

imenje i nazivlje u kemiji i kemijskom inženjerstvu

Uređuje: Marija Kaštelan-Macan

Kako isplivati iz mora podataka, pravila i preporuka

L. Varga-Defterdarović

Zavod za organsku kemiju i biokemiju, Institut Ruđer Bošković,
Bijenička cesta 54, 10 000 Zagreb

Već u osamnaestom stoljeću francuski kemičari Guyton de Morveau (Dijon, 1737. – Pariz, 1816.), Antoine-Laurent de Lavoisier (Pariz, 1743. – Pariz, 1794.), Claude Louis Berthollet (Talloires, 1748. – Arcueil, 1822.) i Antoine François de Fourcroy (Pariz, 1755. – Pariz, 1809.) suočili su se sa sve većim brojem otkrivenih i sintetiziranih kemijskih spojeva i potrebom njihova imenovanja, pa su već 1787. godine o tome sastavili i objavili prva pravila.¹ Neka od tada preporučenih imena još i danas su u upotrebi i predstavljaju trivijalna imena. Švedski kemičar Jøns Jakob Berzelius (Väversunda, 1779. – Stockholm, 1848.) uveo je, 1813. godine, simbole za kemijske elemente temeljene na njihovim latinskim imenima (npr. simbol "Fe" od lat. *ferrum*), a omjer elemenata označio je brojem u eksponentu (npr. H₂O). Simboli kemijskih elemenata zadržani su do danas s iznimkom pisanja njihova omjera u indeksu (npr. H₂O). Sažeti pregled nastanka i načina pisanja konstitucijskih formula od vremena Berzeliusa, pa sve do današnjih dana objavljen je u časopisu KUI.²

Od tada se s više ili manje entuzijazma stalno radi na normiranju nomenklature (imenovanja) i grafičkog prikaza (strukturne formule) kemijskih spojeva. Posljednjih osamdesetak godina obilježila su Ženevska i Lieška pravila (*Geneve rules* i *Liege rules*, 1882. i 1939.), nekoliko izdanja *Plave knjige* (*Blue book*) i to 1957., 1965., 1971. i 1979. te *Vodiča kroz IUPAC-ovu nomenklaturu organskih spojeva* (*A Guide to IUPAC Nomenclature of Organic Compounds*) 1993. i to sve radom u okviru međunarodne Komisije za nomenklaturu (imenovanje) organske kemije (*The Commission on the Nomenclature of Organic Chemistry, CNOC*), danas sastavnice Međunarodne unije za čistu i primijenjenu kemiju (*The International Union of Pure and Applied Chemistry, IUPAC*).³ Izvornici IUPAC-ovih preporuka objavljuju se na engleskome jeziku kao zasebne knjige ili kao članci u IUPAC-ovu časopisu *Pure and Applied Chemistry (PAC)*,⁴ a njihove prijevode i prilagodbu različitim jezicima, u svojim zemljama, provode nacionalna tijela (sekcije, odbori...) zadužena za imenovanje kemijskih spojeva.⁵

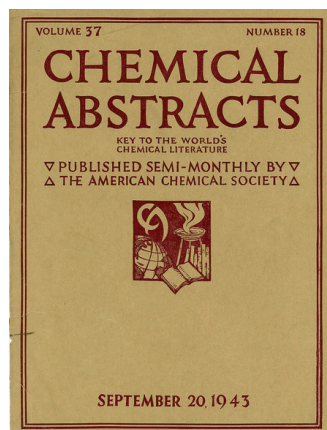
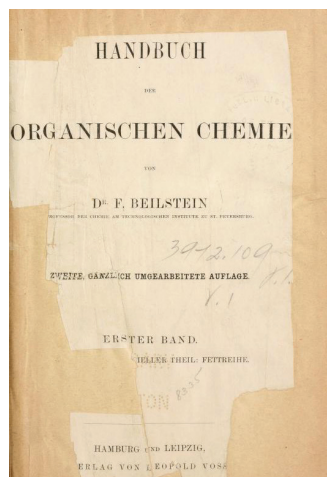
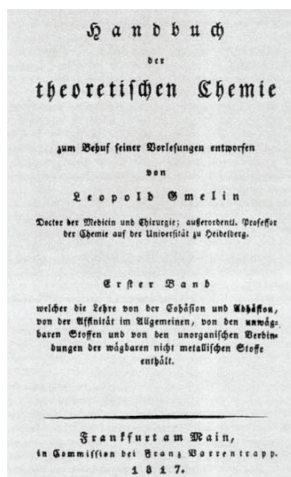
Razvoj prirodnih znanosti, pa tako i kemije, doveo je do novih spoznaja i ogromnog broja novosintetiziranih spojeva, što postaje i temom raznovrsnih publikacija. Pridržavanjem vrlo opsežnih IUPAC-ovih pravila za imenovanje organskih spojeva, otkrivenim i sintetiziranim spojevima moguće je, ovisno o vrsti primijenjene IUPAC-ove nomenklature, pridružiti nekoliko potpuno ispravnih imena,⁶ što može dovesti do nedoumica kod drugih kemičara, jer relativno mali broj njih u potpunosti je ovladao svim do sada donesenim nomenklaturnim pravilima. Pri tome je bitno da, bez obzira koji tip nomenklature je primijenjen, sva ispravna imena moraju predstavljati jednu te istu strukturnu formulu.

Kemija je bila prva prirodna znanost koja je sistematizirala pristup znanstvenim informacijama. Prvi korak napravio je Leopold Gmelin (Göttingen, 1788. – Heidelberg, 1853.) kada je 1817. u Frankfurtu na Majni počeo objavljivati *Handbuch der Theoretischen Chemie* s ciljem da se sakupe, kritički razmotre i objave sve značajne spoznaje (kemijske i fizikalne) o kemijskim spojevima. Potrebno je naglasiti da u to vrijeme još nije postojala podjela na anorgansku i organsku kemiju. Uskoro se broj podataka toliko povećao da je priprema sveobuhvatnih monografija o pojedinom spoju postala toliko zahtjevna da je od petog izdanja, objavljenog 1852. – 1853., obuhvat bio ograničen samo na anorganske spojeve i tako nastaje *Gmelin's Handbuch der Anorganischen Chemie*. Nastalu prazninu popunio je 1881. godine Friedrich Konrad Beilstein (Saint Petersburg, 1838. – Saint Petersburg, 1906.) objavljivanjem prvog izdanja *Handbuch der Organischen Chemie*, koje je tiskano u dvije knjige. Tiskano izdanje *Gmelin's Handbuch der Anorganischen Chemie* prestalo je izlaziti 1997. godine. U 116 godina izlaženja objavljeno je 760 svezaka s oko 240 000 stranica i zajedničko kazalo (*Gmelin Formula Index*) u 35 svezaka. Danas je *Gmelin's Handbuch der Organischen Chemie* uključen u bazu podataka *Reaxys*.⁷ *Gmelin's Handbuch der Organischen Chemie* dugo je vremena imao odlučujući utjecaj na sustavnu nomenklaturu anorganskih kemijskih spojeva.

Handbuch der Organischen Chemie, koju je 1881. utemeljio F. K. Beilstein (*Beilstein's Handbook of Organic Chemistry*) daje pregled organskih spojeva sa stanovišta njihove konstitucije, konfiguracije, konformacije, dobivanja, pročišćavanja, termodinamičkih, fizikalnih i kemijskih svojstava, karakterizacije, analitike i derivata. Tiskano izdanje prestaje izlaziti 1998. godine. U 117 godina izlaženja objavljena su 503 sveska na 440 814 stranica. Do kraja 2008. godine kritički je razmotreno preko dva milijuna originalnih publikacija te je opisano preko deset milijuna struktura i deset milijuna reakcija s literaturnim navodima. Ovi podaci prikupljeni su pregledom znanstvene literature od 1771. godine. Od 2009. godine *Handbuch der organischen Chemie* uključen je u Elsevierovu bazu *Reaxys*, nasljednicu *CrossFire Database*,⁷ kao *Beilstein Database*, koja je najveća baza specijalizirana za organsku kemiju (s preko sedam milijuna spojeva), u kojoj se svakom novom organskom spoju dodjeljuje jedinstven *Beilsteinov registracijski broj* (*Beilstein Registry Number*).

Kako je sakupljanje i kritičko razmatranje sakupljenih podataka zahtijevalo određeno vrijeme, a kemijska se znanost početkom devetnaestog stoljeća vrlo brzo razvijala, 1830. godine je u Leipzigu pokrenuto izdavanje *Pharmaceutisches Central Blatta* (od 1856. izlazi pod imenom *Chemisches Zentralblatt*), koji objavljuje sažetke značajnih publikacija iz područja kemije. Objavljivanje sažetaka publikacija, a ne i njihova kritičkog prikaza u okviru sveobuhvatne monografije čitavog područja, omogućilo je mnogo veću ažurnost u objavljivanju, ali je na korisniku ostalo da prikupi podatke i provede njihovu kritičku analizu. *Chemisches Zentral-*

* Dr. sc. Lidija Varga-Defterdarović
e-pošta: lidija@irb.hr



Slika 1 – Naslovnice nekih od prvih svezaka zbirki monografija kemijskih spojeva Gmelin's Handbuch der Anorganischen Chemie, Handbuch der Organischen Chemie (Beilstein's Handbook of Organic Chemistry) i sažetaka članaka *Chemisches Zentralblatt*, *Chemical Abstract* koji se mogu pregledavati na svjetskoj mreži (internetu)¹³

blatt prestaje izlaziti 1969. godine. Tijekom 140 godina izlazenja *Chemisches Zentralblatt* objavio je više od 900 000 stranica od kojih je 700 000 sadržavalo oko dva milijuna sažetaka međunarodnih publikacija iz područja kemije, dok se abecedni registar protezao na oko 200 000 stranica. Jedinствен sadržaj *Chemisches Zentralblatta* predstavlja nezaobilazan i neprocjenjiv izvor informacija iz područja kemije od početka 19. stoljeća do četrdesetih godina dvadesetog stoljeća. Oko milijun imena kemijskih spojeva i preko 500 000 kemijskih struktura digitalizirano je i čine *Chemisches Zentralblatt Structural Database*.⁸

Godine 1907. u sklopu Američkog kemijskog društva (*American Chemical Society*) počinje izlaziti *Chemical Abstract* (CA), zbirka sažetaka objavljenih članaka iz raznih područja kemije, koja 1956. godine prerasta u *Chemical Abstract Service* (CAS) te u tiskanom obliku izlazi sve do 2009. Danas se kemijski spojevi, prema imenima i strukturama, mogu elektronički pretraživati u bazi *CAS Registry* (najvećoj bazi, utemeljenoj 1964. godine, koja danas broji više od 75 milijuna kemijskih spojeva i 60 milijuna sekvencija),⁹ a objavljeni znanstveni radovi i patenti u bazi podataka *CAPLUS* (najveća i jedinstvena baza s više od 30 milijuna objavljenih članaka i patenata iz različitih područja kemije).¹⁰ Relevantni podaci o nekom spoju mogu se dobiti preko CAS broja (*CAS Registry Number*), međunarodno prihvaćenog broja, na mrežnim stranicama kao što su npr. *National Institute of Standards and Technology* (NIST),¹¹ katalozi dobavljača kemikalija,¹²

različite baze i drugo. CAS broj čine tri skupine brojeva, međusobno odijeljenih spojnicom, koji ne daju podatke o strukturi ili svojstvima spoja, već se dodjeljuju redosljedom prvog spominjanja u CA-u (što je prva skupina veća to je spoj novijeg datuma). Treća skupina brojeva je jednodomenkasti kontrolni broj, koji se dobiva po posebnom algoritmu iz prva dva broja.

Ovdje valja spomenuti da su baze *Beilstein Database* i *Chemical Abstract Service* (CAS) u određenoj mjeri ublažile problem više jednakovrijednih IUPAC-ovih imena uvođenjem "svoje sistematizirane nomenklature, odnosno CAS nomenklature" kao i jedinstvenih registracijskih brojeva za svaki kemijski spoj.

U međusobnoj komunikaciji kemičari diljem svijeta spojeve opisuju imenom dobivenim primjenom dogovorenih pravila, no daleko jednostavniji način je prikaz strukturnom formulom. Važnost pravilnog grafičkog prikaza uočio je i IUPAC te su na tu temu izdane dvije preporuke.^{14,15}

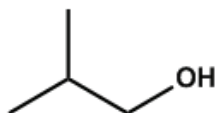
Računala kakva danas poznajemo započiju razvoj tridesetih godina prošlog stoljeća i ubrzo se otkrivaju mogućnosti njihove primjene u raznim područjima ljudskog djelovanja, pa tako i u kemiji, među kojima je i ostvarenje sna o pretvaranju IUPAC-ovih imena kemijskih spojeva u kemijsku formulu i obrnuto, koje započinje algoritmom E. Garfielda, objavljenog 1962. godine.¹⁶ To je i vrijeme sve većeg broja (danas je to preko tisuću dnevno) otkrivenih i sintetiziranih spojeva koje je trebalo imenovati i unijeti u baze *Beilstein* i CA. Upravo to je u *Beilstein Institutu*, 1987. godine, pokrenulo osmišljavanje vlastitog računalnog programa za prepoznavanje grafičkog prikaza kemijskog spoja te imenovanje prema IUPAC-ovoj nomenklaturi, odnosno pretvaranja grafičkoga prikaza u ime i obrnuto, koji je završen 1991. godine pod imenom *AutoNom* (od *Automatic Nomenclature*). Paralelno se i drugdje radilo na računalnim programima za tu namjenu, a primjeri su *BASIC*,¹⁷ *LISP*,¹⁸ te *ACD/Name* i *ACD/ChemSketch*, tada novoosnovane kanadske tvrtke *Advanced Chemistry Development (ACD/Labs)*.¹⁹ *AutoNom* je, sa svojom točnošću oko 85 %, bio vrlo popularan krajem 90-ih godina prošlog i početkom 21. stoljeća. Ostali važniji računalni programi bili su ili jesu *Nemelt* i *DrawIt*,²⁰ *Nomenclator*,²¹ *OPSIN*²² te *ChemDraw* koji od 2011. s *Chem3D* i *ChemFinder* čini paket *ChemOffice*.²³

Sveukupna tiskana izdanja iz područja kemije (znanstveni radovi, patenti, kemijski katalozi, pravna regulativa, udžbenici i dr.), objavljuju se na mnogim svjetskim jezicima, a rad računalnih programa za pretvaranje strukture u ime ili obrnuto temelji se na engleskome, što neminovno podrazumijeva ugradnju programa za prevođenje na druge svjetske jezike i s drugih svjetskih jezika na engleski. Računalni program *Lexichem* sastoji se od tri aplikacije i to za pretvorbu imena u strukturu (*nam2mol*), strukture u ime (*mol2nam*) i prijevod s drugih jezika na engleski i obrnuto (*translate*).²⁴

Upravo nejasnoće i nedoumice do kojih dovodi mogućnost imenovanja kemijskih spojeva s više ispravnih IUPAC-ovih imena (npr. struktura Me_2CO može se imenovati kao acetone, dimetilketon, propanon, pa i pleonazmima 2-propanon ili propan-2-on) nagnala je ovu organizaciju na uvođenje preporučenoga IUPAC-ova imena (*Preferred IUPAC Name, PIN*).²⁵ Danas IUPAC radi na svom globalnom projektu uspostavi jedinstvene oznake nazvane *InChI (International Chemical Identifier)*.²⁶ Algoritam za *InChI* temelji se na nizu IUPAC-ovih preporuka prikaza strukture i pravila za normiranje i kanonizaciju prikaza strukture kako bi se dobila jedinstvena oznaka. Ovo omogućuje računalno pretvaranje grafičkog prikaza kemijskoga spoja u jedinstvenu notaciju *InChI*, koja će se moći ugraditi u bilo koji program za crtanje strukture te biti generirana iz bilo koje baze kemijskih struktura. Kako je *InChI* jedinstvena oznaka svakog pojedinog spoja, duljina mu ovisi o formuli, odnosno kompleksnosti toga spoja, te *InChI* može sadržavati i do 1000 znakova. Ovako velika brojnost znakova onemogućuje pretraživanje svjetskih mreža, pa je stoga osmišljen algoritam koji precizno dijeli notaciju *InChI* po fra-

gmentima. Time je dobivena njegova sažeta notacija nazvana *InChIKey*, koju definiraju kombinacije 27 slova. Notacije *InChI* i *InChIKey* prihvatili su mnogobrojni časopisi i baze, uključujući i velike svjetske baze kao što su *PubChem*²⁷ i *ChemSpider*.²⁸

IUPAC-ov projekt *InChI* započeo je 2000. godine unutar IUPAC-ove *Division VIII (Chemical Nomenclature and Structure Representation Division)* i do danas je postao jedan od najvažnijih međunarodnih standarda prikaza kemijske strukture, no i dalje se radi na uklanjanju uočenih nedostataka i njegovoj nadogradnji i proširenju.²⁹



IUPAC: 2-methylpropan-1-ol

CAS: 78-83-1

Beilstein: 1730878

InChI: *InChI*=1S/C4H10O/c1-4(2)3-5/h4-5H,3H2,1-2H3

InChIKey: ZXEKIIBDNHEJQC-UHFFFAOYSA-N

SMILES: CC(C)CO

Slika 2 – Primjer prikaza kemijskog spoja formulom (grafički) (a), imenom (nomenklatura) (b), registracijskim brojem (c) i notacijom (d).

Iz ovog kratkog prikaza vidljivo je kako dostignuća suvremenih informatičkih tehnologija omogućuju uspješno rukovanje velikim brojem podataka, koji svakodnevno nastaju kao rezultat istraživanja u području kemije. No kao i početkom devetnaestog stoljeća tako je i danas kritičko vrednovanje prikupljenih podataka nužno za njihovu uspješnu primjenu.

Literatura

1. L. B. Guyton de Morveau, A. L. de Lavoisier, C. L. Berthollet i A. F. de Fourcroy, *Méthode de Nomenclature Chimique*, Pariz, 1787.
2. N. Raos, A. Miličević, Načini pisanja konstitucijskih formula, *Kem. Ind.* **61** (9-10) (2012) 435–441.
3. V. Rapić, L. Varga-Defterdarović, Imenje (nomenklatura) i nazivlje (terminologija) organske kemije. I. Šezdeset godina hrvatske organsko-kemijske nomenklature. *Kem. Ind.* **62**(7-8) (2013) 261–270.
4. IUPAC, *Pure and Applied Chemistry*. URL: www.iupac.org/publications/pac/ (30. 11. 2013.).
5. IUPAC, *Nomenclature Books*. URL: www.chem.qmul.ac.uk/iupac/bibliog/books.html (30. 11. 2013.).
6. *A Guide to IUPAC Nomenclature of Organic Compounds* (pripremili: R. Panico, W. H. Powell, J.-C. Richer), Blackwell Scientific Publications, Oxford, 1993.
Hrvatski prijevod: Vodič kroz IUPAC-ovu nomenklaturu organskih spojeva (poglavlje R-1.2), preporuke HKD i HKDI 2001. (urednik prijevoda V. Rapić, prijevod: I. Bregovec, Š. Horvat, K. Majerski, V. Rapić, sveučilišni priručnik), Školska knjiga, Zagreb, 2002.
7. Reaxys, Elsevier. URL: www.elsevier.com/online-tools/reaxys (30. 11. 2013.).
8. *Chemisches Zentralblatt*, InfoChem GmbH. URL: infochem.de/products/databases/czb.shtml (30. 11. 2013.).
9. CAS Registry (CAS Chemical Registry System). URL: www.cas.org/content/chemical-substances (30. 11. 2013.).
10. CAplus. URL: www.cas.org/content/references (30. 11. 2013.).
11. NIST Chemistry WebBook. URL: webbook.nist.gov/chemistry/cas-ser.html (30. 11. 2013.).
12. University of Texas Libraries, Austin, SAD. URL: www.lib.utexas.edu/chem/info/cats.html (30. 11. 2013.).
13. Open Library. URL: openlibrary.org/ (30. 11. 2013.).
14. J. Brecher, Graphical Representation of Stereochemical Configuration (IUPAC Recommendations 2006), *Pure Appl. Chem.* **78** (10) (2006) 1897–1970.
URL: pac.iupac.org/publications/pac/pdf/2006/pdf/7810x1897.pdf (30. 11. 2013.).
15. J. Brecher, Graphical Representation Standards for Chemical Structure Diagrams (IUPAC Recommendations 2008), *Pure Appl. Chem.* **80** (2) (2008) 277–410. URL: pac.iupac.org/publications/pac/pdf/2008/pdf/8002x0277.pdf (30. 11. 2013.).
16. E. Garfield, An algorithm for Translating Chemical Names to Molecular Formulas, *J. Chem. Doc.* **2**(3) (1962) 177–179.
17. D. E. Meyer, S. R. Gould, Microcomputer Generation of Chemical Nomenclature from Graphic Structure Input, *Am. Lab.* **20** (11) (1988) 92–96.
18. K. W. Raymond, A LISP program for the generation of IUPAC names from chemical structures, *J. Chem. Inf. Comput. Sci.* **31** (1991) 270–274.
19. ACD/Name i ACD/ChemSketch, Advanced Chemistry Development, Toronto, Kanada. URL: www.acdlabs.com/products/draw_nom/ (30. 11. 2013.).
20. Namelt i DrawIt, Bio-Rad, SAD. URL: www.bio-rad.com/en-hr/product/iupac-nameit-drawit (30. 11. 2013.).
21. Nomenclator, ChemInnovation Software Inc., San Diego, SAD. URL: www.cheminnovation.com/products/nomenclature.asp (30. 11. 2013.).
22. OPSIN (Open Parser for Systematic IUPAC Nomenclature), University of Cambridge, UK. URL: opsin.ch.cam.ac.uk/ (30. 11. 2013.).
23. ChemDraw, PerkinElmer Informatics, SAD. URL: www.cambridgesoft.com/Ensemble_for_Chemistry/ChemDraw/ (30. 11. 2013.).
24. Lexichem, OpenEye Scientific Software Inc., Santa Fe, SAD. URL: www.eyesopen.com/docs/lexichem/current/pdf/lexichem.pdf (30. 11. 2013.).
25. IUPAC project: Preferred names in the nomenclature of organic compounds. URL: www.iupac.org/nc/home/projects/project-db/project-details.html?tx_wfqbe_pi1%5Bproject_nr%5D=2001-043-1-800 (30. 11. 2013.).
26. IUPAC – International Chemical Identifier. URL: www.iupac.org/nc/home/projects/project-db/project_details.html?tx_wfqbe_pi1%5Bproject_nr%5D=2000-025-1-800 (30. 11. 2013.).
27. PubChem Sketcher. URL: pubchem.ncbi.nlm.nih.gov/edit2/index.html
28. ChemSpider. URL: www.chemspider.com/ (30. 11. 2013.).
29. Project List for Division Chemical Nomenclature and Structure Representation Division. URL: www.iupac.org/nc/home/projects/projects-by-divisions/project-list-for-division.html?tx_wfqbe_pi1%5Bdivision%5D=Chemical%20Nomenclature%20and%20Structure%20Representation%20Division (30. 11. 2013.).