételezéssel csak éjszakára indított mérésrel lehetett volna elérni ugyanilyen felbontást, és ha ennél kevesebb időt szánnánk a mérésre, akkor a **12. ábrán** látható szerkezetet nem tudtuk volna nagy konfidenciával bizonyítani.

A HMBC-spektrumot úgy vettük fel, hogy 8192 inkrementumból csak 512-t mértünk meg, amivel mintegy 94%-kal csökkentettük a mérésidőt. A NUS-spektrum rekonstrukcióját a mérő szoftver automatikusan elvégezte, ami kevesebb mint fél óráig tartott.

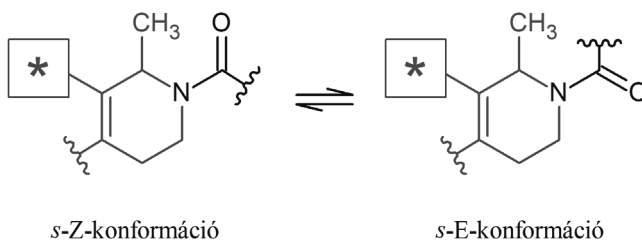


13. Ábra. A vindolin-trimer NUS HMBC-spektrumának egy részlete.

4.2. Döntés konstitúciós és konformációs izoméria között

Egy minta várt főkomponense egy olyan molekula volt, ami egy savamidsoportot tartalmaz, de ismeretlen okból nem a megszokott kettő, hanem négy jelsorozat jelent meg az NMR-spektrumokban, 52:22:18:8 arányban. Túl kevés anyag állt rendelkezésre (kb. 3,5 mg) ahhoz, hogy a három kisebb intenzitású jelsorozatra is el lehessen végezni a szerkezetigazolást [9].

Emiatt nem lehetett azt kizárni, hogy nem egységes az anyag (**14. ábra**), lehetséges, hogy két konstitúciós izomer két-két konformerét láttuk a spektrumokban. Ezt cáfolni úgy lehet, hogy bebizonyítjuk, hogy a különböző jelsorozatokhoz tartozó molekulák dinamikus egyensúlyban vannak egymással, a hőmozgás által a konformerek lassan egymásba alakulnak.



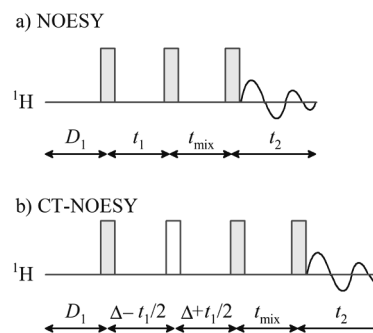
14. Ábra. A négy jelsorozatot adó vegyület szerkezeti képletének egy részlete; a csillag egy olyan részét jelöli a molekulának, amelynek a konstitúciójában nem voltunk biztosak.

A bizonyításhoz két módszert használhatunk. Felfűtjük a mintát, aminek hatására gyorsulnak a konformációs átalakulások, és megfigyeljük a különböző konformerhez tartozó jelek koaleszcenciáját (kiszélesedését és egyesülését).

A másik módszer az, hogy megfigyeljük a konformerek közötti kémiai kicserélődést. Ehhez a NOESY-pulzusprogramot (15. ábra) használjuk (ugyanazt, mint amellyel a ^1H - ^1H térközelségeket detektáljuk). A pulzusprogram elején (t_1) a molekula az egyik konformációs állapotban van, a pulzusprogram közepén (t_{mix}) várakozunk, hogy megtörténjen a konformációs átalakulás, a pulzusprogram végén (t_2) pedig detektáljuk annak tényét, hogy a molekula átkerült egy másik konformációs állapotba. A két konformer eltérő térszerkezete miatt ugyanaz a hidrogénatom két különböző kémiai környezetet érzékel a két konformerben, tehát megváltozik a kémiai eltolódása egy adott hidrogénatomnak a konformációs átalakulás során (például 2,2 ppm-ről 2,5 ppm-re). A mérés során a pulzusprogram elején regisztráljuk az ^1H -atommag kémiai eltolódását az egyik konformációs állapotban (a csúcs helye függőleges tengelyen: 2,2 ppm), a pulzusprogram végén pedig az ^1H -atommag kémiai eltolódását a másik konformációs állapotban (a csúcs helye a vízszintes tengelyen: 2,5 ppm). Természetesen a molekulák egy része a pulzusprogram alatt nem cserél konformációt, illetve egy részük a többszöri csere miatt a kiindulási állapotba tér vissza. Ezeket a molekulákat a spektrum főátlójában detektáljuk (a pulzusprogram elején és a végén azonos az ^1H NMR kémiai eltolódás).

Mivel a négy jelsorozat jelei nagyon közel estek egymáshoz (részlegesen át is fedtek), a kémiai kicserélődés detektálásához rendkívül nagy felbontást szerettünk volna elérni mindkét dimenzióban. A 14. ábrán feltüntetett metilcsoporthoz tartozó dublett jeleken terveztük megfigyelni a kémiai kicserélődést (a dublett jelalak a szomszédos CH-csoport hidrogénatommagjával való csatolás miatt jön létre). A nagy felbontás miatt a csúcsok dublett szerkezete a függőleges dimenzió mentén is láthatóvá vált, ami nehezítette a NUS NOESY-spektrum rekonstrukcióját (16a. ábra).

A dublett felhasadás kiküszöbölésének érdekében házon belül módosítottuk [9] a standard NOESY-szekvenciát (15a. ábra), amelybe beépítettünk egy „constant time” (CT) [17] egységet (15b. ábra), amivel egy új variánsát hoztuk létre a NOESY-szekvenciának. A „constant time” kifejezés arra utal, hogy az első két 90° -os pulzus között eltelt időt állandósítottuk (2Δ), ami a standard szekvenciában egy inkrementumonként változó (t_1) időtartam volt. Ezzel a technikával a NOESY-spektrumban a csúcsok dublett szerkezete szingulettre egyszerűsödött a függőleges tengely mentén.



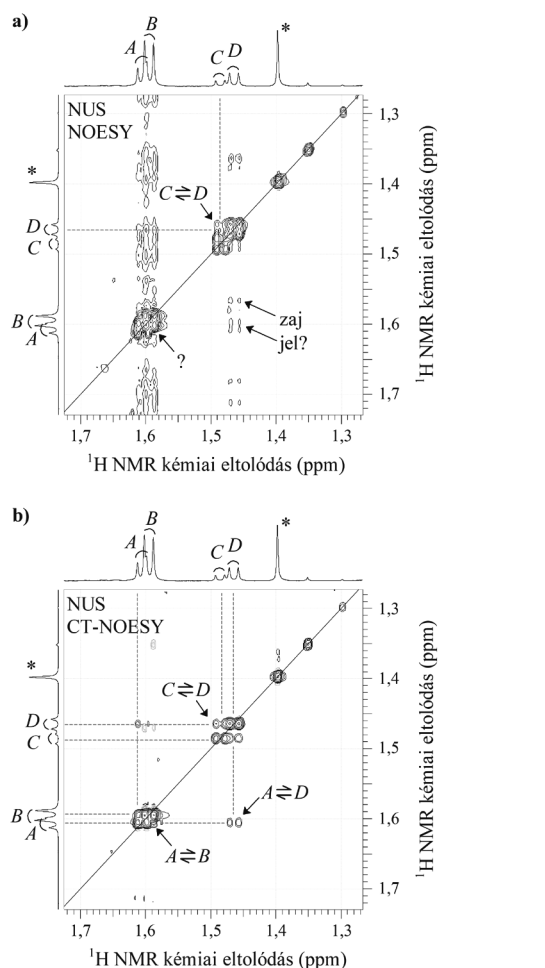
15. Ábra. a) NOESY-kísérlet és b) az általunk kidolgozott „constant time” (CT) NOESY-kísérlet pulzusprogramja; a teli téglalapok 90° -os pulzust jelentenek, az üres téglalap pedig 180° -os pulzust.

A csúcsok alakjának egyszerűsödése könnyebbé tette a CS-algoritmus számára a spektrum rekonstrukcióját, tehát a CT-technika nélkül felvett spektrumhoz képest (16a. ábra) jobb minőséget értünk el (16b. ábra). Bebizonyítottuk, hogy mind a négy jelsorozat ugyanabban a kémiai kicserélődési folyamatban érintett. A következő cserefolyamatokat láttuk: A jelsorozat B-vel és D-vel, illetve C jelsorozat D-vel. Ezzel igazoltuk, hogy a négy jelsorozat négy konformert jelent, tehát egységes volt az anyag. NUS nélkül az 5-6 perces mérések időigénye több mint fél óra lett volna.

4.3. Nyomszennyező szerkezetmeghatározásának elősegítése

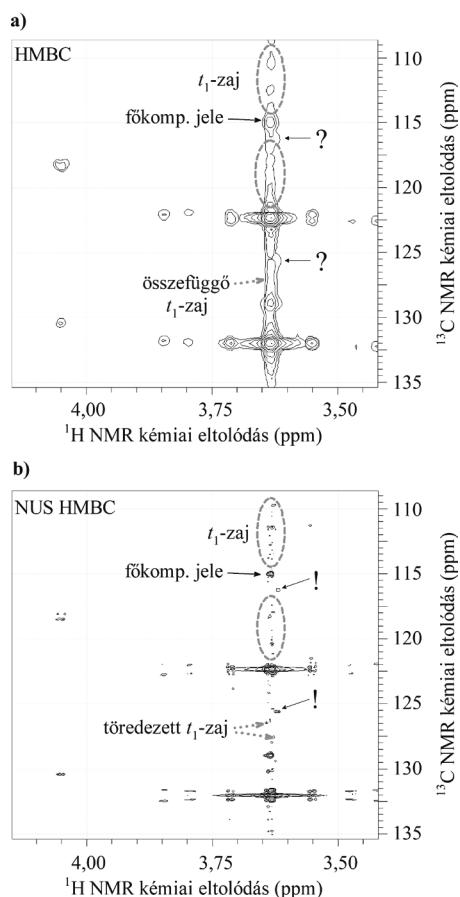
Adott volt egy viszonylag kis molekulatömegű fenilecet-savszármazék, amit egy gyártásban kívántak felhasználni alapanyagként. Az alapanyagban egy nyomszennyezőnek a szerkezetét kellett meghatározni úgy, hogy a főkomponens és a keresett szennyező mellett voltak még más 0,01–0,14%-os mennyiségű szennyezők is a mintában. A szerkezetmeghatározáshoz ^1H - ^{13}C HSQC és ^1H - ^{13}C HMBC-spektrumokat vettünk fel egy meglehetősen tömény (1,4 mol/dm³ koncentrációjú) mintáról [9]. Még egy 800 MHz-es, hűtött mérőfejjel ellátott NMR-spektrométeren is igen nagy kihívást jelent megmérni egy ilyen mintát, ha a szennyezők szerkezetmeghatározása a feladat.

A NUS előnyét a hagyományos mintavételezéshez képest a 17. ábrán mutatjuk be. A 17a. ábrán a hagyományos mintavételezésű HMBC-spektrum található. A főkomponens metilencsoportjához tartozó ^1H NMR-jel 3,63 ppm-nél található. Nagy intenzitása azt okozza, hogy az összes 2D NMR-spektrumban a 3,63 ppm-es oszlop mentén végig futó spektrumhibát okoz, amit t_1 -zajnak hívunk. Ebből a zajból a főkomponenshez tartozó leggyengébb, 4-5-kötéses HMBC-korrelációk alig emelkednek ki, detektálásukhoz olyan csúcskategóriás spektrométerre van szükség, amin a t_1 -zaj mértéke viszonylag kicsi. A 3.3. fejezetben bemutatunk a NUS-nak azt a tulajdonságát, hogy zajszinten lévő csúcsmagasságokat rosszul rekonstruál. Ennek az a következménye, hogy a t_1 -zaj rekonstrukciója elromlik, széttöredeznek a t_1 -zajhoz tartozó csúcsok a szomszédos oszlopok mentén. Emiatt a hasznos jelek detektálása könnyebbé válik, amit az alábbi példa is szemléltet.



16. Ábra. a) NUS NOESY-spektrum; b) a kémiai kicserélődés bizonyításához használt NUS CT-NOESY-spektrum, a négy jelsorozat A, B, C és D szimbólumokkal jelöltük

Az egyik 0,14%-os szennyezőnek az egyik ^1H NMR-rezonanciája 3,62 ppm-nél volt. Ezt a jelet az ^1H NMR-spektrumban a főkomponens 3,63 ppm-nél lévő jele teljesen elfedte. A nyomszennyezőhöz tartozó jeleket a HMBC-spektrumban sem vettük észre (**17a. ábra**, kérdőjellel jelölve), mert részben átfedtek a főkomponenshez tartozó, 3,63 ppm oszlopban lévő t_1 -zajjal. Ezzel szemben a NUS-HMBC-spektrumban (**17b. ábra**) a t_1 -zaj a korábban említett okok miatt széttöredezett, ezáltal felismerhetővé váltak a nyomszennyezőhöz tartozó korrelációk ^1H : 3,62 ppm, ^{13}C : 116,2 és 125,6 ppm-nél (felkiáltójelekkel jelölve). Megjegyezzük, ha a kis koncentrációjú komponensek mérését a túl kis jel/zaj-viszony is akadályozza, a jelek detektálhatósága önmagában a NUS használatától nem javul.



17. Ábra. a) A fenilecetsav-származék hagyományos mintavételezéssel készült HMBC-spektruma (16 ismétlés; 512 inkrementum; 3,2 óra mérésidő); b) ugyanennek a mintának a NUS HMBC-spektruma (16 ismétlés; 1024 inkrementum mintavételezése a 8192-ből; 6,9 óra mérésidő). A t_1 -zaj a 3,63 ppm-es oszlop mentén végig fut.

Ez a példa arra világított rá, hogy a szennyezésvizsgálat során előnyünkre tudtuk fordítani a NUS egyik fő hiányosságát, a zajrekonstrukcióra való képesség hiányát. Megjegyzendő, hogy ezt az effektust csak abban az esetben lehet kihasználni, ha megvan a műszer képessége arra, hogy jól csillapítsa a t_1 -zajt a spektrum felvétele során.

5. Összefoglalás

Ahhoz, hogy kellően magas konfidenciaszinten tudjunk egy szerkezetet meghatározni, sokkal inkább szükséges a 2D NMR-módszerek alkalmazása, mint ami a jelenlegi általános gyakorlat. Ehhez nemcsak a technológia meglétére van szükség, hanem megfelelő szemléleti tudatosságra is. Az elmúlt évtizedek hardveres fejlesztései az NMR-mérések érzékenységét a sokszorosára növelték, azonban a klaszikus 2D NMR-spektrumok felvételi ideje inhereus technikai okokból mégis relatíve lassú maradt. Az utóbbi évek szoftveres fejlesztései – mint például a NUS spektrum-rekonstrukciós algoritmusok – lehetővé tették, hogy nagyságrendekkel gyorsabban jó minőségű 2D NMR-felvételek készüljenek, ezáltal rövid átfutási idő mellett is jelentősen növelve a magas szerkezeti konfidencia elérhetőségét. A

módszer általunk történt továbbfejlesztését is alkalmazva, saját reprezentatív gyógyszeripari példákon keresztül mutattuk be a NUS rendkívüli hasznosságát.

Hivatkozások

1. Szántay, Cs. ifj.; Sánta, Zs.; Háda, V. *Magy. Kém. Foly.* **2016**, *122*, 124-133.
2. Szántay, Cs. Jr., Ed. *Anthropic Awareness: the human aspects of scientific thinking in NMR spectroscopy and mass spectrometry*. Elsevier: New York, **2015**.
<https://doi.org/10.1016/C2013-0-06828-4>
3. Borman, S. (2014). Tug Of War Over Promising Cancer Drug Candidate. *Chem. Eng. News*, on-line edition, available online at <http://cen.acs.org/articles/92/web/2014/05/Tug-War-Over-Promising-Cancer.html>
<https://doi.org/10.1021/cen-09221-notw1>
4. Carvalho, E.M., Periera, F.A., and Junker J. How well does NMR behave in natural products structure determination? A survey of natural products published in 2007 and 2008. Poster presented at the 50th Experimental NMR Conference, 2009, March 29 – April 3, Asilomar, California.
5. Elyashberg, M.E., Williams, and A.J., Blinov, K.A. *Contemporary Computer-Assisted Approaches to Molecular Structure Elucidation*. RSC: Cambridge, **2012**.
<https://doi.org/10.1039/9781849734578>
6. Elyashberg, M.E., and Williams, A.J. (2015). Computer-Based Structure Elucidation from Spectral Data. *The Art of Solving Problems*. Springer: Heidelberg, **2015**.
<https://doi.org/10.1007/978-3-662-46402-1>
7. Kazimierczuk, K.; Orekhov, V. *Angew. Chem. Int. Ed.* **2011**, *50*, 5556-5559.
<https://doi.org/10.1002/anie.201100370>
8. Szigetvári, A.; Keglevich, A.; Keglevich, P.; Dékány, M.; Hazai, L.; Szántay, Cs. Jr. *Struct. Chem.* **2019**, *30*, 795-804.
<https://doi.org/10.1007/s11224-018-1267-1>
9. Szigetvári, A.; Szántay, Cs. Jr. *Magn. Reson. Chem.* **2020**, 1-23 (online megjelenés).
<https://doi.org/10.1002/mrc.5085>
10. Hore, P. J. *Mágneses magrezonancia*. Nemzeti Tankönyvkiadó: Debrecen-Budapest, **2004**.
11. Summers, M. E.; Marzilli, L. G.; Bax A. *J. Am. Chem. Soc.* **1986**, *108*, 4285-4294.
<https://doi.org/10.1021/ja00275a008>
12. Bain, A. D.; Burton, I. W. *Concepts Magn. Reson.* **1996**, *8*, 191-204.
[https://doi.org/10.1002/\(SICI\)1099-0534\(1996\)8:3<191::AID-CMR3>3.0.CO;2-U](https://doi.org/10.1002/(SICI)1099-0534(1996)8:3<191::AID-CMR3>3.0.CO;2-U)
13. Vitorge, B.; Bieri, S.; Humam, M.; Christen, P.; Hostettmann, K.; Muñoz, O.; Loss, S.; Jeannerat, D. *Chem. Commun.* **2009**, 950-952.
<https://doi.org/10.1039/b820478k>
14. Giraudeau, P.; Frydman, L. *Ann. Rev. Anal. Chem.* **2014**, *7*, 129-161.
<https://doi.org/10.1146/annurev-anchem-071213-020208>
15. Claridge, T. D. W.; Mayzel, M.; Kupče, E. *Magn. Reson. Chem.* **2019**, *57*, 946-952.
<https://doi.org/10.1002/mrc.4887>
16. Shchukina, A.; Kasprzak, P.; Dass, R.; Nowakowski, M.; Kazimierczuk, K. *J. Biomol. NMR* **2017**, *68*, 79-98.
<https://doi.org/10.1007/s10858-016-0068-3>
17. K. Furihata, H. Seto *Tetrahedron Lett.* **1998**, *39*, 7337-7340.
[https://doi.org/10.1016/S0040-4039\(98\)01574-3](https://doi.org/10.1016/S0040-4039(98)01574-3)

NMR spectroscopy in the pharmaceutical industry: towards higher structural confidence via nonuniform sampling

In this article we discuss the topic of structural confidence, involving both the technology and the mindset required to assure a highly secure chemical structure. The chemical structures published in the scientific literature are much less accurate than it is generally thought, and their validity depends on the comprehensiveness of the measured spectroscopic data as well as on the prudence of the interpretation of these analytical data. Today, structure elucidation mainly relies on the concerted use of nuclear magnetic resonance (NMR) spectroscopy and high-resolution mass spectrometry (HRMS). However, the acquisition of a comprehensive set (^1H NMR, ^{13}C NMR, two-dimensional ^1H – ^1H connections through bonds and through space, and ^1H – ^{13}C connectivity through one bond and multiple bonds) of NMR data – which is needed to achieve a truly high level of structural confidence together with an in-depth spectral characterization – often requires a lot of time that we cannot afford in practice. A comprehensive set of experimental NMR data can however be collected in a relatively short time by using “accelerated acquisition” techniques such as nonuniform sampling which is now becoming increasingly popular among NMR spectroscopists.

As a case in point, in the first part of the paper we present the structure elucidation of the product of a simple chemical reaction at two different levels of structural confidence. First, the successful hydrolysis of ethyl 1-ethylindole-5-carboxylate was verified at a lower confidence level using MS and ^1H NMR. That approach is quite resource-friendly and is seemingly thoughtful, and in fact it represents the prevailing practice for “routine” structure elucidation in the field of organic synthesis. Secondly, we demonstrate that by the careful (e.g., unbiased by the knowledge of the chemical context and by the structural expectation stemming thereof) interpretation of the spectra one can find alternative structures that are also consistent with the measured data. By ruling out these alternative structures via a more comprehensive set of NMR measurements, one can reach a higher structural confidence regarding the originally proposed structure. High-confidence structure elucidation is especially important in the pharmaceutical industry because of the regulatory requirements for meeting high quality standards.

For the sake of better understanding the time issues associated with the acquisition of conventional two-dimensional (2D) NMR spectra, we provide a brief introduction to 2D NMR spectroscopy by the example of HMBC measurements. We explain that these 2D NMR measurements are time-consuming due to the inherent need for several hundred steps of incrementation of the t_1 parameter. Then, we discuss what happens when some t_1 increments are intentionally and regularly left out from the measurement: either the resolution decreases or the correctly interpretable chemical shift range narrows (spectral aliasing). When the t_1 increments are left out irregularly, i.e., nonuniform sampling (NUS) is applied, the conventional Fourier transform cannot be used for spectrum processing.

We give a simplified explanation of how “compressed sensing” (CS), a family of the spectrum reconstruction algorithms that processes NUS data, works. In general, these methods can find the highest peaks in a spectrum but fail to reconstruct data points at

the noise level. Using NUS and CS, as a rule of thumb the time requirement of the 2D NMR measurements decreases from the order of ten minutes to a few minutes if the signal-to-noise ratio is adequate, i.e., the sample is not too dilute.

Although there are several other novel approaches to accelerating data acquisition, NUS has become particularly popular in recent years because spectrometer manufacturers have implemented the newest NUS acquisition and processing algorithms in the standard spectrometer control software.

In the last section of the paper we present the application of NUS in three interesting structure elucidation problems coming from our own practice. One of these involves an unexpected vindoline trimer molecule whose structure could be solved only by recording 2D NMR spectra in which the resolution in the ^{13}C dimension was extremely high. For example, the HMBC spectrum was recorded using 8,192 t_1 increments (ca. 0.025 ppm/row for a 200-ppm spectral width) from which only 512 increments were sampled. Therefore, the measurement time was reduced by ca. 94% and the spectrum could be acquired within one hour. Without using NUS, the structure elucidation of the vindoline trimer would have been exceedingly time-consuming and cumbersome.

Our next example of the application of NUS is about solving a challenging structure-verification problem. We expected a compound giving two signal sets due to the presence of amide rotamers, but instead four signal sets appeared in the NMR spectra. Hence, we could not decide whether the sample was a mixture of two constitutional isomers or a single compound that had four conformers. Using an in-house-modified NOESY pulse sequence combined with NUS, we were able to see the chemical exchange between four conformers, proving that the sample was uniform in terms of the analyte’s constitution. Our modification of the NOESY sequence involved the incorporation of a constant-time element that simplified the multiplet structure of the peaks in the indirect ^1H dimension. Accordingly, the number of the peaks decreased, and their intensity increased, facilitating the reconstruction of the NUS dataset by CS algorithms.

In the last example we show that NUS can be used and is useful even if the dynamic range in the spectrum is as high as ca. $10^3:1$, revealing peaks in the 2D NMR spectrum that belong to impurities whose concentration is as low as 0.14% relative to the main component. In this case we took advantage of the fact that the t_1 noise becomes ill-reproduced in the NUS spectra. We were able to observe some HMBC correlations for the trace impurity in the NUS spectrum that partially overlapped with the t_1 noise in the traditionally sampled spectrum and could therefore not be noticed there. Using a high-end NMR-spectrometer is essential to acquire such high-quality NMR spectra.

In conclusion, we recommend a more extensive use of 2D NMR spectroscopy in order to gain higher structural confidence in general. In our experience NUS is a convenient and extremely useful method for accelerating the acquisition of the 2D NMR spectra of NMR samples that are not extremely dilute, both in routine and non-routine applications