

## A Beilstein-kézikönyv számítógépesítése

A *Beilstein Handbook of Organic Chemistry*, amely több mint 1,5 millió szerves vegyület szerkezeti, fizikai és kémiai tulajdonságaira és előállítására vonatkozó információkat tartalmaz, a legrégebb és legismertebb kézikönyv a kémia területén. A vegyészek által közzétett Chemical Abstracts (CA) és a Beilstein-kézikönyv között alapvető különbség, hogy a Chemical Abstracts Service (CAS) intézet – a CA-referálólappal készítője – csak kivonatolja és indexeli a kémiai tárgyú publikációkat, a Beilstein-kötetekbe viszont a szerves vegyületekre vonatkozó, a Beilstein Institute intézetnél dolgozó több mint száz vegyész által tudományosan értékelt, ellenőrzött, megbízható tényadatok kerülnek be. A Beilstein-kézikönyv 1830-tól kezdve dolgozza fel a szakirodalmat, míg a CA csak 1907-től. A két kiadvány azonban jól kiegészíti egymást, mivel a Beilstein-kézikönyvbe csak hosszú átfutási idővel, évek múlva kerülnek be az új eredmények, a legfrissebb publikációkról csak a hetente megjelenő CA-füzetekből szerezhetnek tudomást a vegyészkutatók.

### A Beilstein-kézikönyv használatát elősegítő programok

A több mint 300 kötetből álló kézikönyvsorozat használatát megnehezíti, hogy nem könnyű megtalálni egy-egy vegyületre vonatkozó leírás helyét a nagyszámú kötet között, főleg ha csak alkalmanként keresőkről van szó. Az eligazodás megkönnyítésére a Beilstein Institute már korábban kidolgozott egy számítógépes programot, a *Beilstein Keyt*, amely egy sor kérdésre adott válasz alapján elvezet ahhoz a kötethez, amelyben az adott vegyületre vonatkozó leírás található. A kötetszám birtokában az adott kötet kumulatív vagy az egyes kötetek képletmutatójából már könnyen megkereshető a vegyület pontos helye. Nemrégiben kereskedelmi forgalomba került a *SANDRA* (Structure and Reference Analyzer) nevű, IBM PC mikroszámítógépre készült program, amely lehetővé teszi, hogy felrajzoljuk a képernyőre a bennünket érdeklő vegyület szerkezeti képletét, majd ennek alapján néhány másodperc alatt 40 oldalnyi vagy még ennél is nagyobb pontossággal megadja a keresett vegyület helyét a kézikönyvben. A legutóbbi kiegészítő kötetekre azonban már kisebb a pontosság (80–100 oldalnyi).

### A Beilstein számítógépes adatbázis létrehozását megelőző piackutatás

1983-ban a Beilstein Institute elhatározta, hogy számítógépre viszi a kézikönyv anyagát és online hozzáférhetővé teszi. Első lépésként széles körű felmérést végeztek, hogy megismerjék a szerves kémiai információk iránti igényeket és azok beszerzésének módját. A világ különböző részein dolgozó kutatóknak és információs szakembereknek kiküldött 21 767 kérdőívből 4007 (18,41%) érkezett vissza. A visszaküldött kérdőívek alapján a vizsgált minta reprezentatívnak tekinthető. A válaszadók 86,5%-a a szerves kémia, 49,8%-a a biokémia, 40,9%-a a farmakológia, 31%-a a toxikológia és 30,6%-a a polimerkémia területén végez rendszeres információkeresést. A válaszadók 80,5%-a kutató és oktató, 7,5%-a információs szakember és 4,6%-a egyéb foglalkozású.

#### A felmérés eredménye

1. Igény van a rendszerezett fizikai és kémiai tulajdonságadatokat tartalmazó gyűjteményekre, a jelenlegi információforrások nem kielégítőek. A válaszadók 56,9%-a jelezte a tényadatokat tartalmazó adatbázisok iránti igényt. Az online elérhető adatbázisok mindössze 20,8% szerint tartalmaznak elegendő fizikai-kémiai tényadatot.

2. A másodlagos kémiai információforrások használatában a CA vezet (87,5%), viszont meglepően magas a Beilstein használatának aránya (48,3%). Részesedése még az USA-ban is – a nyelvi nehézségek ellenére – 46,2%-os. Az elektronikus úton elérhető információk piacát egyértelműen a CAS adatbázisok uralják.

3. Meglepően nagy a primer szakirodalom szerepe a specifikus információk megszerzésében. A fizikai adatokat a megkérdezettek 54,2%-a közvetlenül a primer forrásból keresi, de mégha a másodlagos információforrásban meg is található a keresett adat, annak helyességét a legtöbb vegyész ellenőrzi az eredeti publikációban.

4. A kémiai információk elektronikus úton való keresése egyre elterjedtebb, a megkérdezettek 68,7%-ának van terminálja, ill. személyi számítógépe, vagy tervezi azok beszerzését külső adatbázisok elérésére. A megkérdezettek jelentős hányada az adatbázisok előnyét a gyorsaságban és a nagy információmennyiségben látja, ugyanakkor drágának találja őket, és nincs megelégedve sem az elérhető pontossággal, sem a teljességgel.

5. A vegyészek leginkább az összegképlet és a szisztematikus kémiai elnevezések alapján keresnek, míg az információs szakemberek előnyben részesítik a kulcsszavakkal vagy a CAS Registry Numbrakkal való keresést. A kifinomultabb keresési lehetőségek – szubstruktúra és Markush-képlet keresés – még nem terjedtek el széles körben.

6. A specifikus információk keresése általában több lépcsőben, különböző információforrások igénybevételével történik. A megkérdezett vegyészek 40%-a a keresést nyomtatott tájékoztató kiadványokban, kézikönyvekben kezdi, az információs szakemberek 52%-a viszont először az online információszolgáltató központokhoz fordul. Harmadik információforrásként gyakori a belső, nem elektronikus információs rendszerek használata. Érdekes tendencia, hogy mégha valamelyik információforrásban sikerült is megtalálni a kívánt információt, a teljes biztonság kedvéért mind a vegyészek, mind az információs szakemberek elvégzik a keresést valamennyi rendelkezésre álló információforrásban.

7. A felmérésből kitűnt, hogy a vegyészutatók információs igényeik nagyobbik részét a kémiai vagy valamilyen természettudományi alapképzéssel és megfelelő információtechnikai ismeretekkel is rendelkező, információszolgáltatásra specializálódott szakembereken keresztül elégítik ki. A megkérdezett vegyészek mindössze 16,4%-a végez maga rendszeresen online kereséseket. Ennek az az oka, hogy a kémiai szakirodalom az egyik legkomplexebb és a legszélesebb körre kiterjedő információgyűjtemény, és már a hagyományos úton való hatékony keresés is megfelelő szakismereteket igényelt. Az információkeresést végzők iránti követelmények az online adatbázisok és az új keresési lehetőségek megjelenésével és elterjedésével tovább növekedtek.

## A Beilstein Online

A felmérés tapasztalatainak figyelembevételével beindult intenzív fejlesztőmunka eredményeképpen 1988-ban kereskedelmi forgalomba kerül a Beilstein-adatbázis: A Beilstein Online. Elsőként a Dialog szolgáltatóközponton lesz elérhető, majd az STN-hálózaton is, de tárgyalások folynak más szolgáltatóközpontokkal is.

A Beilstein Online adatbázisnak két különálló állománya lesz: a vegyületeket nyilvántartó *Beilstein Structure* és a tényadatokat tartalmazó *Beilstein Factual Data* állomány. A Beilstein Structure az online hozzáférhetőség mellett mágnesszalagon is megvásárolható lesz. Az online keresés díját még nem ha-

tározták meg, de azt már eldöntötték, hogy a nyomtatott változat előfizetői kedvezményt kapnak.

A szerkezeti állomány keresésére vagy a DARC, vagy a HTSS, vagy a Beilstein Institute által kidolgozott kémiai szerkezetkereső szoftvert fogják alkalmazni. A Dialog rendszeren a Beilstein adatbázis bevezetésével nyílik majd először lehetőség grafikus vegyületkeresésre.

A tényadatokat tartalmazó faktografikus állomány a Dialog vagy a Beilstein Institute által kifejlesztett szoftverrel lesz kereshető. A faktografikus állomány két részből fog állni:

- ◆ az 1830–1979 közötti, ellenőrzött, értékelt adatok és
- ◆ az 1980-tól kezdve publikált, nem ellenőrzött adatok, amelyek csak online lesznek hozzáférhetőek, majd kritikai elemzésük után az értékelt adat kerül a helyükre. Így a Beilstein-adatbázis sokkal naprakészebb lesz, mint nyomtatott változata.

A vegyületek két csoportja lesz megkülönböztethető. Az egyik csoportba a nagyon fontos, elterjedten használt vegyületek (a Beilsteinben található összes vegyület 5%-a) tartoznak, amelyekre nagyon sok információ áll rendelkezésre (Large Information Compound = LIC). Ezeknek az információknak csak egy része lesz online elérhető. Azokra a vegyületekre, amelyekre kevesebb információ áll rendelkezésre (Small Information Compounds = SIC), a nyomtatott változatban szereplő valamennyi információ hozzáférhető lesz.

## A szerkezeti állomány

A Beilstein Structure File (Beilstein Registry Connection Table = BRCT, kapcsolódási táblázatok nyilvántartása) mintegy 1,5 millió vegyület teljes szerkezeti ábrázolását fogja tartalmazni, beleértve a sztereokémiai jellegzetességeket is, ellentétben a CAS Registry File kétdimenziós ábrázolásával. A másik eltérés a CAS adatbázistól, hogy míg a CAS Registry Numbrak nincsenek semmilyen összefüggésben a vegyületek szerkezetével, a Beilstein által adott nyilvántartási számok (ún. Lawson-számok) "értelmes számok", amelyek kapcsolatba hozhatók a szerkezeti képlettel.

A vegyületek szerkezetét olyan paraméterekkel tartják nyilván, mint pl. a szénkötések  $\pi$ -vegyértékelektronjainak száma, az atomok kapcsolódása, a sztereoatomok és sztereokötések, a lokalizált hidrogének, a normálistól eltérő vegyértékek és atomtömegek, a lokalizált és delokalizált töltések, páratlan vegyértékelektronok (gyökök), hidrogénizotópok, tautomer csoportok stb.

A Beilstein Factual File mintegy 7,5 millió faktografikus rakordot fog tartalmazni 1830-ig visszamenően, az adatbázisban szereplő szerves vegyületek kb. 400 tulajdonságára vonatkozó információval. A 70 legfontosabb tulajdonságra a számszerű érték szerepel, a többire csak irodalmi hivatkozás található. A fizikai és kémiai tulajdonságadatokat hierarchikus elrendezésben találhatók, így a 4.2.6 osztály tartalmazza a spektrumokra vonatkozó információkat, ezen belül a 4.2.6.1 az NMR-re, a 4.2.6.1.1 ezen belül meghatározott atommagra, a 4.2.6.2 pedig az ESR-re vonatkozik. A Beilstein Factual File fő tulajdonságosztályai a következők:

1. Elsődleges adatok
2. Alaki adatok
3. Előállításra vonatkozó adatok
4. Fizikai tulajdonságok
  - 4.1 Szerkezeti és energiajellemzők
    - 4.1.1 Konformáció
    - 4.1.2 Vázjellemzők
    - 4.1.3 Elektromos momentum
    - 4.1.4 Elektromos polarizálhatóság
    - 4.1.5 Kapcsolási jelenségek
    - 4.1.6 Molekuladeformáció és molekulapotenciálok
    - 4.1.7 Molekula-energiaparaméterek
  - 4.2 Tiszta vegyületek fizikai tulajdonságai
    - 4.2.1 Fizikai állapot
      - 4.2.1.1 Kristályok
      - 4.2.1.2 Folyadékok
      - 4.2.1.3 Gázok
    - 4.2.2 Egyéb fizikai és mechanikai tulajdonságok
    - 4.2.3 Transzportjelenségek
    - 4.2.4 Hőadatok
    - 4.2.5 Optikai tulajdonságok
    - 4.2.6 Spektrumok
      - 4.2.6.1 NMR
      - 4.2.6.2 ESR
      - 4.2.6.3 NQR
      - 4.2.6.4 Rotációs spektrumok
      - 4.2.6.5 Vibrációs spektrumok
      - 4.2.6.6 Elektronspektrumok
      - 4.2.6.7 Emissziós spektrumok
      - 4.2.6.8 Egyéb spektroszkópiai módszerek
    - 4.2.7 Mágneses tulajdonságok
    - 4.2.8 Elektromos tulajdonságok

- 4.2.9 Elektrokémiai viselkedés
- 4.2.10 Egyéb tulajdonságok
- 4.3 Többkomponensű rendszerek fizikai tulajdonságai
  - 4.3.1 Oldatok viselkedése
  - 4.3.2 Folyadék/folyadék rendszerek
  - 4.3.3 Folyadék/szilárd rendszerek
  - 4.3.4 Folyadék/gőz rendszerek
  - 4.3.5 Egyéb mechanikai tulajdonságok
  - 4.3.6 Transzportjelenségek
  - 4.3.7 Energiaadatok
  - 4.3.8 Határfelületi jelenségek
  - 4.3.9 Adszorpció
  - 4.3.10 Asszociáció
5. Kémiai viselkedés
6. Fiziológiai viselkedés és alkalmazások
7. Származékok jellemzése

A faktografikus állomány fő felhasználási lehetőségei:

- ◆ egy adott vegyületre vonatkozó valamennyi információ megkeresése,
- ◆ meghatározott tulajdonsággal rendelkező valamennyi vegyület megkeresése,
- ◆ adatok összehasonlítása, tulajdonságadatok kiszámítása, interpolálása, extrapolálása az adatbázisban meglévő adatok alapján (e lehetőség megtevése csak később kerül sor).

A Beilstein Online valószínűleg az egyik leginkább keresett online adatbázis lesz, elsősorban a vegyészek körében.

/HELLER, S. R.: *New access to data from the Beilstein Institute: Beilstein Online and SANDRA.* = Database, 10. köt. 4. sz. 1987. p. 47–52.

WOLMAN, Y.: SANDRA (Structure and Reference Analyzer). = *Journal of Chemical Information and Computer Sciences*, 27. köt. 3. sz. 1987. p. 144–145.

JOCHUM, C.—MORICZ, P.: *Information needs and use in organic chemistry: a worldwide survey from the Beilstein Institute.* = Database, 10. köt. 4. sz. 1987. p. 41–46./

(Novák Teréz)