

Multikomponensű reakciók kémiai informatikai értékelése virtuális könyvtárfejlesztéshez

HAVASI Dávid^{a,b,*} és ANGYAL Márton^a

^a Kémiai és Környezeti Folyamatmérnöki Tanszék, Vegyészmérnöki és Biomérnöki Kar,
Budapesti Műszaki és Gazdaságtudományi Egyetem, 1111 Budapest, Műegyetem rkp. 3., Magyarország
^b Mcule.com Kft., Bartók Béla út 105-113., 1115 Budapest, Magyarország

1. Bevezetés

A gyógyszerkutatás korai fázisában a számítógépes módszereknek köszönhetően nagyszámú szerkezet vizsgálatára nyílik lehetőség.¹ Egy virtuális szűrés során akár többmillió molekula is gyorsan vizsgálható, ugyanakkor a jelenleg polcról elérhető vegyületek száma mindössze tízmillió nagyságrendű. Napjainkra a szerves félvezetők kutatásában is megjelentek ezek a módszerek, ami tovább növeli az új vegyületek iránti igényt.² A polcról elérhető és az egyedi szintézissel előállítható vegyületek közötti szállítási idő- és költségbeli különbséget a virtuális vegyületek igyekeznek áthidalni. Ezek jellemzője, hogy elérhető alapanyagokból 1-2 szintetikus lépésben állíthatók elő elterjedten alkalmazott reakciótípusokkal, mint acilezés vagy C-C keresztkapcsolás.³ Az ilyen típusú vegyületekkel szemben elvárás, hogy nagy valószínűséggel elkészíthetők legyenek, valamint alacsony költséggel és szintézisidővel rendelkezzenek. Ezt a vegyületeket létrehozó szabályokkal lehet biztosítani, amelyek azonosítják a reakció szempontjából fontos alszerkezeteket, valamint kiszűrhetik a mellékfolyamatokat.⁴ Ilyen célból elterjedten alkalmazzák a SMARTS⁵ és SMIRKS⁶ reprezentációkat, mivel a reaktivitás mellett a transzformációkat is leírhatják. A jelenleg elérhető virtuális adatbázisok többek közt a Hartenfeller és munkatársai⁷ által gyűjtött robusztus reakció szabályhalmazán alapulnak.⁸ A terek vizsgálata alapján látható, hogy igazán nagyszámú terméket minimum két egymást követő reakcióval lehet elérni.

2. Multikomponensű reakciók

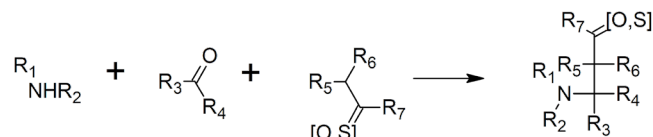
A multikomponensű reakciók (MCR-ek) olyan átalakulások, amelyeknél kettőnél több reaktáns kombinációjával jön létre a termék.⁹ Az MCR-ek különösen előnyösek, mivel egyetlen lépésben lehetővé teszik egy nagyobb kémiai tér gyors és hatékony felépítését. Fontos azonban különbséget tenni az egyes MCR-ek között aszerint, hogy a résztvevő reaktánsok változtathatók-e. Az alkinnek [2+2+2] cikloaddíciója (trimerizáció) nikkkel katalizátorral benzolgyűrűt alkot, azonban három különböző alkinvegyület kombinálására nem alkalmazható egyetlen termék eléréséhez¹⁰ – a reakció jellemzően szimmetrikus vagy intramolekuláris rendszerekben működik szelektíven. A tiofének előállításá-

hoz használható Gewald-reakcióban a-cianoészter, aldehid vagy keton, valamint kén vesz részt.¹¹ Az elemi kénalkalmazása miatt, csak a két további reaktáns lesz változtatható. Hasonló korlátai vannak az Asinger-reakciónak is az ammónia és a kén miatt, valamint a két keton használata is vegyes termékösszetételt eredményez.¹² A Pauson–Khand [2+2+1] cikloaddíció során alkin, alkén és szén-monoxid kapcsolódik össze öttagú gyűrűt zárva, a változtathatóságot a karbonilként beépülő CO korlátozza.¹³ A Strecker-féle aminosav szintézis első lépése kondenzáció aldehid, cianid és ammónia között, amelynek eredményeként egy a-aminonitril alakítható ki.¹⁴ A Hantzsch-féle piridinszintézis során összesen háromféle komponens alkotja meg a termék gyűrűs szerkezetét: ketoészter, aldehid és ammónia.¹⁵ A Bucherer–Bergs-reakció hidantoinok előállítására alkalmas karbonil vegyületekből, ammónia- és cianidforrás segítségével.¹⁶ Látható, hogy a felsorolt reakcióknál legfeljebb kettő komponensük ad változtatható alszerkezetet a termékben.

3. A vizsgálatba bevont multikomponensű reakciók

A következőkben olyan MCR-k rövid bemutatásával folytatjuk, amelyek legalább három variálható reaktánsra épülnek, kitérve a változtatható alszerkezetekre, valamint az egyes reakciókhoz bemutatjuk a SMIRKS reprezentációkat is, amelyek enumerációhoz és keresésekhez használhatók.

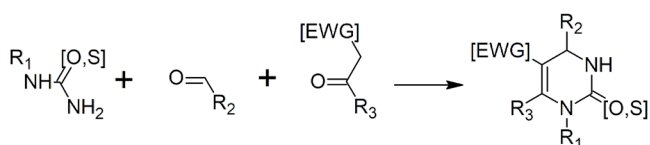
A Mannich-reakció egy vegyes-aldol-reakcióként értelmezhető, amelyben amin, karbonil vegyület és egy CH-savas komponens savas protonja vesz részt^{17,18} – szakirodalmi adatok alapján karbonil és tiokarbonil vegyületek is használhatók.¹⁹ A reakció eredményeként b-amino-karbonil vegyületek alakíthatók, különféle Lewis- és Bronsted-savak jelenlétében.



1. Ábra. A Mannich-reakció vizsgálatban használt sémája.
SMIRKS: [NX3!H0;!\$(*=[O,S,N]):10]-[#6].
[#6;\$([c,CX4H0])]-[CX3H1:20]=[OX1].[#6]-[CX4H2:30]-
[A;\$([CX3]=[OX1,SX1])]>>[NX3:10]-[CX4:20]-[CX4:30]

* Főszerező. Tel.: +36 1 463 3103; e-mail: havasi.david@vbk.bme.hu

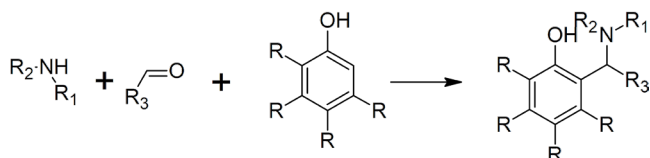
A Biginelli-reakció 3,4-dihidropirimidin-2(1H)-on előállítására alkalmas, eredeti formájában karbamid, aril aldehid és etil acetoacetát kapcsolódásából.²⁰ Újabb kutatásokban sikerült megvalósítani különféle karbamidszármazékok, valamint a ketoészterek mellett ketoszulfonátok alkalmazásával is. Amennyiben nem szimmetrikus az alkalmazott karbamid / tiokarbamid komponens, kétféle termék képződése várható, ugyanakkor ha csak az egyik nitrogén szubsztituált, akkor egyféle termék képződik.^{21,22}



2. Ábra. A Biginelli-reakció vizsgálatban használt sémája (EWG: karbonil és szulfonil funkciók)

SMIRKS: [NX3!H0;!\$([#6]):10]-[CX3R0:11](=[OX1,SX1:12])- [NX3H2:13].[#6;\$([c,CX4H0]):1]-[CX3H1:20]=[OX1].[#6]-[CX3:30] (= [OX1])- [CX4H2:31]-[A;\$([CX3]=[OX1])[OX2H0]),\$([SX4] (= [OX1])(=[OX1])[OX2H0,NX3,#6]),\$([CX2]#[NX1]),\$([NX3] (~[OX1])~[OX1]))>>[NX3:10]1-[CX3:11](=[*:[12])]-[NX3:13]- [CX3:20]=[CX3:30]-[CX4:31]-1

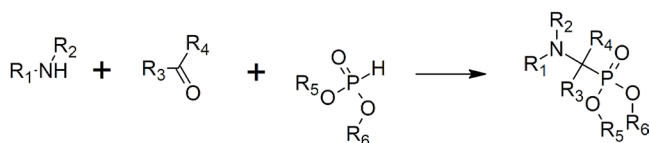
A Betti-reakcióval egy a-aminobenzilfenol alakítható ki egy fenolos komponens, aldehid és amin kapcsolásával.²³ Alapvetően a Mannich-reakció egy speciális esetének tekinthető, a karbonil CH savas hidrogénje helyett a fenol orto helyzetű protonja vesz részt a reakcióban. Jellemzően naftol és fenol típusú vegyületekkel írják le a szintézist, különféle szubsztituációs mintákkal.^{24,25}



3. Ábra. A Betti-reakció vizsgálatban használt sémája.

SMIRKS: [NX3!H0;!\$([*]=[O,S,N]):10]-[#6].[#6;\$([c,CX4H0]):1]- [CX3H1:20]=[OX1].[OX2H1]-[c]([cH1:30])~[cH0]>>[NX3:10]- [CX4:20]-[c:32]

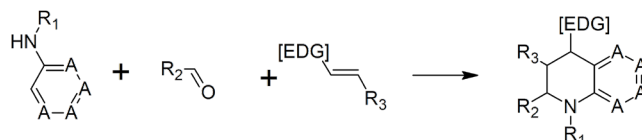
Az a-aminofoszfátok előállítására alkalmas Kabachnik–Fields-reakció során amin, karbonil és egy dialkil foszfonát kapcsolódik össze.^{26,27} A reakció alapvetően egy iminképződésből és az azt követő hidrofoszfonilációból áll, amelyet Lewis-savak és dehidralószerek gyorsíthatnak, ugyanakkor újabb szintetikus eredmények azt mutatják, hogy a reakció mikrohullámú fűtéssel akár oldószer és katalizátor nélkül is kivitelezhető.²⁸ A résztvevő amin és karbonil mellett a foszfonát is kétféle szubsztituenssel rendelkezhet.



4. Ábra. A Kabachnik–Fields-reakció vizsgálatban használt sémája.

SMIRKS: [NX3H2;!\$([*]=[O,S,N]):1]-[#6].[#6]-[CX3R0;\$([CX3H1]),\$([CX3H0])([CX4])([CX4,c]):2]=[OX1].[#6] [PX4H1:3](=[OX1])([OX2H0])[OX2H0]>>[NX3:1]-[CX4:2]-[PX4:3]

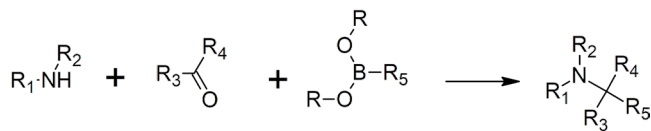
A Povarov- / Grieco-reakció karbonil, amin és elektrondús alkén (gyakran ciklopentadién) reakciójával 6-tagú heterociklusok kialakítására használható.^{29,30} Az alkén esetében fontos az elektrondús csoport jelenléte az egyik oldalon, ez ugyanis meghatározó a reakció szempontjából, valamint biztosítja a szelektív termékképződést.³¹ A résztvevő aminos esetében heteroaromás gyűrűhöz kapcsolódó, valamint szekunder aminos is alkalmasak lehetnek a reakcióra.^{32,33}



5. Ábra. A Povarov- / Grieco-reakció vizsgálatban használt sémája (EDG: az alkén szempontjából fontos, hogy elektrondús szubsztituensek, pl. fenil, vinil, éter, amin, stb.).

SMIRKS: [NX3!H0;!\$([*]=[O,S,N]):10]-[c:11]([cH:12])[aH0].[#6;\$([c,CX4H0]):1]-[CX3H1:20]=[OX1].[A;\$([SX2,NX3,OX2,PX3]),\$([CX3]=[CX3]),\$([c]):1)-[CX3H1:30]=[CX3;\$([CX3H2]),\$([CX3H1] [CX4]):31]>>[NX3:10]1-[c:11][c:12]-[CX4:20]-[CX4:30]-[CX4:31]-1

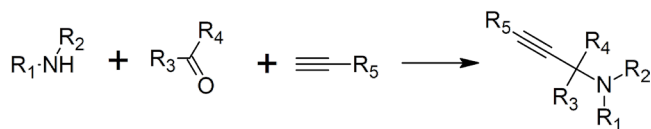
A Petasis-reakció során aromás vagy vinil boronsavak segítségével szubsztituált aminos állíthatók elő primer vagy szekunder aminosból, valamint aldehidekből vagy ketonokból.³⁴ A reakció első leírása óta igazolták, hogy boronsavak mellett boronátokkal és különféle heterociklusos bórvegyületekkel is működik a kapcsolás – katalitikus és katalizátortmentes körülmények között is.^{35,36}



6. Ábra. A Petasis-reakció vizsgálatban használt sémája.

SMIRKS: [NX3!H0;!\$([*]=[O,S,N]):1]-[#6].[#6]-[CX3R0;\$([CX3H1]),\$([CX3H0])([CX4])([CX4,c]):2]=[OX1].[BX3] ([OX2])([OX2])[#6;\$([c,CX3,CX2]):3]>>[NX3:1]-[CX4:2]-[#6:3]

Az aldehidek-aminok-alkinek MCR-ben való katalitikus kombinációját Li és munkatársai írták le először.³⁷ A röviden csak A3 kapcsolásnak nevezett reakció terméke egy propargil amin, ami hasznos építőkö különféle gyűrűs, biológiai és kémiai aktív szerkezet előállításához.³⁸ A reakció körülményeit tekintve átmenetifémek, például réz, ezüst vagy cink vegyületek használhatók katalizátorként.^{39,40}

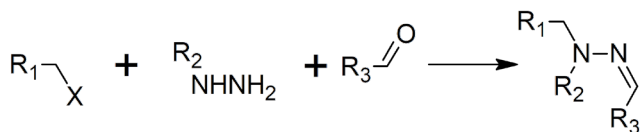


7. Ábra. Az A3 kapcsolási reakció vizsgálatban használt sémája.

SMIRKS: [NX3!H0;!\$([*]=[O,S,N]):1]-[#6].[#6]-[CX3R0;\$([CX3H1]),\$([CX3H0])([CX4])([CX4,c]):2]=[OX1].[#6] [CX2]#[CX2H1:3]>>[NX3:1]-[CX4:2]-[CX2:3]

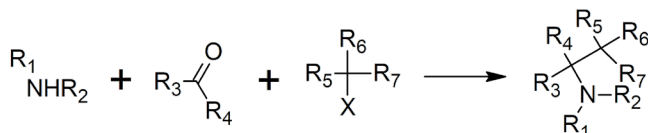
Egy nemrég megjelent munkában Hao és munkatársai triszubsztituált hidrazonok szintézisét mutatták be alkil halogenidek, aminos és aldehidek felhasználásával. A reakció

bázis jelenlétében hajtható végre, a különféle alkil halogenidek közül a kloridok bizonyultak a legreaktívabbnak. A termékként előállítható hidrazonok hasznos intermedierek lehetnek különféle heterociklusok szintéziséhez.⁴¹



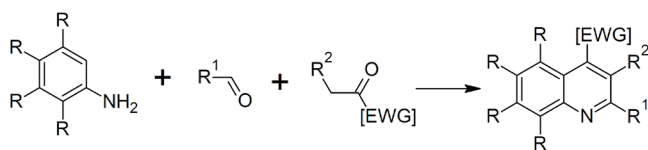
8. Ábra. A Hao-féle reakció vizsgálatban használt sémája, halogenidként fluorid, klorid, bromid és jodid is használható.
SMIRKS: [FX1,CIX1,BrX1,IX1]-[CX4H2:10]-[c].[NX3H2:20]-[NX3H1;!\$(*=[O,S,N]):21]-[#6].[#6;\$([c,CX4H0])]-[CX3H1:30]=[OX1]>>[CX4:10]-[NX3:21]-[NX2:20]=[CX3:30]

Gaunt és munkatársai cink katalizátor vagy látható fény segítségével alfa pozícióban elágazó aminok szintézisét mutatták be karbonil alkilatív aminálási mechanizmussal. A reakcióban hatékonyan alkil jodidok és bromidok bizonyultak megfelelően reaktívnak, ugyanakkor mind primer, szekunder és tercier halogenidekre is mutattak be példákat.^{42,43}



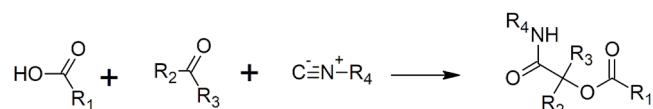
9. Ábra. A Gaunt-féle reakció vizsgálatban használt sémája.
SMIRKS: [NX3!H0;!\$(*=[O,S,N]):1]-[#6].[#6]-[CX3R0;\$([CX3H1]),\$([CX3H0])([CX4])([CX4,c]):2)]=[OX1].[BrX1,IX1]-[CX4:3]>>[NX3:1]-[CX4:2]-[CX4:3]

A Doebner-reakcióban első leírása szerint anilin, egy aldehid, valamint piroszőlősav kondenzációjából kinolinszármazékok állíthatók elő, szervesetlen sav jelenlétében.⁴⁴ Az első publikálása óta rámutattak, hogy az amin és az aldehid komponensek mellett a piroszőlősav is kiváltható változtatható szerkezetekkel, pl. más ketoészterekkel, különféle savkatalizátorok jelenlétében.⁴⁵



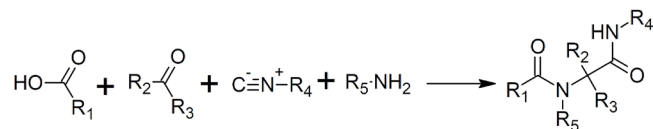
10. Ábra. A Doebner-reakció vizsgálatban használt sémája (EWG: észter és karbonsav funkciók).
SMIRKS: [NX3H2;!\$(*=[O,S,N]):10]-[c;\$([c1cccc1]):11](-[a;!\$([cH]))~[cH:12].[#6;\$([c,CX4H0])]-[CX3H1:20]=[OX1].[CX4!H0&!H1R0:31]-[CX3:30]=[OX1][A;\$([CX3]=[OX1])([OX2H0,OX2H1])>>[NX2:10]1:[c:11]:[c:12]:[c:30]:[c:31]:[c:20]:1

Az izocianidot használó MCR-ek egyike a Passerini-reakció, amelyben karbonsavak és karbonil vegyületek kapcsolhatók, egy a-aciloxi amidot kialakítva.⁴⁶ A reakció különféle karbonilvegyületekkel, köztük gyűrűs ketonokkal is kivitelezhető, limitációja elsősorban az izocianidok elérhetősége.⁴⁷ A szintézisre legnagyobb hatása az oldószernek van.⁴⁸



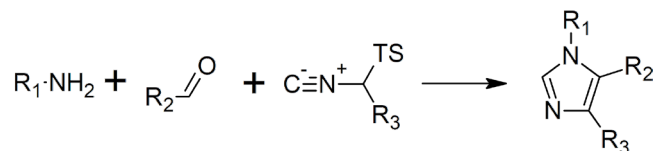
11. Ábra. A Passerini-reakció vizsgálatban használt sémája.
SMIRKS: [#6][CX3]=[OX1][OX2H1:10].[#6]-[CX3R0;\$([CX3H1]),\$([CX3H0])([CX4])([CX4,c]):20)]=[OX1].[CX1:30]~[NX2:31]-[#6]>>[OX2:10]-[CX4:20]-[CX3:30]=[OX1]-[NX3:31]

Az Ugi-reakció ugyanazokat a komponenseket használja, mint a Passerini, kiegészítve egy aminnal, ezáltal egy reakcióban négy változtatható komponenst kombinálva, jellemzően katalizátor nélkül.⁴⁹ Az egyes reaktánsok szubsztituensei széles körben változtathatók, pl. az izocianid esetén terc-butil csoport, a karbonsav esetén különféle aromás gyűrűk is kapcsolódhatnak.⁵⁰



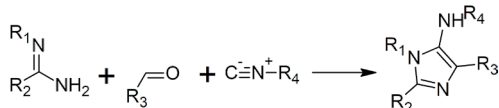
12. Ábra. Az Ugi-reakció vizsgálatban használt sémája.
SMIRKS: [#6][CX3:10]=[OX1][OX2H1].[NX3!H0;!\$(*=[O,S,N]):20]-[#6].[#6]-[CX3R0;\$([CX3H1]),\$([CX3H0])([CX4])([CX4,c]):30)]=[OX1].[CX1:40]~[NX2:41]-[#6]>>[CX3:10]-[NX3:20]-[CX4:30]-[CX3:40]=[OX1]-[NX3:41]

A van Leusen-reakció tozilmetil izocianidok (TosMIC) felhasználásával aldehidek és primer aminok összekapcsolására alkalmas, imidazolgyűrűt zárva.⁵¹ A reakció során a tozil funkció távozik, ugyanakkor a reakció szempontjából egy további szubsztituens (alifás, aromás) megengedhető.⁵² A reakció alapvetően bázis jelenlétében zajlik, és szobahőmérsékleten is végrehajtható.⁵³



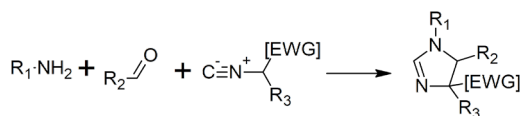
13. Ábra. A van Leusen-reakció vizsgálatban használt sémája (TS: p-toluolszulfonil csoport).
SMIRKS: [NX3H2;!\$(*=[O,S,N]):10]-[#6].[#6]-[CX3H1:20]=[OX1].[CX1:30]~[NX2:31]-[CX4!H0:32]-[SX4;\$(*=(O)=(O)c1ccc(C)cc1)]>>[NX3:10]1:[c:20]:[c:32]:[NX2:31]:[c:30]:1

A savkatalizált Grobbeck–Blackburn–Bienaymé-reakcióban egy izocianid segítségével kondenzált aromás rendszer alakítható ki, savas közegben.^{54,55,56} A reakcióba bevonható alapanyagok körét az izocianid mellett az amin is limitálja, mivel annak egy orto helyzetben nitrogént tartalmazó gyűrűhöz kell kapcsolódnia – viszont nem csak hat, hanem öttagú heterociklusok is bevonhatók. Az aldehidek számos szubsztituenssel használhatók, így a reakció többféle variációra ad lehetőséget.^{57,58}



14. Ábra. A Groebke–Blackburn–Bienaymé-reakció vizsgálatban használt sémája (R_1 és R_2 egy gyűrű részei).
SMIRKS: [NX3H2:10]-[c:11]~[n:12]~[a;!\$([n])].[#6;\$([c,CX4H0])]-[CX3H1:20]=[OX1].[CX1:30]~[NX2:31]-[#6]>>[nX2:10]1:[c:11]:[nX3:12]:[c:30]([NX3H1:31]):[c:20]:1

Az Orru-reakció nagyban hasonlít a van Leusen-reakcióra (amin, aldehid és izocianid kapcsolódása), azonban ebben az esetben az izocianid nem TosMIC típusú, hanem más elektronszívó csoporttal (pl. észter) van ellátva. Ennek köszönhetően a reakció termékeként 2-imidazolinok alakíthatók ki. A szintézis nátrium-szulfát jelenlétében, akár szobahőmérsékleten is végrehajtható.⁵⁹



15. Ábra. Az Orru-reakció vizsgálatban használt sémája (EWG: elektronszívó csoport, pl. észterek).
SMIRKS: [NX3H2:10]-[c:11]~[n:12]~[a;!\$([n])].[#6;\$([c,CX4H0])]-[CX3H1:20]=[OX1].[CX1:30]~[NX2:31]-[CX4H0:32]-[A;\$([CX3]=[OX1])[OX2H0]),\$([SX4]=[OX1])(=[OX1])[OX2H0,NX3,#6]),\$([CX2]#[NX1]),\$([NX3]~[OX1])~[OX1]);!\$([S(=O)(=O)c1ccc(C)cc1])>>[NX3:10]1-[CX4H1:20]-[CX4:32]-[NX2:31]=[CX3H1:30]-1

4. Reakciók termékpotenciáljának értékelése

A virtuális tér bővítése szempontjából fontos, hogy a kialakítandó alszerkezet ritkának számít-e a jelenleg elérhető vegyületek között, valamint hasznosítható-e potenciálisan (pl. nem kerül-e kizárásra gyógyszerkutatási szempontok miatt⁶⁰). A kialakított SMIRKS-ek alapján első körben megvizsgáltuk az Mcule In Stock adatbázisában (6.317.976 szerkezet),⁶¹ hogy az egyes reakciók által előállított termék alszerkezetek milyen számban találhatók meg, az adatokat az 1. Táblázat tartalmazza. Az egyes reakciótípusokkal potenciálisan elérhető kémiai tér méretét közelíteni az alapanyagokra kapott találati számok alapján lehetséges. Ennek feltérképezéséhez az Mcule Fast Delivery Building Blocks Class 2 (317.751 szerkezet) halmazát használtuk (a kiválasztás szempontja az alapkövek könnyű hozzáférhetősége volt).⁶² Izocianidok esetében a kiválasztott halmaz egy szerkezetet sem tartalmazott, így ezeket egy szintetikus kémiai beszállító adatbázisában vizsgáltuk (Key Organics BIONET Building Blocks, 271.148 szerkezet).⁶³ A reakciók segítségével elérhető maximális kombinatorikus tér a találati számok szorzata alapján becsülhető (1. Táblázat), amely segítheti pl. áralapú alapanyag leválogatás kialakítását.⁶⁴

A vizsgálat alapvetően rámutat, hogy egyes reakciók alkalmazását elsősorban a legspeciálisabb reaktáns típus korlátozza, mint amilyen az izocianid (kivéve az Ugi-reakció esetében, ahol a 4. komponens ellensúlyozza ezt)

vagy a Kabachnik–Fields-reakció esetében a foszfonátok. Ezzel szemben az olyan reakciók, mint az A3 kapcsolás, a Mannich- vagy a Petasis-reakció jelentős előnyt élveznek, mivel azok gyakran használt szintetikus építőkövekre támaszkodnak, amelyeket például acilezési, redukív aminálási, Sonogashira- és Suzuki reakciók is hasznosítanak, így nagyobb számban elérhetők. A reakciók segítségével elérhető kémiai tér mérete tovább növelhető olyan reakciók azonosításával, amelyek alkalmasak a szűk keresztmetszetet jelentő alapanyagok kialakítására, pl. az izocianidok előállíthatók primer aminok felhasználásával.⁶⁵ A MCR-ek esetében megfigyelhető, hogy döntő többségük igényel amin és karbonil (aldehid vagy keton) funkciót, a különbség a további komponensekben figyelhető meg. Habár sok esetben még nem vizsgálták teljeskörűen a reakciók mechanizmusát, egyes vizsgálatokban az imin vagy iminium ion képződését tekintik elsődleges lépésnek,^{65,66} így a reakciókat érdemes lehet két lépésben is kezelni enumeráció során (pl. ha az enumerációs környezet csak 2-2 komponens kombinációját engedi meg), első lépésben kialakítva ezt az intermediert.

1. Táblázat. Az egyes reakciók termékeinek alszerkezeti találati, valamint az egyes reakciókkal potenciálisan elérhető kémiai tér mérete (a *-al jelölt reakciók izocianidot igényelnek)

Reakció megnevezése	Termék találati szám	Potenciálisan elérhető termékszám
Mannich	26	2.48×10^{13}
Biginelli	9715	3.07×10^9
Betti	421	5.44×10^{12}
Kabachnik-Fields	409	1.51×10^{10}
Grieco/Povarov	1944	7.40×10^{10}
Petasis	140513	2.28×10^{13}
A3	84	4.46×10^{12}
Hao	105	3.30×10^{10}
Gaunt	20485	9.91×10^{12}
Doebner	11043	1.07×10^{10}
Passerini*	902	2.23×10^{11}
Ugi*	54995	1.99×10^{16}
van Leusen*	4148	1.74×10^{10}
Groebke-Blackburn-Bienayme*	83	1.43×10^{10}
Orru*	0	6.44×10^9

5. Összegzés

A MCR-k lehetővé teszik, hogy egyetlen reakciólépésben, legalább három különféle alapanyag kombinációjával új termékeket állítsunk elő. Virtuális könyvtárak fejlesztése szempontjából ez kiemelten hasznos lehet, mivel a valós szintézis egy lépésben történik, nem igényel köztes tisztítást, szemben egy kétlépéses szintézissel. Jelen munkában

15 különböző MCR-t vizsgáltunk meg virtuális könyvtárépítés szempontjából, amihez minden reakció esetében bemutattuk a keresésekhez és enumerációhoz használható SMARTS / SMIRKS mintázatokat. A reaktánsok SMARTS-ai lehetővé teszik a potenciálisan elérhető virtuális tér méretének becslését, a SMIRKS-ek pedig a tényleges enumerációhoz használhatók. Az elemzés alapján látható, hogy számos reakció olyan nagyszámban rendelkezésre álló vegyületre épít, mint az aminos, aldehidek vagy ketonok, viszont egyes esetekben ritkább vegyületekre van szükség, ami jelenleg limitálhatja az elérhető tér diverzitását – ugyanakkor rámutat arra is, hogy mely alapanyagok számát lehet célszerű növelni.

Köszönetnyilvánítás

Jelen munka támogatásban részesült az Nemzeti Kutatási, Fejlesztési és Innovációs Hivatal 2020-1.1.2-PIACI-KFI-2020-00050 azonosítószámú pályázat keretében.

Hivatkozások

1. Zhao H. *ACS Med. Chem. Lett.*, **2024**, 15, 4, 436–440
<https://doi.org/10.1021/acsmmedchemlett.4c00093>
2. Omar, Ö. H.; del Cueto, M.; Nematiamram, T.; Troisi, A. J. *Mater. Chem. C*, **2021**, 9, 13557–13583
<https://doi.org/10.1039/D1TC03256A>
3. Takács, G.; Havasi, D.; Sándor, M.; Dohánics, Zs.; Balogh, Gy. T.; Kiss, R. *ACS Med. Chem. Lett.* **2023**, 14, 9, 1188–1197
<https://doi.org/10.1021/acsmmedchemlett.3c00146>
4. Patel, H.; Ihlenfeldt W-D.; Judson, P. N.; Moroz, Y. S.; Pevzner, Y.; Peach, M. L.; Delannée, V.; Tarasova, N. I.; Nicklaus, M. C. *Sci Data*, **2020**, 7, 384
<https://doi.org/10.1038/s41597-020-00727-4>
5. Daylight. SMARTS: a language for describing molecular patterns [Internet] [2025 április 28.]
<http://www.daylight.com/dayhtml/doc/theory/theory.smarts.html>
6. Daylight. SMIRKS: A Reaction Transform Language [Internet] [2025 április 28.]
<http://www.daylight.com/dayhtml/doc/theory/theory.smirks.html>
7. Hartenfeller, M.; Eberle, M.; Meier, P.; Nieto-Oberhuber, C.; Altmann, K-H.; Schneider, G.; Jacoby, E.; Renner, S. J. *Chem. Inf. Model.* **2011**, 51, 12, 3093–3098
<https://doi.org/10.1021/ci200379p>
8. Roughley, S. D.; Jordan, A. M. *J. Med. Chem.* **2011**, 54, 10, 3451–3479
<https://doi.org/10.1021/jm200187y>
9. Dömling, A.; Wang, W.; Wang, K. *Chem. Rev.* **2012**, 112, 6, 3083–3135
<https://doi.org/10.1021/cr100233r>
10. Agenet, N.; Buisine, O.; Slowinski, F.; Gandon, V.; Aubert, C.; Malacria, M. *Organic Reactions*, **2007**
<https://doi.org/10.1002/0471264180.or068.01>
11. Gewald, K.; Schinke, E.; Böttcher, H. *Ber.* **1966**, 99, 94–100
<https://doi.org/10.1002/cber.19660990116>
12. Asinger, F.; Offermanns, H. *Angew. Chem. Internat. Edit.* **1967**, 6, 11, 907–919
<https://doi.org/10.1002/anie.196709071>
13. Khand, I. U.; Knox, G. R.; Pauson, P. L.; Watts, W. E. *J. Chem. Soc., Perkin Trans. 1*, **1973**, 975–977
<https://doi.org/10.1039/P19730000975>
14. Singh, S. K.; Zhu, C.; La Jeunesse, J.; Fortenberry, R. C.; Kaiser, R. I. *Nat. Commun.* **2022**, 13, 375
<https://doi.org/10.1038/s41467-022-27963-z>
15. Hantzsch, A. *Ber. Dtsch. Chem. Ges.* **1881**, 14, 1637–1638.
<https://doi.org/10.1002/cber.18810140214>
16. Bucherer, H. T.; Steiner, W. *J. Prakt. Chem.* **1934**, 140, 291–316
17. Mannich, C.; Krösche, W. *Arch. Pharm. Pharm. Med. Chem.*, **1912**, 250, 647–667.
<https://doi.org/10.1002/ardp.19122500151>
18. Manabe, K.; Kobayashi, S. *Org. Lett.*, **1999**, 1, 12, 1965–1967
<https://doi.org/10.1021/ol991113u>
19. Bepalova, G. V.; Sedavkina, V. A.; Kharchenko, V. G. *Chem. Heterocycl. Compd.* **1979**, 15, 1035
<https://doi.org/10.1007/BF00473853>
20. Biginelli, P. *Chemische Berichte*, **1891**, 24, 1317
<http://dx.doi.org/10.1002/cber.189102401228>
21. Ryabukhin, S., V.; Plaskon, A. S.; Ostapchuk, E. N.; Volochnyuk, D. M.; Tolmachev, D. M. *Synthesis*, **2007**, 417–427
<https://doi.org/10.1055/s-2007-965881>
22. Ranu, B. C.; Hajra, A.; Jana, U. *J. Org. Chem.* **2000**, 65, 19, 6270–6272
<https://doi.org/10.1021/jo000711f>
23. Betti, M. *Gazz. Chum. Ital.* **1900**, 30 II, 301
24. Olyaei, A.; Sadeghpour, M. *RSC Adv.*, **2024**, 14, 11811–11848
<https://doi.org/10.1039/D4RA01256A>
25. Cimarelli, C.; Fraton, D.; Mazzanti, A.; Palmieri, G. *Eur. J. Org. Chem.*, **2011**: 2094–2100.
<https://doi.org/10.1002/ejoc.201001611>
26. Kabachnik, M. I.; Medved, T. Y. *Dokl. Akad. Nauk SSSR* **1952**, 83, 689–692.
27. Fields, E. K. *J. Am. Chem. Soc.* **1952**, 74, 6, 1528–1531
<https://doi.org/10.1021/ja01126a054>
28. Keglevich, Gy.; Bálint, E. *Molecules* **2012**, 17, 11, 12821–12835
<https://doi.org/10.3390/molecules171112821>
29. Povarov, L. S.; Mikhailov, B. M. *Akad. Nauk SSSR, Ser. Khim.* **1963**, 953–956.
30. Larsen, S. D.; Grieco, P. A. *J. Am. Chem. Soc.* **1985**, 107, 1768.
<https://doi.org/10.1021/ja00292a057>
31. Kanth, S. R.; Reddy, G. V.; Maitraie, D.; Narsaiah, B.; Rao, P. S.; Kumar, K. R.; Sridhar, B. *J. Fluor. Chem.* **2006**, 127, 1211–1221
<https://doi.org/10.1016/j.jfluchem.2006.06.017>
32. Lemos, B. C.; Filho, E. V.; Fiorot, R. G.; Medici, F.; Greco, S. J.; Benaglia, M. *Eur. J. Org. Chem.* **2022**, e202101171
<https://doi.org/10.1002/ejoc.202101171>
33. Masdeu, C.; de los Santos, J. M.; Palacios, F.; Alonso, C. *Top. Curr. Chem. (Z)* **2023**, 381, 20
<https://doi.org/10.1007/s41061-023-00428-7>
34. Petasis, N. A.; Akritopoulou, I. *Tetrahedron Lett.* **1993**, 34, 4, 583–586
[https://doi.org/10.1016/S0040-4039\(00\)61625-8](https://doi.org/10.1016/S0040-4039(00)61625-8)
35. Wu, P.; Givskov, M.; Nielsen, T. E. *Chem. Rev.* **2019**, 119, 20, 11245–11290
<https://doi.org/10.1021/acs.chemrev.9b00214>
36. Koolmesiter, T.; Södergren, M.; Scobie, M. *Tetrahedron Lett.* **2002**, 43, 34, 5965–5968
[https://doi.org/10.1016/S0040-4039\(02\)01265-0](https://doi.org/10.1016/S0040-4039(02)01265-0)
37. Wei, C.; Li, Z.; Li, C-J. *Synlett* **2004**, 9, 1472–1483
<https://doi.org/10.1055/s-2004-829531>
38. Volkova, Y.; Baranin, S.; Zavarzin, I. *Adv. Synth. Catal.* **2021**, 363, 40.
<https://doi.org/10.1002/adsc.202000866>
39. Jesin, I.; Nandi, G. C. *Eur. J. Org. Chem.*, **2019**: 2704–2720.
<https://doi.org/10.1002/ejoc.201900001>

40. Farhi, J.; Lykakis, I. N.; Kostakis, G. E. *Catalysts* **2022**, *12*, 6, 660
<https://doi.org/10.3390/catal12060660>
41. Hao, W.; Gao, S.; Cui, H.; Ding, D.; Jiang, S.; Zhang, C.; Ji, Y.; Zhang, G. *J. Org. Chem.* **2024**, *89*, 4, 2605–2621
<https://doi.org/10.1021/acs.joc.3c02621>
42. Phelps, J. M.; Kumar, R.; Robinson, J. D.; Chu, J. C. K.; Flodén, N. J.; Beaton, S.; Gaunt, M. J. *J. Am. Chem. Soc.* **2024**, *146*, 13, 9045–9062
<https://doi.org/10.1021/jacs.3c14037>
43. Smith, M. A.; Kang, R. J. D.; Kumar, R.; Roy, B.; Gaunt, M. *J. Chem. Sci.*, **2024**, *15*, 14888–14898
<https://doi.org/10.1039/D4SC03916E>
44. Döbner, O. *Justus Liebigs Ann. Chem.*, **1887**, 242: 265–289.
<https://doi.org/10.1002/jlac.18872420302>
45. Komatsu, H.; Shigeyama, T.; Sugimoto, T.; Nishiyama, H. *J. Org. Chem.* **2023**, *88*, 17, 12816–12820
<https://doi.org/10.1021/acs.joc.3c01123>
46. Passerini, M.; Simone, L. *Gazz. Chim. Ital.* **1921**, *51*, 126–129
47. Sztanó, G.; Dobi, Z.; Soós, T. *ChemistryOpen* **2023**, *12*, e202200083
<https://doi.org/10.1002/open.202200083>
48. Ferdeghini, C.; Wu, M.; Ranjan, P.; Würdemann, M. A.; Pyschik, J.; Mitsos, A.; Ruijter, E.; Orru, R. V. A.; Hansen, T. *J. Org. Chem.* **2025**, *90*, 14, 5000–5007
<https://doi.org/10.1021/acs.joc.5c00236>
49. Ugi, I.; Meyr, R.; Fetzer, U.; Steinbrückner, C. *Angew. Chem.* **1959**, *71*, 373–388.
<https://doi.org/10.1002/ange.19590711110>
50. Rocha, R. O.; Rodrigues, M. O.; Neto, B. A. D. *ACS Omega* **2020**, *5*, 2, 972–979
<https://doi.org/10.1021/acsomega.9b03684>
51. Oldenziel, O. H.; van Leusen, D.; van Leusen, A. M. *J. Org. Chem.* **1977**, *42*, 19, 3114–3118
<https://doi.org/10.1021/jo00439a002>
52. Zheng, X.; Ma, Z.; Zhang, D. *Pharmaceuticals* **2020**, *13*, 3, 37
<https://doi.org/10.3390/ph13030037>
53. Geigle, S. N.; Petersen, A. C.; Satz, A. L. *Org. Lett.* **2019**, *21*, 22, 9001–9004
<https://doi.org/10.1021/acs.orglett.9b03406>
54. Groebke, K.; Weber, L.; Mehlin, F. *Synlett* **1998**, 661–663
<https://doi.org/10.1055/s-1998-1721>
55. Blackburn, C.; Guan, B.; Fleming, P.; Shiosaki, K.; Tsai, S. *Tetrahedron Lett.* **1998**, *39*, 3635–3638
[https://doi.org/10.1016/S0040-4039\(98\)00653-4](https://doi.org/10.1016/S0040-4039(98)00653-4)
56. Bienaymé, H.; Bouzid, K. *Angew. Chem. Int. Ed.* **1998**, *37*, 2234–2237
[https://doi.org/10.1002/\(SICI\)1521-3773\(19980904\)37:16%3C2234::AID-ANIE2234%3E3.0.CO;2-R](https://doi.org/10.1002/(SICI)1521-3773(19980904)37:16%3C2234::AID-ANIE2234%3E3.0.CO;2-R)
57. An, Y.; Lee, J.; Seo, H.; Bae, S.; Kang, J.; Lee, J.; Kim, J.; Nam, M. H.; Song, M.; Hwang, G. G. *Org. Lett.* **2023**, *25*, 24, 4445–4450
<https://doi.org/10.1021/acs.orglett.3c01366>
58. Martini, C.; Mardjan, M. I. D.; Basso, A. *Beilstein J. Org. Chem.* **2024**, *20*, 1839–1879.
<https://doi.org/10.3762/bjoc.20.162>
59. Bon, R. S.; Hong, C.; Bouma, M. J.; Schmitz, R. F.; de Kanter, F. J. J.; Lutz, M.; Spek, A. L.; Orru, R. V. A. *Org. Lett.* **2003**, *5*, 20, 3759–3762
<https://doi.org/10.1021/ol035521g>
60. Baell, J. B.; Holloway, G. A. *J. Med. Chem.* **2010**, *53*, 7, 2719–2740
<https://doi.org/10.1021/jm901137j>
61. Mcule In Stock 2025. március 9-i verzió [Internet] [2025 április 25.]
<https://mcule.com/database/>
62. Mcule Fast Delivery Building Block Class 2 2025. április 20-i verzió [Internet] [2025 április 25.]
<https://mcule.com/database/>
63. Key Organics BIONET Building Blocks 2025. márciusi verzió [Internet] [2025 április 25.]
<https://www.keyorganics.net/downloads-bionet-databases/>
64. Sanchez-Garcia, R.; Havasi, D.; Takács, G.; Robinson, M. C.; Lee, A.; von Delft, F.; Deane, C. M. *Digital Discovery*, **2023**, *2*, 103–111
<https://doi.org/10.1039/D2DD00071G>
65. Si, Y.-X.; Zhu, P.-F.; Zhang, S.-L. *Org. Lett.* **2020**, *22*, 22, 9086–9090
<https://doi.org/10.1021/acs.orglett.0c03472>
66. Domingo, L. R.; Aurell, M. J.; Sáez, J. A.; Mekelleche, S. M. *RSC Adv.*, **2014**, *4*, 25268–25278
<https://doi.org/10.1039/C4RA02916J>
67. Peshkov, V. A.; Pereshivko, O. P.; van der Eycken, E. V. *Chem. Soc. Rev.*, **2012**, *41*, 3790–3807
<https://doi.org/10.1039/C2CS15356D>

Cheminformatic evaluation of multicomponent reactions for the development of virtual chemical libraries

In the early stage of drug discovery nowadays millions of molecules can be screened via computational methods to identify structures that could potentially be utilized in the development of novel drugs. Although the number of in-stock available molecules is estimated to be between 10-20 million, virtual molecules opened the possibility to investigate billions of structures. As novelty is highly needed, the further development and extension of the chemical space is still interesting, however most of the currently available virtual libraries are based on the same reaction types and usually combining these reactions in 2-3 consecutive steps. Virtual molecules are also expected to be synthesizable with a 60-80% success rate – which is greatly affected by reactivity, purification and the stability of substructures under the synthetic conditions.

Multicomponent reactions (MCR) are a unique class of synthetic methods as they are capable to combine three or more components in one synthetic step. Although the first types of these reactions were already identified in the late 19th century, there is an emerging interest recently in their utilization in the development of drug-like molecules. Certain types of multicomponent reactions rely on some reactants bearing a specific functional group and some fixed components, like the Gewald-, Asinger-, Pauson-Khand-, Strecker-, Hantzsch- or Bucherer-Bergs reactions. With a fixed component like sulfur or ammonia, only the other components will be variable, thus a similar library size to the A+B type reactions could be expected. Reactions where all three (or more) components have a variable substituent part could generate much larger combinatorial libraries, therefore in the present study we aimed to collect these reaction types, create their cheminformatic representations and evaluate their potential in extending the chemical space. As part of the work, 15 multicomponent reactions were identified: the Biginelli to form dihydropyrimidones, the Mannich for α -amino carbonyls, the Betti reaction generating α -aminobenzylphenols, the Grieco/Povarov ring closing reaction, the Kabachnik-Fields synthesis for aminophosphonate formation, the Petasis reaction for amine substitution, the A3 coupling for propargyl amines, the Doebner synthesis of quinoline and two recently developed methods by Hao et al. and Gaunt et al. A separate group of multicomponent reactions of isocyanides was also included, covering the Passerini, Ugi, van Leusen, Groebke-Blackburn-Bienayme, Orriu reactions. To evaluate the potential of the reactions, first the variability of substructures was investigated

based on scientific literature. Based on the findings, the general schemes of all reactions were constructed (Figures 1-15), which served as a basis for creating the SMIRKS reaction representation of all involved reactions. Based on the reactant and product SMARTS of the reaction SMIRKS, the availability of the compounds containing the desired substructures was evaluated. The investigation showed that many of the product types have a very low representation among the in-stock molecules.

By evaluating the availability of building blocks required for a reaction, an initial estimation of the potential chemical space could be provided. Therefore, the building block SMARTS were searched, and a product count estimation was provided based on the results. It should be noted that the actually available chemical space could be much smaller due to multiple factors: competing functionalities that could interfere with reactive centers under the synthesis conditions; lability of substructures; duplicate formation from functionally different molecules, etc. Even if these considerations are considered, it is clearly visible that billions of virtual structures could be accessed via multicomponent reactions. It is also visible that some reactions, e.g. the A3 coupling or the Petasis reaction utilize building blocks that are available in very large numbers, due to their participation in the most commonly applied reaction types, like amines and acids in amide formation, or acetylenes and boronic acids in the Sonogashira and Suzuki couplings. On the other hand, we can see that many of the reactions are mainly limited by one of the required components, like in the case of isocyanides. However, as the interest for these reactions is growing, there could be a rise in the number of available building blocks in the future, allowing the faster spread of isocyanide-based multicomponent reactions. Another limitation of multicomponent reactions currently is that compared to the commonly applied synthetic reactions in medicinal chemistry have a much greater validation database, including various synthetic protocols. In the future with the continuous extension of validation data for MCRs, we could expect a much wider utilization, especially if mild, efficient and more generalizable conditions could be developed. Both can highly be boosted by high throughput experimentation techniques. The main goal of the present study was to provide a collection of multicomponent reactions and their SMIRKS representations for researchers, allowing them to go beyond the traditional virtual spaces, that are based on regular 2-3 step syntheses.