



Szegedi Tudományegyetem

Természettudományi és Informatikai Kar

Alkalmazott és Környezeti Kémiai Tanszék

Diplomamunka

Fémionok szerepének kvantumkémiai vizsgálata fém-oxid nanorészecskék toxicitásában

Készítette:

Tasi Domonkos Attila

Vegyész MSc szak

Témavezető:

Dr. Kónya Zoltán

tanszékvezető egyetemi tanár

Dr. Tasi Gyula

egyetemi docens

Szeged

2017

Tartalmi összefoglalás és kulcsszavak

Leszczynski és mtsai. 2011-ben felállítottak egy nano-QSAR-modellt fém-oxid nanorészecskék kólibaktériumokra vonatkozó citotoxicitásának leírására. A QSAR-modellezéshez a PM6 szemempirikus kvantumkémiai módszert használták. A számított 12 kvantumkémiai deszkriptor közül végül is csak egyetlen egy, a fém-oxid nanorészecskéket alkotó fématomok ionjainak a PM6 standard képződéshője bizonyult megfelelőnek. Wu és mtsai. – kiegészítve az eredeti modellt – egy jobbat állítottak fel. Modelljük két deszkriptorra épült: a fémionok PM6 standard képződéshőjére és az ionok polarizációs erejére. A szerzők javasoltak egy mechanizmust is a fém-oxid nanorészecskék toxicitásának az értelmezésére. Munkánk során a közölt nano-QSAR-modellek jóságát vizsgáltuk és sikerült egy realisabb modellt is felállítanunk, fizikailag realisabb deszkriptorok alkalmazásával.

Kvantumkémiai és matematikai statisztikai módszerek segítségével megállapítottuk, hogy a Leszczynski és mtsai. által felállított nano-QSAR-modell alapját képező adatok több mint fele, 17 fémion közül 9, hibás vagy nagy hibával terhelt, így fizikailag irreális. A Wu és mtsai. által felállított modell esetén is 23 deszkriptor értéke valamilyen oknál fogva nem bizonyult megfelelőnek. Lineáris regressziókat végeztünk a toxicitási adatok és a fémionokra vonatkozó, általunk számított PM6 képződéshők, illetve a rendelkezésünkre álló kísérleti értékek között. Kiderült, hogy a szerzők által közölt PM6 értékek jelentősen eltérnek a kísérleti adatoktól. Végül is a fémionok standard képződéshője helyett kísérleti és magasabb szinteken számított elméleti ionizációs energiát használtunk deszkriptorként.

Az ionizációs energiák kvantumkémiai számításához első közelítésben a G2 és a CCSD(T) módszereket alkalmaztuk. A továbbiakban Dr. Csontos József és mtsai. – együttműködésben velünk – nagy pontosságú kvantumkémiai számításokat végeztek: meghatározták a modell felállításához szükséges összes ionizációs energia lehető legpontosabb értékét. A számítási eredmények igen jól megegyeznek a kísérleti értékekkel. Mivel a kísérleti értékek nem állnak minden esetben rendelkezésre (pl. W(VI)), ez a tény fontos a modell alkalmazhatósága szempontjából.

Kulcsszavak: nanorészecske, fémionok, kólibaktérium, QSAR, ionizációs energia

Tartalomjegyzék

1. Bevezetés és irodalmi összefoglaló	2
2. Célkitűzés.....	8
3. Elméleti összefoglaló	9
3.1 QSAR-modell felépítése	9
3.2 Modellek validálása	11
3.3 Kvantumkémiai módszerek.....	12
3.3.1 Kvantumkémiai bázisok.....	12
3.3.2 Hartree-Fock-Roothaan-Hall modell	14
3.3.3 Csatolt klaszter (CC) módszer	17
3.3.4 Szemiempirikus módszerek	18
3.4 Fémionok képződéshője.....	19
3.5 Paraméterbecslés a legkisebb négyzetek módszerével	21
3.6 Lineáris és multilineáris regresszió.....	22
4. Eredmények	26
4.1 Modellek jóságának vizsgálata	26
4.2 Fématomok ionizációs energiája, mint deskriptor	33
4.3 Fém-oxid nanorészecskék toxicitásának becslése.....	40
5. Összefoglalás	43
6. Felhasznált irodalom.....	44
7. Köszönetnyilvánítás.....	47
8. Nyilatkozat.....	48

1. Bevezetés és irodalmi összefoglaló

A klasszikus fizika törvényei szerint működő makrovilág és a kvantumfizika szerint működő mikrovilág között helyezkedik el a nanovilág. A kvantumfizika határesetben azonban tartalmazza a klasszikus fizikát, így a kvantumfizika alkalmas a nanovilág fizikai-kémiai leírására. A nanovilág olyan része a térnek, melyben az élettelen vagy élő anyag atomokból önszerveződéssel jön létre úgy, hogy az alkotóknak (részecskéknél) legalább az egyik dimenziójuk, azaz a méretük 100 nm vagy ennél kisebb [1,2].

A nanorészecskék szerves részét képezik a modern életnek, napjainkban széleskörűen alkalmazzák őket [3]. Nanoanyagok (nanorészecskék, nanocsövek, nanoszálak, nanotűk, stb.) megtalálhatók különféle elektronikai eszközökben, kozmetikai készítményekben, naptejekben, napelemekben, gyógyszerekben, öntisztító ablakokban, sőt még folttaszító ruhákban is [4,5,6]. A nanotudomány folyamatos fejlődése azonban kockázattal jár, ugyanis a nanoméretű anyagok toxikusak lehetnek élő szervezetek számára [7,8]. Ennek a vizsgálata azonban nem egyszerű, mivel nagyon nehéz pontosan specifikálni nanorészecskék esetén a toxikusságot meghatározó jellemzőket (a részecskék mérete, alakja, felülete, kémiai összetétele, stb.) [9]. Ezért – az esetleges toxikusság elkerülése érdekében – a nanotechnológiában intelligens megoldásokat, eljárásokat fejlesztenek ki. Az intelligens megoldások például a nanoanyagok környezetbarát kémiai szintézisét és funkcionálizálását jelentik egyszerű szerves anyagokból (fémek, fém-oxidok, szénnanocsövek, stb.), a biológiai szintézisek preferálását a hagyományos szintézisek helyett, továbbá elméleti tervezési megfontolások alkalmazását az eljárások hatékonyságának és a nanorészecskék tulajdonságainak a javítására [10,11].

A nanorészecskéken belül két alapvető osztályt különböztetünk meg; a szénnanorészecskéket és a fém/fém-oxid nanorészecskéket [3]. Fém-oxid nanorészecskéket használnak fel a kozmetikai iparban, a textiliparban, víztisztító berendezésekben [12], továbbá alkalmazzák őket autóiipari katalitikus konverterekben is. A fém-oxid nanorészecskék, köszönhetően méretüknek és felületük morfológiájának, egyedülálló fizikai és kémiai tulajdonságokkal rendelkeznek. Nemrég kimutatták, hogy a fém-oxid nanorészecskék – eltérően a makro és a mikro változatoktól – toxikusak néhány organizmusra [13].

A mindennapi életben használt nanoanyagok száma gyorsan növekszik, így szükséges gyors módszerek kifejlesztése toxikusságuk és környezetre való kockázatuk becslésére. A toxicitás elméleti (*in silico*) előrejelzése az utóbbi évtizedekben egyre nagyobb jelentőségre tett szert, általában szerkezet-aktivitás összefüggést (SAR) vagy kvantitatív szerkezet-aktivitás összefüggést (QSAR) alkalmaznak a toxicitás modellezésére [14,15]. Az utóbbi esetben a kapcsolat konkrét matematikai formában is meg van fogalmazva.

Legtágabb értelemben a számítógépes toxikológia matematikai és számítógépes modellek révén segíti felbecsülni anyagok veszélyességét és kockázatát az emberi életre és a környezetre [16]. Legegyszerűbb esetben az anyagok egyszerűsített kémiai szerkezetéből próbálunk következtetni azok toxikusságára. Ez a módszer jelenti a legkisebb pénz- és időráfordítást, sőt új anyagok esetén még predikciót is tudunk tenni a toxikusságukról előállításuk előtt.

A sejthalál természetes formáit már a 19. században feljegyezték biológusok [17]. Akkoriban a biológusok érdeklődése elsősorban a sejtek életének és szaporodásának megértése felé irányult, csak később, az 1960-as évektől kezdve került az érdeklődés középpontjába a sejthalál jelensége. Az inaktív módon lejátszódó, szabályozott molekuláris folyamatok által "programozott" sejtelhalást – amely gyakorlatilag egy időzített önmegsemmisítés – nevezzük apoptózisnak. Ez és egyéb alternatív formái mind energiát igénylő, relatíve lassú folyamatok. Ezzel szemben a nekrosis ártalmas mechanikai és kémiai hatások által bekövetkező, általában egész sejtcsoportokat érintő masszív sejtelhalás. Ez minden esetben egy energiát nem igénylő, passzív folyamat. A sejthártya felszakad, kontrollálatlan anyagcsere következik be a külső és belső celluláris tér közt, melynek végkifejleteként a sejtet az őt ért külső hatás megöli.

Különböző anyagok toxikusságát – kvantitatív módon – az 50%-os halálozási koncentrációjukkal (EC_{50}) jellemezhetjük. Ez megadja azt a koncentrációértéket, melynél a kiindulási sejtállomány fele elhal. Egyéb jelölések is használatosak a szakirodalomban: LD_{50} , LC_{50} , stb.

A kísérleti tapasztalatok azt mutatják, hogy a nanorészecskék – az anyagi minőségüktől függően – nem képesek átjutni a sejtmembránon az intracelluláris térbe egy bizonyos méret felett. Az intracelluláris térbe való kerülés diffúzió, endocitózis, ioncsatornák és transzporter fehérjék révén lehetséges. Ezek a folyamatok általában nagymértékben függnak a nanorészecskék összetételétől, méretétől, morfológiájától és

töltésétől [18]. A nanorészecskék toxikussága függ a részecskék méretétől, koncentrációjától és a felületen lévő esetleges funkcionális csoportoktól [19].

Leszczynski és mtsai. 2011-ben publikáltak egy nano-QSAR-modellt fém-oxid nanorészecskék toxikusságának becslésére [20,21,22]. A cikkekben 17 fém-oxid szerkezete és kólibaktériumokra vonatkozó toxicitása között próbáltak kapcsolatot megállapítani. A kólibaktériumok különösen jó kísérleti alanyok, mivel a természetben fontos szerepet játszanak szerves anyagok lebontásában. A kísérletek során a legkevésbé toxikusnak a TiO₂, míg a legtoxikusabbnak a CoO nanorészecske bizonyult. A vizsgált fém-oxidok – melyeknek a mérete 15 és 90 nm közé esett – kvantumkémiai szempontból nagy rendszereknek számítanak, ezért a szerzők egy gyorsabb, de pontatlanabb szemiempirikus kvantumkémiai módszert használtak. A nanorészecskék szerkezetének és tulajdonságainak pontos jellemzése érdekében először nagyméretű klaszterekre végeztek számításokat. Az egyes fém-oxid nanorészecskék toxicitásának jellemzéséhez azok EC₅₀ értékét vették alapul. A QSAR-modell felállításához PM6 szemiempirikus módszerrel [23] összesen 12 kvantumkémiai deskriptor értékét határozták meg (1. táblázat).

Szimbólum	Deskriptor
HOF	Az oxid klaszter standard képződéshője
TE	Az oxid klaszter teljes energiája
EE	A klaszter elektronenergiája
Core	Mag-mag taszítási energia a klaszteren belül
CA	Az oxid klaszter felszíne
CV	Az oxid klaszter térfogata
HOMO	A klaszter legmagasabb betöltött molekulapályájának energiája
LUMO	A klaszter legalacsonyabb nem betöltött molekulapályájának energiája
GAP	A HOMO és LUMO közötti energiakülönbség (energiagát)
ΔH_{Clust}	A klaszter felületéről történő Me ⁿ⁺ fémionok leválásának entalpiája
ΔH_{Me^+}	Fémionok standard képződéshője
ΔH_{L}	Az oxid klaszter rácsenergiája

1. táblázat: A kiválasztott 12 kvantumkémiai deskriptor [20].

A deskriptorokat három különböző osztályba sorolták. Az első osztályba kerültek azon jellemzők, melyek a klaszter méretével lineárisan változnak (HOF, TE,

EE, Core, CA és CV). A második osztályba kerültek azon deskriptorok, melyek változása elhanyagolható 15 és 90 nm között (HOMO, LUMO, GAP). A harmadik osztályba a maradék három deskriptor került (ΔH_{Clust} , ΔH_{Me^+} , ΔH_{L}). A kiindulási cellaegység koordinátáit krisztallográfiai adatok segítségével határozták meg. Ezután folyamatosan növelték a rácsparamétereket mindhárom dimenzió mentén. Első lépésként a kiindulási cella egyik dimenzióját a kétszeresére növelték. Második lépésként már a megnövelt struktúrát használták kiindulásként és a második dimenzióját is megduplázták. Végül a harmadik dimenzióját is megduplázták. Ezt az eljárást addig folytatták, míg a klaszter mérete el nem érte a 12 Ångströmöt. Ezek után a klasztereket – mint a tanulmányozott nanorészecskék molekulamodelleit – használták kvantum-kémiai számításokra PM6 szemempirikus szinten a MOPAC 2009 programcsomag segítségével. A számítások két lépésből álltak: (1) a molekulageometria optimalizálása megnövelt precíziós kritériumokkal, (2) a deskriptorok számítása az optimalizált geometrián.

Vizsgálataik szerint az egyes fém-oxid nanorészecskéket alkotó fématomok ionjainak a standard képződéshői (ΔH_{Me^+}) a legmegfelelőbb jellemzők. Ok-okozati összefüggést is találtak a kiválasztott deskriptor és a kísérleti adatok között. Ugyanis a szerzők véleménye szerint egy megfelelő QSAR-modell felállításához nem elegendő csak statisztikai korreláció megléte, hanem szükséges, hogy a deskriptorok és a prediktorok között ok-okozati összefüggés is fennálljon. A 2. táblázatban megtalálhatók a kísérletileg meghatározott EC_{50} koncentrációk (dimenziójuk mol/L) és a fém-oxidokat felépítő atomok ionjainak PM6 módszerrel számított képződéshői. A szerzők az alábbi nano-QSAR összefüggéshez jutottak:

$$-\lg(EC_{50}) = 2.59 - 0.50 \Delta H_{\text{Me}^+}. \quad (1)$$

Értelmezésük szerint a fém-oxid nanorészecskék toxikusságát a vízben kioldódó fémionok okozzák.

Fém-oxid	Me ⁿ⁺	pMEC ₅₀	ΔH _{Me+} (kcal/mol)	Készlet
ZnO	Zn(II)	3.45	662.44	T
CuO	Cu(II)	3.20	706.25	T
V ₂ O ₃	V(II)	3.14	1097.73	V ₁
Y ₂ O ₃	Y(II)	2.87	837.15	T
Bi ₂ O ₃	Bi(III)	2.82	1137.40	T
In ₂ O ₃	In(III)	2.81	1271.13	T
Sb ₂ O ₃	Sb(III)	2.64	1233.06	V ₁
Al ₂ O ₃	Al(III)	2.49	1187.83	T
Fe ₂ O ₃	Fe(III)	2.29	1408.29	T
SiO ₂	Si(IV)	2.20	1686.38	T
ZrO ₂	Zr(IV)	2.15	1357.66	V ₁
SnO ₂	Sn(IV)	2.01	1717.32	T
TiO ₂	Ti(IV)	1.74	1575.73	T
CoO	Co(II)	3.51	601.80	V ₂
NiO	Ni(II)	3.45	596.70	V ₂
Cr ₂ O ₃	Cr(III)	2.51	1268.70	V ₂
La ₂ O ₃	La(III)	2.87	1017.22	V ₂

2. táblázat: A fém-oxid nanorészecskék EC₅₀ koncentrációi és a fémionok képződéshői (1.modell) [20].

Némely szerző szerint a fenti QSAR-modell nem tekinthető igazi nano-QSAR-modellnek, mivel a kísérleti adatok nem függenek a részecskék méretétől [24]. A nano-QSAR-modell szerintük három tagból áll:

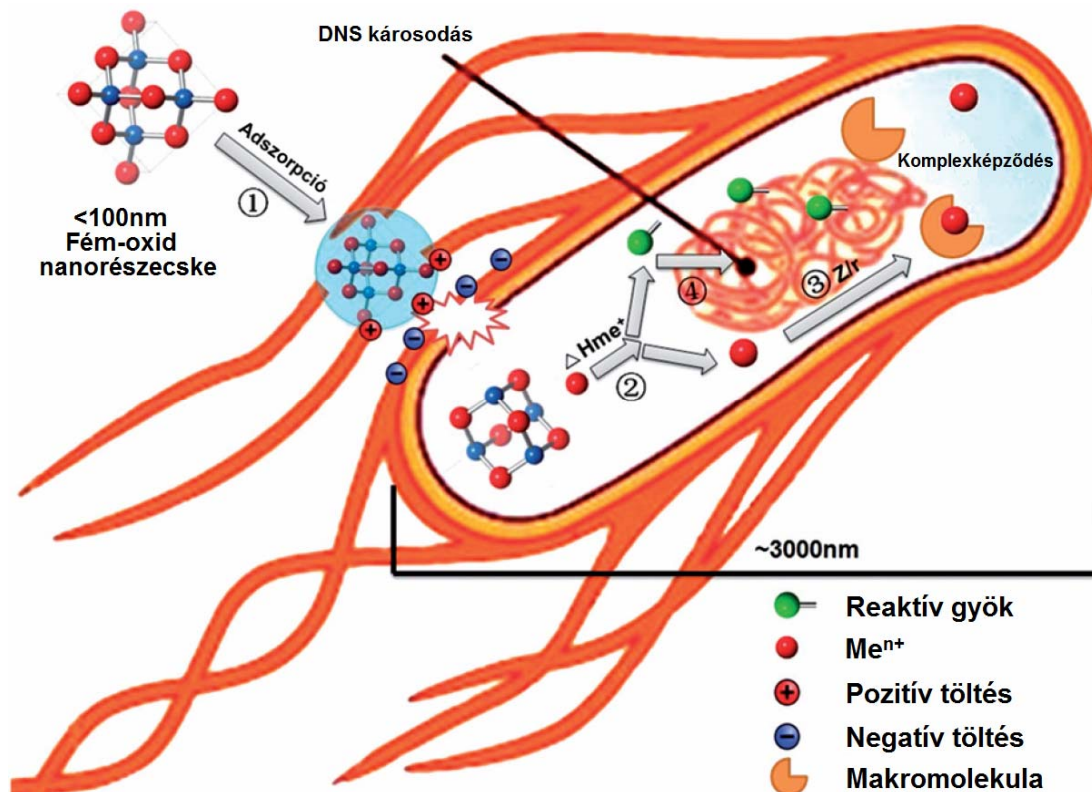
$$\pi = \alpha P(s) + \beta Q(d) + \gamma S(p), \quad (2)$$

ahol π a nanorészecske valamilyen tulajdonsága (pl.: EC₅₀ koncentráció) és $\{\alpha, \beta, \gamma\}$ skaláris paraméterek. A $P(s)$ függvény írja le, hogy a vizsgált tulajdonság hogyan függ a nanorészecske méretétől. A $Q(d)$ függvény a hagyományos QSAR modell, ahol d a deskriptorok vektora. A harmadik tag, az $S(p)$ függvény speciális effektusokat ír le. A p vektor például figyelembe veheti a nanorészecskék alakját vagy a szolvatációs effektusokat. Az esetleges rejtett paraméterekkel is itt számolunk el. A gyakorlatban általában elhanyagoljuk a harmadik tagot.

Wu és mtsai. legutóbbi cikkükben [25] az előző QSAR-modellt továbbgondolva – 26 deskriptort figyelembe véve – egy újabb, pontosabb modellt állítottak fel

(2.modell). A 26 deszkriptorból azonban végül is csak kettő bizonyult megfelelőnek: a fém-oxid nanorészecskéket alkotó fématomok ionjainak a standard képződéshői (ΔH_{Me^+}), illetve az egyes ionok polarizációs ereje (Z/r). A standard képződéshőket ők is a PM6 szemiempirikus módszerrel határozták meg (MOPAC 2012). A felállított modelljükkel összesen 51 fém-oxid nanorészecske kólibaktériumokra vonatkozó citotoxicitását becsülték meg.

A szerzők felvázoltak egy lehetséges mechanizmust is fém-oxid nanorészecskék kólibaktériumokra (*Escherichia coli*) való toxicitására [25]. A mechanizmus négy lépésből áll (1. ábra). Először a fém-oxid nanorészecskék akkumulálódnak (adszorbeálódnak) a sejtmembránon és a membrán felszínén kölcsönhatásba lépnek bizonyos fehérjékkel. Ezután a fém-oxid nanorészecskék fizikai-kémiai reakciók útján bejutnak a citoplazmába, ahol is a gyengén savas körülmények következtében a nanorészecskék felületén felszakadnak a fém-oxigén kötések. Így fémionok és szabad elektronok jönnek létre. A fémionok makromolekuláris komplexképződésben vesznek részt, míg a szabad elektronok kölcsönhatnak a sejtben található különféle molekulákkal, melyek során reaktív oxigéntartalmú gyökök keletkeznek (reactive oxygen species /ROS/). Végül ezek a gyökök a baktériumban DNS károsodást és apoptózist okoznak.



1.ábra: Fém-oxid nanorészecskék toxicitásának feltételezett mechanizmusa kólibaktériumokra [25].

2. Célkitűzés

Leszczynski és mtsai. 2011-ben felállítottak egy nano-QSAR-modellt fém-oxid nanorészecskék E. Coli baktériumokra vonatkozó toxicitására [20]. A modell felállításához PM6 szemiempirikus módszerrel [23] előzetesen 12 kvantumkémiai deskriptort számítottak. Ezek közül azonban csak egy bizonyult megfelelőnek: a fém-oxid nanorészecskéket alkotó fématomok ionjainak PM6 standard képződéshője. Értelmezésük szerint a fém-oxid nanorészecskék toxikusságát a vízben kioldódó fémionok okozzák. Wu és mtsai. legutóbbi [25] cikkükben – a modellt továbbgondolva és 26 deskriptort figyelembe véve – egy pontosabb modellt állítottak fel. A 26 deskriptorból végül is csak kettő bizonyult megfelelőnek: a fémionok PM6 standard képződéshője, illetve az egyes ionok polarizációs ereje. A szerzők javasoltak egy lehetséges mechanizmust is, mely leírja a fém-oxid nanorészecskék toxicitását kólibaktériumokra nézve. A mechanizmus lényege, hogy a fém-oxid nanorészecskék bejutnak a citoplazmába, ahol felszakadnak a fém-oxigén kötések; fémionok és szabad elektronok keletkeznek. A szabad elektronok kölcsönhatnak különféle molekulákkal, s reaktív oxigéntartalmú gyökök keletkeznek. Végül ezek a gyökök okoznak DNS-károsodást és apoptózist baktériumokban. Munkánk során célul tűzzük ki, hogy a fent említett QSAR-modellek helyességét megvizsgáljuk kvantumkémiai és matematikai statisztikai eszközök segítségével. Először is szeretnénk górcső alá venni a deskriptorként alkalmazott PM6 standard képződéshők minőségét. Ezt követően lineáris regressziók végrehajtását tervezzük a toxicitási adatok és a fémionokra vonatkozó, általunk számított PM6 képződéshők, illetve a rendelkezésünkre álló kísérleti adatok között (egyres fémionokhoz tartozók ionizációs energiák, és az ionok standard képződéshői). A későbbiekben szeretnénk nagy pontosságú kvantumkémiai módszerrel is meghatározni a kérdéses deskriptorok értékét.

Végző célunk egy megbízhatóbb nano-QSAR-modell felállítása, melynek segítségével megbecsülhetjük egy tetszőleges fém-oxid nanorészecske citotoxicitását E. coli baktériumra nézve.

3. Elméleti összefoglaló

3.1 QSAR-modell felépítése [16]

A QSAR/QSPR (kvantitatív szerkezet-hatás, kvantitatív szerkezet-tulajdonság) módszerek alapja az, hogy egy vegyület (anyag) fizikai, kémiai és biológiai tulajdonságait determinálja annak szerkezete. Mivel az anyagszerkezet pontos leírása nem könnyű feladat, értelemszerűnek tűnik néhány olyan egyszerű szerkezeti jellemző (un. *deskriptor*) kiválasztása, melyek döntő mértékben meghatározzák a modellezni kívánt hatást/tulajdonságot. Ha sikerül egy ilyen modellt felállítani, akkor becslést (predikciót) tudunk tenni ismeretlen hatású/tulajdonságú anyagok hatására/tulajdonságára azok egyszerűsített kémiai szerkezetének ismeretében.

Egy QSAR/QSPR-modell felépítéséhez megfelelő számú és minőségű kísérleti adat szükséges, melyek döntő mértékben megszabják a modellt, s így a predikció minőségét. A kísérleti adatok minősége mellett fontos tényező a szerkezeti jellemzők helyes megválasztása is. Azt, hogy mely jellemzők a legalkalmasabbak egy jó minőségű modell felállítására, úgy tudjuk eldönteni, hogy kezdetben igen sokat veszünk figyelembe, s később statisztikai és egyéb módszerekkel kiválasztjuk a legmegfelelőbbeket.

Sokáig QSAR összefüggéseket kizárólag molekulacsaládokra, homológ sorokra állítottak fel. Ezek esetén a molekulák alapszerkezete ugyanis azonos: egymástól csak szubsztituensekben, funkcionális csoportokban különböznek. Később általánosabb modelleket is létrehozottak. A QSAR modellezés területén Hansch és Fujita végeztek úttörő munkát az 1960-as években [26,27]. QSAR-modelljeikben vegyületek biológiai hatását hozták kapcsolatba azok hidrofób és elektronszerkezeti tulajdonságaival [28,29].

A molekuláris jellemzőket számos szempont szerint csoportosíthatjuk. Előállításuk módja alapján megkülönböztetünk empirikus és számított deskriptorokat. Az empirikus deskriptorok előnye, hogy jól meghatározott fizikai-kémiai tartalommal rendelkeznek, azonban előállításuk költséges és időigényes. Ezzel szemben a számított jellemzők előállítása könnyű és költséghatékony, de az általuk felépített modellt nehezebb kémiaiilag értelmezni.

Dimenziójuk alapján egy deskriptor lehet *null-dimenziós* (0D), *egydimenziós* (1D), *kétdimenziós* (2D), valamint *háromdimenziós* (3D). A null-dimenziós molekulajellemzők az összetételt leíró skalárértékek, úgymint a moláris tömeg vagy

éppen a különböző típusú atomok száma a molekulában. Az egydimenziós deskriptorok lehetnek például a molekulafragmensek listái, míg a kétdimenziós jellemzőket a molekula szerkezeti képletéből gráfelmélet segítségével állítjuk elő. A háromdimenziós deskriptorok pedig a molekula geometriájából adódnak.

Egy sikeres QSPR/QSAR-modell felállításához megfelelő minőségű kísérleti adat, a szerkezetet, tulajdonságot jól leíró, egymástól független deskriptor, megfelelő regressziós módszer, illetve validáló eljárás szükséges. Nagyon fontos, hogy kísérleti adataink azonos kísérleti körülményekből származzanak, lehetőleg egy adott személy egy adott módszer alapján végezze el a mérést, valamint az megfelelő pontosságú és reprodukálható legyen, ugyanis ezekből a kísérleti eredményekből fog felépülni a kiindulási adathalmazunk. Lényeges szempont az is, hogy az experimentális értékek hibái normális eloszlásúak legyenek, kerüljük az adatok ismétlődését és az extrapolációt. A vizsgált adathalmaz elemeinek egymáshoz viszonyított tulajdonságai alapján különbséget tehetünk *homogén és inhomogén* adathalmazok között (homogén adathalmazban a molekulák azonos vagy nagyon hasonló vegyülettípusokból kerülnek ki). A homogén adathalmazokra jobb QSAR/QSPR-modell állítható fel, ámbar a modell predikciós képessége csökken, így csak az adathalmazban szereplő vegyületekhez hasonló tulajdonságú anyagokra használható fel.

Fontos feladat azonosítani az esetleges kiugró pontokat (mérési hibákat) is adataink halmazában. Érdekes azonban csak akkor eltávolítani a kiugró pontokat, ha több teszt egyöntetűen erre az eredményre vezet. Továbbá, nem szabad megfélekednünk arról sem, hogy egyes kiugró értékek akár valós mérési eredmények is lehetnek.

A QSAR/QSPR-modellezés során leggyakrabban multilineáris regressziót alkalmaznak. Ha két deskriptorral végzett multilineáris regresszióval, azaz három paraméterrel (két meredekség és egy tengelymetszet) jól le tudjuk írni a valóságot, akkor nem szabad több független változót vagy nemlineáris összefüggést alkalmaznunk. Ha ezt tesszük, akkor a túlillesztés hibáját követjük el. Az ellenkező eset is bekövetkezhet, ha a szükségesnél kevesebb paramétert alkalmazunk. Ebben az esetben alulillesztésről beszélünk.

A független változók értékei nem lehetnek befolyással egymásra, így célszerű a változók korrelációi alapján korrelációs analízist végezni a változók közötti kapcsolatok felderítésére. Ha erősen korreláltak, akkor nagy valószínűséggel a becsült paraméterek is azok lesznek. Ennek megfelelően érdemes a független változók korrelációs mátrixát

előállítani, s alaposan megvizsgálni. Az erősen korrelált deskriptorok valamelyikét el kell távolítani. Ellenkező esetben az illesztésünk nem lesz szignifikáns.

Miután kiválasztottuk az n számú megfelelő deskriptort, akkor alapvetően kétféle elvet alkalmazhatunk a legjobb modell kiválasztására: a deskriptorok összes lehetséges kombinációjával elvégezzük az illesztést ($2^n - 1$) vagy lépésenként haladva adjuk hozzá, vagy vesszük el a deskriptorokat egy adott értékig.

3.2 Modellek validálása

A statisztikában validálás alatt egy regressziós modell statisztikai mutatóinak meghatározását értjük. A validálás speciális esete a külső validálás, melynek során a modellt olyan adatpontok halmazán vizsgáljuk, melyek nem szerepelnek az illesztett adathalmazban (tanítóhalmazban). Belső (kereszt)validálásról beszélünk, ha a validálásra szánt pontokat az illesztés során is felhasználjuk. Egyszerű szabályként a következőt alkalmazhatjuk: az adataink 80%-át használjuk illesztésre (*modellező halmaz*), 20%-át pedig külső validálásra (*kiértékelő halmaz*). A modellező halmaz elemeinek 80%-a a *tanítókészletbe*, 20%-a pedig a *tesztkészletbe* kerül.

Ha kevés mérési pontunk van, akkor kizárólag belső validálást érdemes alkalmazni, melyet a legegyszerűbben a „*leave-one-out*”-módszerrel (*LOO*) valósíthatunk meg. Ezen módszer során elhagyunk az adatsorból egymásután egy-egy pontot, majd a többi pontból *MLR-WLS* eljárással határozzuk meg a becsült paramétervektort. Ha elegendő adattal rendelkezünk, akkor érdemes szisztematikusan több pontot is elhagyni az illesztésből („*leave-more-out*”-módszer, *LMO*).

A validálási eljárással ellenőrizhetjük, hogy az általunk használt deskriptorok jól leírják-e a tanítóhalmazt, és jó becsléssel szolgálnak-e. Továbbá információt szolgáltat a modellünk megbízhatóságáról, valamint segít meghatározni a felhasználási területet is.

3.3 Kvantumkémiai módszerek [30]

A kvantumkémia elsődleges célja molekulák geometriájának, elektronszerkezetének és fizikai-kémiai tulajdonságainak a meghatározása kvantumkémiai eszközök felhasználásával. A kvantumkémiai módszereknek két típusát különböztetjük meg: hullámfüggvény-alapú és sűrűségfüggvény-alapú módszerek. A hullámfüggvény-alapú módszerek tovább oszthatók szemempirikus és *ab initio* módszerekre. Az *ab initio* módszerekben a lehető legkevesebb közelítést alkalmazzuk.

A Born-Oppenheimer-féle közelítés az alapja a hagyományos kvantumkémiai módszereknek. Az atommagok tömege sokkal nagyobb, mint az elektronok tömege, így az atommagok sokkal lassabban mozognak. Ennek megfelelően az atommagok és az elektronok mozgását szétválaszthatjuk: a magokat fixáljuk és csak az elektronok mozgását vizsgáljuk. Rögzített magkoordináták esetén megoldva az elektronok mozgására felírt Schrödinger-egyenletet, megkapjuk a rendszer elektronenergiáját, mint a magkoordináták függvényét. Ha ehhez hozzáadjuk a konstansnak tekintett mag-mag tasztítást, megkapjuk egy adott magkonfiguráció esetén a teljes energiát. Ha ezt a számítást sok térbeli magelrendezés esetén elvégezzük, akkor a rendszer potenciális energia-hiperfelületéhez (PES) jutunk.

3.3.1 Kvantumkémiai bázisok

A kvantumkémiai számítások legfontosabb összetevői az atompályák, az ún. bázisfüggvények, melyek révén előállíthatjuk a rendszer elektron-hullámfüggvényét. A számítások során alkalmazott bázisfüggvények halmazát egyszerűen bázisnak nevezzük. Egy halmaz nagyságát a számosságával jellemezhetjük. Minél nagyobb ez, a számításaink általában annál pontosabbá válnak, bár költségesebbé is, mivel nő az adott számításhoz szükséges gépidő. A kiterjesztett viriáltétel segítségével jellemezhetjük bázisaink jóságát [31].

A Slater-pályákat (*Slater-type orbital*, *STO*), helyesebben Slater-típusú függvényeket, valamint a Gauss-pályákat (*Gauss-type orbital*, *GTO*), helyesebben Gauss típusú függvényeket, régóta alkalmazzák – mint bázisfüggvényeket – a kvantumkémiaiában. Kezdetben Slater-típusú pályákat alkalmaztak, annak ellenére, hogy az integrálok egy részének a számítása ebben az esetben általában csak numerikusan lehetséges.

A Slater-pályák Descartes-alakja a következő:

$$\chi_{n,l,m}^{STO}(r;\zeta) = Nx^l y^m z^n e^{-\zeta r}, \quad (3)$$

ahol ζ a STO pályakitevője, N a normálási tényező. A $\lambda = (l, m, n)$ rendezett számhármast impulzusmomentumnak, a $[\lambda] = l + m + n$ számot pedig impulzusmomentum-indexnek nevezzük. A $[\lambda]$ értékétől függően (0, 1, 2, ...) beszélünk s , p , d , ... primitív függvényekről. A Gauss-pályákat Samuel Francis Boys vezette be először a kvantumkémiába [32]. Előnyük, hogy a többcentrumú molekuláris integrálokat sokkal könnyebb Gauss-pályákkal meghatározni, mint Slater-pályákkal. Ez annak köszönhető, hogy két különböző atomon centrált Gauss-pálya szorzata, egy konstans faktortól eltekintve, egy harmadik centrumon lokalizált Gauss-pályát eredményez. A Gauss-pálya Descartes-alakja:

$$\phi_{l,m,n}^{GTO}(r;\alpha) = Nx^l y^m z^n e^{-\alpha r^2}, \quad (4)$$

ahol N a normálási tényező és α a GTO pályakitevője.

Míg a Slater-pályák a hidrogénszerű atompályákhoz hasonló tulajdonságokkal rendelkeznek, addig a Gauss-pályák jelentősen eltérnek azoktól. Így egy atompálya közelítésére elegendő egy Slater-pálya, a Gauss-pályákból azonban több szükséges. Tehát 1 STO közelítésére több primitív GTO lineáris kombinációját kell alkalmaznunk. Így kapjuk az összetett (kontrahált) GTO bázisfüggvényt (*CGTO*):

$$\chi^{STO}(r;\zeta) \approx \Phi^{CGTO} = \sum_{i=1}^k c_i g_i, \quad (5)$$

ahol k a kontrakció mértéke és a c_i skalárok az un. expanziós koefficiensek. Egy adott Slater-pálya előállításához ugyanolyan szimmetriájú primitív Gauss-pályákat kell alkalmaznunk.

A legegyszerűbb bázisok közé tartozik az STO- n G, melyben minden atompályát (STO-t) n darab primitív Gauss-függvény lineáris kombinációjából állítunk elő. Pople és mtsai. dolgozták ki a felhasított vegyérték (*split valance*) bázist, melyben a törzselektronok pályáit szokásos módon, azaz Gauss-pályák lineáris kombinációjaként állítjuk elő. A vegyértékelektronok pályáit azonban felhasítjuk 2, 3, vagy akár több részre (DZ, TZ, QZ,...) és azokra külön-külön báziskészletet rendelünk hozzá. Például: a 3-21G bázisban a törzselektronok pályáit 3 darab Gauss-függvény lineáris kombinációjaként állítjuk elő (az első szám erre vonatkozik). Mivel a kötőjel után két szám szerepel, így a vegyértékhéj DZ minőségű lesz, vagyis a vegyértékelektronokra 2

báziskészletet használunk. Az egyik 1 primitív Gauss-függvényből, a másik 2 primitív Gauss-függvény lineáris kombinációjából adódik. A bázisokat gyakran egészítjük ki polarizációs és diffúz függvényekkel. A polarizációs függvények, olyan bázis-függvények, melyekhez tartozó $[\lambda] = l + m + n$ nagyobb, mint az atom legkülső betöltött pályájának mellék-quantumszáma. Ezek a függvények a bázist flexibilisebbé teszik, így javítják az elektronszerkezet leírását. A polarizációs függvényeket *-gal vagy zárójelben megfelelő betűkkel jelöljük a bázis neve után. Egyre nagyobb teret hódítanak, az ún. Dunning-féle korrelációs konzisztens bázisok. Ezek a bázisok már beépítve tartalmazzák a polarizációs függvényeket. A diffúz függvények kisebb $[\lambda]$ értékűek és kis exponensűek, + jellel szokták őket jelölni Pople-féle bázisok esetén. A Dunning-féle bázisoknál az *aug*- előtag jelöli diffúz függvények jelenlétét. A számításokhoz használt bázist a módszer után szoktuk feltüntetni: HF/3-21++G**, CCSD(T)/aug-cc-pV5Z.

A szemempirikus kvantumkémiai és a G_n (G1-G4) számítások során a bázis-függvényeket nem kell megadni, mivel a bázisok be vannak építve („drótozva”) a módszerekbe. Bázisfüggvényeket természetesen a mai napig is fejlesztenek, melyek az EMSL Basis Set Exchange (bse.pnl.gov/bse/portal) weboldalon érhetők el. Itt körülbelül 580 bázis található. Meg kell jegyeznünk, hogy az említett cc-aug-pVnZ ($n=1-5$) bázisfüggvények és a G_n módszerek sajnos csak a kriptonig alkalmazhatóak.

3.3.2 Hartree-Fock-Roothaan-Hall modell [33]

A kvantumkémia kezdete óta az elektronok mozgására vonatkozó Schrödinger-egyenlet közelítő megoldása a cél. Az esetek többségénél (a legegyszerűbbeket kivéve) több-elektron problémával állunk szemben. A közelítő megoldások kiindulási pontja (az *ab initio* módszerek esetén) a Hartree-Fock közelítés, melynek lényege, hogy az elektron-elektron taszítást csak átlagolva vesszük figyelembe, így a sokelektronos esetet visszavezethetjük effektív egyelektronos problémára.

A molekuláris sokrészecskés elektron-hullámfüggvényt általában egyrészecske-spinpályákból állítjuk elő. Mivel az elektron térbeli mozgása és spinje között nincsen csatolás, a két mozgást szeparálhatjuk:

$$\varphi(\mathbf{x}) = \psi(\mathbf{r})\sigma(\omega), \quad (6)$$

ahol \mathbf{x} az elektron összes koordinátája (tér és spin), \mathbf{r} a helyvektor és ω a formális spinkoordináta, így a $\psi(\mathbf{r})$ a térbeli pálya és a $\sigma(\omega)$ a spinfüggvény.

Az elektronnak – a spinjéből adódóan – két állapota lehetséges (*spin-up* és *spin-down*):

$$\varphi(\mathbf{x}) = \begin{cases} \psi(\mathbf{r})\alpha(\omega) \\ \psi(\mathbf{r})\beta(\omega) \end{cases}. \quad (7)$$

Ezeket a spinpályákat spin-megkötéses (*restricted*) pályáknak hívják. Zárt héjú rendszerről beszélünk, ha a térbeli pályákon két-két elektron van, nyílt héjú rendszerről, ha van olyan pálya, amelyiken csak egy elektron helyezkedik el.

Írjuk fel a Hamilton-operátort, egy N -elektronos rendszerre, melyben az elektronok között nincs kölcsönhatás:

$$\hat{H} = \sum_{i=1}^N h(i). \quad (8)$$

A $h(i)$ operátort effektív egyelektron-operátornak nevezzük, mely tartalmazza az adott elektron kinetikus és potenciális energiáját. A sajátfüggvényei az alábbi $\{\varphi_j\}$ spinpályák lesznek:

$$h(i)\varphi_j(\mathbf{x}_i) = \varepsilon_j\varphi_j(\mathbf{x}_i). \quad (9)$$

Így az elektron-hullámfüggvény szorzatként állítható elő, melyet *Hartree-szorzatnak* nevezünk:

$$\psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = \varphi_i(\mathbf{x}_1)\varphi_j(\mathbf{x}_2)\dots\varphi_k(\mathbf{x}_N). \quad (10)$$

Azonban ha bármely két elektron koordinátáját felcseréljük, a (10) hullámfüggvény előjele nem változik meg. Vagyis ez az alak nem felel meg a fermionokra vonatkozó antiszimmetria elvnek (Pauli-elv). Ennek kiküszöbölésére a hullámfüggvényt Slater-determináns formájában szoktuk felvenni:

$$\psi(x_1, x_2, \dots, x_N) = (N!)^{-1/2} \begin{vmatrix} \varphi_i(x_1) & \varphi_j(x_1) & \cdots & \varphi_k(x_1) \\ \varphi_i(x_2) & \varphi_j(x_2) & \cdots & \varphi_k(x_2) \\ \vdots & \vdots & & \vdots \\ \varphi_i(x_N) & \varphi_j(x_N) & \cdots & \varphi_k(x_N) \end{vmatrix}, \quad (11)$$

ahol $(N!)^{-1/2}$ un. normálási tényező.

A rendszer energiáját következőképpen kapjuk meg:

$$E = \langle \psi | \hat{H} | \psi \rangle. \quad (12)$$

A variációs elv kimondja, hogy a közelítő hullámfüggvénnyel számított energia felső korlát az egzakt energiára nézve. Egyenlőség csak akkor áll fenn, ha a közelítő hullámfüggvény megegyezik az egzakt hullámfüggvénnyel:

$$\tilde{E} \geq E_{\text{egzakt}}. \quad (13)$$

Az Eckart-tétel értelmében az a legjobb hullámfüggvény, melynél a legkisebb közelítő energiát kapjuk, mivel

$$\text{Ha } \tilde{E} \rightarrow E_0, \text{ akkor } \tilde{\psi} \rightarrow \psi_0. \quad (14)$$

Ahhoz, hogy az optimális spinpályákat megkapjuk, az E_0 -t minimalizálnunk kell a spinpályák szerint, melyből az alábbi egyenletrendszert kapjuk:

$$\left[h(i) - \sum_{b \neq a} \hat{K}_b(i) + \sum_{b \neq a} \hat{J}_b(i) \right] \varphi_a(i) = \varepsilon_a \varphi_a(i), \quad (15)$$

ahol \hat{J}_b a Coulomb-operátor, \hat{K}_b a kicserélődési-operátor. Bizonyítható, hogy:

$$\left[\hat{J}_a(i) - \hat{K}_a(i) \right] \varphi_a(i) = 0. \quad (16)$$

A (15) és a (16) egyenleteket figyelembe véve, definiálható a Fock-operátor:

$$f(i) = h(i) + \sum_b \hat{J}_b(i) - \hat{K}_b(i) = h(i) + v^{HF}(i), \quad (17)$$

ahol $h(i)$ az egyelektron Hamilton-operátor, $v^{HF}(i)$ az effektív egyelektron potenciál-operátor. Így a Hartree-Fock egyenletek általános alakja:

$$f(i)\varphi(x_i) = \varepsilon\varphi(x_i). \quad (18)$$

Azonban a Fock-operátor függ a sajátértékeitől, mivel a Hartree-Fock potenciál $v^{HF}(i)$ értéke függ a többi elektron spinpályájától. Vagyis a Hartree-Fock egyenletek megoldásához iteratív eljárás alkalmazása szükséges, melyet önkonzisztens tér (*self-consistent-field*, *SCF*) módszernek nevezünk. Lényege, hogy a kezdeti spinpályákból kiindulva egy átlagos $v^{HF}(i)$ értéket számolunk az egyes elektronokra, majd megoldjuk a sajátértékegyenleteket, melyekből új spinpályákat nyerünk. Ezeket felhasználva megismételjük az eljárást addig, amíg a Fock-operátort felépítő spinpályák nem egyeznek meg az újonnan számítottakkal.

A sajátértékegyenletek megoldásával a Hartree-Fock spinpályákat $\{\varphi_k\}$ és a megfelelő pályae energiákat $\{\varepsilon_k\}$ kapjuk. Az N db spinpálya közül a legalacsonyabb energiájúak lesznek a betöltött pályák, míg a fennmaradóak a betöltetlen (*virtuális*) pályák. A Hartree-Fock egyenleteket mindig véges bázison oldjuk meg. Ahogy növeljük a bázisfüggvények számát, a Hartree-Fock energia csökken; a határértékét Hartree-Fock (limit) energiának nevezzük.

3.3.3 Csatolt klaszter (CC) módszer

A Hartree-Fock szinten számított energiaértékek nem egyeznek meg a rendszer tényleges energiájával. A Hartree-Fock módszerben a hullámfüggvény csak egy determinánssal van közelítve, ami nem azonos az egzakt hullámfüggvénnyel. A fő probléma azonban mégis abból származik, hogy a módszer az elektronok kölcsönhatását átlagos elektrontaszítási potenciállal veszi figyelembe. Az elektronok közvetlenül befolyásolják egymás mozgását, vagyis az elektronok mozgása korrelált. Az egzakt (nem-relativisztikus) energia és a Hartree-Fock-limithez tartozó energia különbséget ezért korrelációs energiának nevezzük. A kvantumkémia magasabb rendű módszereinek ezt az elektronkorrelációt valamilyen módon számítaniuk kell. A korrelációs kölcsönhatás számítására a Møller-Plesset-féle perturbációs (MP), a konfigurációs kölcsönhatás (CI) és a csatolt klaszter (CC) módszereket használhatjuk.

A csatolt klaszter módszer során a hullámfüggvényt az alábbi alakban vesszük fel:

$$\Psi = e^{\hat{T}} \Phi_0, \quad (19)$$

ahol \hat{T} a gerjesztő operátor ($\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 \dots$) és Φ_0 referencia hullámfüggvény. A \hat{T}_1 és \hat{T}_2 operátor hatása a Φ_0 -ra:

$$\hat{T}_1 \Phi_0 = \sum_i \sum_a t_i^a \Phi_i^a, \quad (20)$$

$$\hat{T}_2 \Phi_0 = \sum_i \sum_{j>i} \sum_a \sum_{b>a} t_{ij}^{ab} \Phi_{ij}^{ab}. \quad (21)$$

A (19) egyenlet sorbafejtése után az alábbi formulát kapjuk:

$$\Psi = e^{\hat{T}} \Phi_0 = \left(1 + \hat{T} + \frac{1}{2} \hat{T}^2 + \frac{1}{3!} \hat{T}^3 + \dots \right) \Phi_0 \quad (22)$$

Behelyettesítve ($\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 \dots$)-t:

$$\Psi = \left(1 + \hat{T}_1 + \hat{T}_2 + \frac{1}{2} \hat{T}_1^2 + \hat{T}_3 + \hat{T}_1 \hat{T}_2 + \frac{1}{3!} \hat{T}_1^3 + \hat{T}_4 + \frac{1}{2} \hat{T}_2^2 + \dots \right) \Phi_0 \quad (23)$$

A CCSD szinten a \hat{T}_1 és \hat{T}_2 gerjesztéseket vesszük figyelembe, a CCD-nél csak a \hat{T}_2 gerjesztéseket. A CCSDT szinten már \hat{T}_1 , \hat{T}_2 , és \hat{T}_3 , azaz a háromszoros gerjesztések is megtalálhatók. A CCSD(T) módszer napjainkban a kvantumkémia arany módszere, melynél a háromszoros gerjesztést nem iteratív, hanem perturbatív módon veszik figyelembe [34].

3.3.4 Szemiempirikus módszerek

Közepes (~50 atom), és nagyobb molekulákra napjainkban is végeznek szemiempirikus kvantumkémiai számításokat [35]. A MOPAC (Molecular Orbital Package) egy szemiempirikus kvantumkémiai programcsomag, amelynek segítségével tanulmányozni lehet molekulákat, atomokat, lineáris polimereket, stb. és azok kémiai reakcióit [36]. Több módszer is elérhető a MOPAC-on belül, mint például: MNDO, MINDO/3, AM1 (SAM1) és a PM3 (PM6 és PM7). Ezek mind SCF-típusú módszerek, melyek – megfelelő közelítésekkel – figyelembe veszik az elektronok közötti elektrostatikus tasztási és kicserélődési kölcsönhatásokat. Felhasználják atomok és molekulák kísérletileg meghatározott adatait is, mint például a legstabilabb izotópok atomtömegét, képződéshőket, dipólusmomentumokat és ionizációs energiákat.

Az NDDO-alapú (Neglect of Diatomic Differential Overlap) MNDO (Modified Neglect of Diatomic Overlap) módszert elsősorban ionizációs energia, dipólusmomentum és molekulák szerkezetének meghatározására alkalmazták [23,37]. 1985-ben közölték az AM1 (Austin Model 1) módszert, amit a PM3 (Parametric Method number 3) követett. A PM3 módszer kifejlesztése után három nagyobb probléma merült fel: a mag-mag kölcsönhatásokat leíró közelítések módosítása; egy teljesen szisztematikus paraméter-optimalizáció a főcsoport elemeire (nagyobb hangsúlyt fektetve ezáltal a biológiailag fontos vegyületekre); és a módszer kiterjesztése az átmeneti fémekre. Ezen fejlesztések eredményeképpen jött létre a PM6 módszer, melyben négylépcsős optimalizációval 70 elemet paramétereztek. Az irodalmi adatok a WebBook termokémiai és a Cambridge Structural Database (CSD) molekulageometriai adatbázisokból származnak.

A PM6 módszer a korábbi szemiempirikus módszerekhez képest a fentiekből adódóan megbízhatóan jobban teljesít. A korábbi módszerekkel elvégezve ugyanazokat a számításokat (képződéshők, kötéstávolságok, kötésszögek, dipólusmomentumok és ionizációs energiák) majd összevetve az irodalmi értékekkel, az eltérés a PM6 módszer esetében a legkisebb. A módszer publikálása után azonban kiderült néhány hiba, ilyen például, hogy a szilárd anyagok esetén (pl. kristályszerkezetek) nagyobb számítási hibát produkál; az elektrostatikus tasztást nem veszi sok atom esetén megfelelően figyelembe; továbbá az, hogy az NDDO formalizmus is módosításra szorul. Meg kell jegyeznünk, hogy a valós kémiai kísérletek nagy része nem gázfázisban játszódik le, így

a szilárd anyagok számítása során keletkező hiba korrigálása megkülönböztetett figyelmet igényel.

Az intermolekuláris kölcsönhatások becslésének pontosítása érdekében két új módszert fejlesztettek ki: a PM6-DH+-t és a PM6-DH2-t. Azonban ezek a módszerek alkalmatlannak bizonyultak képződéshők pontos számítására. További fejlesztések eredményeképpen (az említett hibák korrigálása mellett) kidolgozták a PM7 módszert. Ebben említésre méltó előrelépés volt még a nem-kovalens kölcsönhatások (diszperziós kölcsönhatás, hidrogénkötés, halogénkötés) leírásának pontosítása, aminek segítségével dimerek (pl. víz, ammónia, benzol) kötése erősségének számítása jobb eredményre vezetett.

3.4 Fémionok képződéshője

Az atomok és ionok elektronkonfigurációjának meghatározása kvantummechanikai alapon történt. A Bohr-féle atommodell [38] volt az első, mely képes volt értelmezni elméletileg az elemek periódusos rendszerét. A modell alapja, hogy az elektronok az atommag körül ún. héjakba rendeződnek, s a héjak alhéjakból épülnek fel. Minden elektronhoz négy kvantumszám tartozik: n (főkvantumszám), l (mellékquantumszám), m_l (mágneses kvantumszám) és m_s (spinkvantumszám). Minden héjhoz egy adott n főkvantumszám ($n = 1, 2, 3, \dots$) tartozik, melyen maximálisan $2n^2$ elektron helyezkedhet el. Az alhéjak megkülönböztetése az l értéke ($l = 0, 1, 2, \dots, n - 1$) alapján lehetséges ($l=0$: s , $l=1$: p , $l=2$: d , $l=3$: f). A Pauli-elv szerint egy atomban két elektron nem lehet ugyanabban az energiaállapotban, azaz nem lehet mind a négy kvantumszáma azonos. Így egy adott l értékhez $2l + 1$ állapot (cella) tartozik, amelyek $2(2l + 1)$ elektron befogadására képesek. Minden ilyen cellában két elektron helyezkedhet el, ellentétes spinnel.

Minden elektronkonfigurációnak több elektronállapot, azaz több spektroszkópai term felel meg [39]. Ezeket a termeket akár egyszerű kombinatorikai eszközök segítségével is meghatározhatjuk. Fontos, hogy a betöltött héjakat, alhéjakat nem vesszük figyelembe. Az alapállapot, azaz a legmélyebb energiájú állapot meghatározásához kvantumkémiai számítások szükségesek. Azonban kisebb rendszámú elemek esetén egyszerű elvek is alkalmazhatók: felépítési elv, Pauli-elv és Hund-szabályok. Az felépítési elv (más néven az energiaminimumra való törekvés elve) lényege, hogy az atomi egyrészecske állapotok a legalacsonyabb energiájútól sorban

felfelé töltődnek fel. A Hund-szabályok közül a legfontosabb a maximális spinmultiplicitás elve.

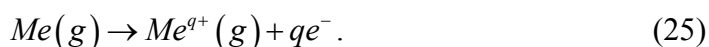
Az atomi spektroszkópai termék jelölésére a következő szimbólum használatos:

$${}^{2S+1}L_J \quad (24)$$

Ahol S a teljes spinquantumszám, L a teljes mellékkquantumszám és J a teljes (impulzusmomentum) quantumszám. (Egy elektronra a szimbólumok: s, l, j .) Az atomi állapotokat következőképpen jelöljük az L értékétől függően: az $L = 0, 1, 2, 3, 4, \dots$ értékeknek rendre az S, P, D, E, F, ... szimbólumok felelnek meg. Pl. a 3P_0 szimbólum egy olyan állapotot jelöl, amelyre $S = 1$, $L = 1$ és $J = 0$. Az $M_S = 2S + 1$ számot spinmultiplicitásnak, vagy egyszerűen multiplicitásnak nevezzük. Az $M_S = 1, 2, 3, 4, 5, 6, 7, \dots$ értékeknek megfelelően szingulett, dublett, tripllett, kvartett, kvintett, szekstett, szeptett, ... állapotokról beszélünk.

Zárthéjú rendszerek esetén az elektronok párosítva vannak, így $S = 0$, azaz az állapot mindig szingulett. A megfelelő spektroszkópai term pedig mindig 1S_0 lesz. Ilyen esetben valamilyen spin-korlátozott (restricted, R) kvantumkémiai számítást alkalmazhatunk. Nyíltéjú rendszerek esetén meg kell adnunk a spinmultiplicitás értékét, és valamilyen nem spin-korlátozott (unrestricted, U) kvantumkémiai számítást végzünk. Minden komolyabb kvantumkémiai számítás első lépéseként általában RHF (restricted Hartree-Fock) vagy UHF (unrestricted Hartree-Fock) számítást végzünk. A kvantumkémiai programok inputjában pontosan meg tudjuk adni a rendszer töltését és spinmultiplicitását.

Tekintsük egy izolált fématom ionizációját (q darab elektront távolítunk el):



Ennek megfelelően a fémion standard képződéshője:

$$\Delta H_{Me^+} = \Delta H_{Me} + IE , \quad (26)$$

ahol ΔH_{Me} a fématom standard képződéshője és IE az elektron(ok) eltávolításához szükséges ionizációs energia. A ΔH_{Me} és IE adatokat megfelelő kvantumkémiai elméleti szinten számíthatjuk, vagy kísérleti adatokat [40] is felhasználhatunk. Meg kell említenünk, hogy kvantumkémiai módszerekkel – a hibák kiegyenlítődése miatt – általában pontosabban tudunk számítani energiakülönbségeket, mint egyedi energiaértékeket.

3.5 Paraméterbecslés a legkisebb négyzetek módszerével [16,41,42]

A modellalkotásnak egy vitathatatlanul fontos pontja a meghatározott kísérleti adatok pontos kiértékelése. Az eljárás során fontos szerepet kap a paraméterbecslés; azon belül is a lineáris és nemlineáris legkisebb négyzetek módszerének alkalmazása a legközkedveltebb.

Tekintsünk m darab (X_i, Y_i) kísérleti adópárt, ahol Y_i a függő és X_i a független változó mért értéke. Legyen adott egy $f(x, \mathbf{p})$ függvénnyel megadott matematikai modell, ahol x a független változó pontos értéke és \mathbf{p} pedig az n darab meghatározandó paramétert tartalmazó vektor. Feltételezzük a következő összefüggések fennállását:

$$X_i = x_i + \varepsilon_{xi} \quad (28)$$

$$Y_i = y_i + \varepsilon_{yi} \quad (29)$$

$$y_i = f(x_i; \mathbf{p}), \quad (30)$$

ahol x_i és y_i az X_i és az Y_i mérési adatok pontos értéke, ε_{xi} és ε_{yi} valószínűségi változók pedig kísérleti adatok hibái. A legkisebb négyzetek (LS) módszerének alkalmazásakor általában feltételezzük, hogy az ε_{xi} és ε_{yi} hibák Gauss-eloszlásúak és korrelálatlanok.

Az optimális paraméterek meghatározásához – első közelítésben – az alábbi célfüggvényt minimalizáljuk:

$$S(\mathbf{p}) = \sum_{i=1}^m w_i \varepsilon_{yi}^2(\mathbf{p}) = \sum_{i=1}^m w_i (Y_i - f(X_i, \mathbf{p}))^2, \quad (31)$$

ahol a w_i súlyfaktorok következőképpen számítandók, ha az X_i értékekben kicsi a hiba:

$$w_i = \frac{1}{\sigma_{Y_i}^2 + \left(\frac{\partial f(x, \mathbf{p})}{\partial x} \Big|_{X_i} \right)^2 \sigma_{X_i}^2}. \quad (32)$$

Azt a paraméterbecslési eljárást, mely az előbbi célfüggvényt használja, általános legkisebb négyzetek módszerének (GLS) nevezzük. A súlyozott legkisebb négyzetek módszeréhez (WLS) a $w_i = \sigma_{Y_i}^{-2}$ súlyfaktorok szükségesek. Abban az esetben, ha sem a függő (Y), sem a független változó (X) hibája nem ismert, w_i -t 1-nek választjuk (közönséges legkisebb négyzetek módszere, OLS).

Tegyük fel, hogy $\hat{\mathbf{p}}$ az optimális paramétervektor, azaz ahol az $S(\mathbf{p})$ függvénynek globális minimuma van a paraméterek (fizikai) értelmezési tartományán.

Ebben az esetben az illesztés jóságára jellemző reziduális variancia (s_{r_2}) következőképpen adható meg:

$$s_{r_2} = \frac{S(\hat{\mathbf{p}})}{m-n} \quad (33)$$

A \mathbf{V} variancia-kovariancia mátrix szükséges a becsült paraméterek sztenderd hibájának és korreláltságának a számításához:

$$\mathbf{V} = 2s_{r_2} \mathbf{H}^{-1}(\hat{\mathbf{p}}), \quad (34)$$

ahol $\mathbf{H}^{-1}(\hat{\mathbf{p}})$ a $\hat{\mathbf{p}}$ globális minimumban számított Hesse-mátrix inverze. Ha $\mathbf{V} = \{v_{ij}\}$, akkor a p_i paraméter $SE(p_i)$ standard hibáját következőképpen kaphatjuk meg:

$$SE(p_i) \approx 2\sqrt{v_{ii}}. \quad (35)$$

Ez jól megfelel a 95%-os szignifikancia-szinten számított konfidencia-intervallumnak.

A paraméterbecsléshez szükségünk van egy jó minimalizáló módszerre, a \mathbf{H} mátrixra és inverzére. A minimalizálást szimplex-módszerrel, az első és a második deriváltak számítását centrális differencia formulákkal, az invertálást pedig LU-faktorizációval oldottuk meg.

Ha az F mennyiség a $\hat{\mathbf{p}}$ függvénye, azaz $F = \varphi(\hat{\mathbf{p}})$, akkor F szórása következőképpen számítható ki:

$$SE(F) = \sqrt{(\text{grad } F)^T \mathbf{V} \text{grad}(F)}. \quad (36)$$

Ezzel a formulával kezelhető a hibaterjedés, azaz a hibás mennyiségekből származtatott mennyiségek hibáinak a számítása.

3.6 Lineáris és multilineáris regresszió

Adataink kiértékelésnél körültekintően kell kiválasztani a kiértékeléshez legmegfelelőbb eljárást, hiszen érékes statisztikai információkat veszíthetünk a nem megfelelő módszer alkalmazásával. A regressziós módszerek közül adataink kiértékeléséhez a teljes legkisebb négyzetek módszerére (TLS) épülő multilineáris regressziót (MLR) alkalmaztuk.

Egyszerű lineáris regresszió esetén egyetlenegy regressziós változónk (független változónk) van, melynek az $\{X_i\}$ értékeit pontosan ismerjük, azaz $X_i = x_i$. Az egyetlen függő változónk $\{Y_i\}$ értékei véletlenszerű hibával terheltek: $Y_i = y_i + \varepsilon_i$. Az y_i a

pontos érték, és az $\{\varepsilon_i\}$ hibák azonos eloszlásúak és korrelálatlanok, és a hibák várható értéke nulla és a szórásuk azonos.

A kétparaméteres lineáris regresszió matematikai modellje a következő:

$$Y_i = aX_i + b + \varepsilon_i, \quad (37)$$

$$y_i = aX_i + b, \quad (38)$$

ahol a és b az ismeretlen paraméterek. Az LS módszer filozófiája szerint az optimális paraméter-értékek minimalizálják a következő célfüggvényt:

$$S(a, b) = \sum_{i=1}^m (Y_i - y_i)^2 = \sum_{i=1}^m (Y_i - aX_i - b)^2, \quad (39)$$

A paraméterek becsült optimális értékével (\hat{a}, \hat{b}) a célfüggvény értéke minimális lesz. A becsült értékek mellett azonban azok megbízhatóságára is kíváncsiak vagyunk. Ennek jellemzésére adott megbízhatósági szinten megbízhatósági (konfidencia) intervallumot határozunk meg. Az x és az y változók közötti lineáris kapcsolat mértékének a jellemzésére jól használható az r_{xy} un. lineáris korrelációs (Pearson-féle) koefficiens. Teljesül rá a következő egyenlőtlenség: $-1 \leq r_{xy} \leq 1$; tökéletesen lineáris kapcsolat esetén az értéke ± 1 .

Az illesztés jóságát jellemző legfontosabb adat az s_r maradék (reziduális) szórás vagy más néven maradék hiba:

$$s_{r2} = \frac{S_{\min}}{m-2} = \frac{\sum_{i=1}^m (Y_i - \hat{y}_i)^2}{m-2} = \frac{\sum_{i=1}^m e_i^2}{m-2}, \quad (40)$$

$$s_r = \sqrt{s_{r2}}, \quad (41)$$

A szabadsági fokok száma $(m-2)$, mivel két paraméter értékét mérési adatokból becsüljük. Az e_i maradék pedig az ε_i hiba becsült értéke.

Gyakran előfordul, hogy a kísérleti adataink között un. kiugró pont van, azaz olyan pont, amely a többi ponttól eltérő viselkedésű, s ki kellene hagynunk a kiértékelésből. A kiugró pontok diagnosztizálását feltétlenül meg kell oldani a kísérleti adatok kiértékelése során. Természetesen, a kiugró pont elhagyása helyett a kísérletet megismételhetjük, s az új adattal helyettesíthetjük a régit.

Multilineáris regressziós modell esetén az adataink következőképpen néznek ki:

$$\begin{array}{ccccccc}
\mathbf{Y} & \mathbf{X}_1 & \mathbf{X}_2 & \dots & \mathbf{X}_k & & \\
Y_1 & X_{11} & X_{12} & & X_{1k} & & \\
Y_2 & X_{21} & X_{22} & & X_{2k} & & \\
\vdots & & & & & & \\
Y_m & X_{m1} & X_{m2} & & X_{mk} & &
\end{array} \quad (42)$$

Minden sor egy adatpontnak felel meg, és feltételezzük, hogy

$$Y_i = y_i + \varepsilon_i = a_0 + \sum_{j=1}^k a_j X_{ij} + \varepsilon_i. \quad (43)$$

Ez azt jelenti, hogy a modellünk a paraméterekben lineáris. Mátrixalakban is felírhatjuk a kérdéses modellt:

$$\mathbf{Y} = \mathbf{X}\mathbf{a} + \boldsymbol{\varepsilon}, \quad (44)$$

ahol

$$\mathbf{Y} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_m \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} 1 & X_{11} & X_{12} & \dots & X_{1k} \\ 1 & X_{21} & X_{22} & & X_{2k} \\ \vdots & & & & \\ 1 & X_{m1} & X_{m2} & \dots & X_{mk} \end{pmatrix}, \quad (45)$$

$$\mathbf{a} = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_k \end{pmatrix}, \quad \boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_m \end{pmatrix}. \quad (46)$$

OLS/WLS esetben az \mathbf{Y} és az $\boldsymbol{\varepsilon}$ oszlopvektorok valószínűségi vektorváltozók, s feltételezzük, hogy a független változók értékei pontosan ismertek, továbbá a hibák véletlenszerűek, függetlenek, és szórásuk konstans. OLS esetben az utóbbiak még azonosak is. A paramétereink száma: $n = k + 1$ (k a regressziós koefficiensek száma). Az LS módszer filozófiájának megfelelően, a paraméterek (regressziós koefficiensek és a regressziós állandó) meghatározása érdekében a következő célfüggvényt kell minimalizálnunk:

$$S(\mathbf{a}) = \boldsymbol{\varepsilon}^T \cdot \boldsymbol{\varepsilon} = (\mathbf{Y} - \mathbf{X}\mathbf{a})^T (\mathbf{Y} - \mathbf{X}\mathbf{a}) = \sum_{i=1}^m \left[Y_i - \left(a_0 + \sum_{j=1}^k a_j X_{ij} \right) \right]^2 = \sum_{i=1}^m e_i^2 \quad (47)$$

(WLS esetben a célfüggvény még tartalmazni fog egy \mathbf{W} súlymátrixot is.) A minimum szükséges feltételéből

$$\frac{\partial}{\partial \mathbf{a}} S(\mathbf{a}) = \frac{\partial}{\partial \mathbf{a}} [(\mathbf{Y} - \mathbf{X}\mathbf{a})^T (\mathbf{Y} - \mathbf{X}\mathbf{a})] = \mathbf{0} \quad (48)$$

kapjuk, hogy:

$$\hat{\mathbf{a}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}. \quad (49)$$

Az \mathbf{X} mátrix az un. modell-mátrix (tervezési mátrix). Ennek megfelelően a számított értékek:

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\mathbf{a}} = \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y} = \mathbf{H}\mathbf{Y}, \quad (50)$$

$$\mathbf{H} = \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T, \quad (51)$$

A \mathbf{H} mátrixot kalap mátrixnak nevezik az angol szakirodalomban, ugyanis ennek segítségével kapjuk meg az \mathbf{Y} mérési adatokból az $\hat{\mathbf{y}}$ számított értékeket. Az $\mathbf{e} = \mathbf{Y} - \hat{\mathbf{y}}$ maradékok pedig következőképpen néznek ki:

$$\mathbf{e} = \mathbf{Y} - \hat{\mathbf{y}} = \mathbf{Y} - \mathbf{H}\mathbf{Y} = (\mathbf{1} - \mathbf{H})\mathbf{Y}. \quad (52)$$

Az illesztés maradék (reziduális) szórásnégyzete:

$$s_{r,2} = \frac{\mathbf{e}^T \cdot \mathbf{e}}{m-n} = \frac{(\mathbf{Y} - \mathbf{X}\hat{\mathbf{a}})^T (\mathbf{Y} - \mathbf{X}\hat{\mathbf{a}})}{m-n} = \sum_{i=1}^m \frac{(Y_i - \hat{y}_i)^2}{m-n}. \quad (53)$$

Multilineáris esetben a statisztikai mutatók lényegében ugyanúgy működnek, mint egyszerű lineáris esetben.

4. Eredmények

4.1 A modellek jóságának vizsgálata

Munkánk során először a Leszczynski és mtsai., illetve a Wu és mtsai. által vizsgált, összesen 52 fém-oxid nanorészecskéket alkotó izolált fématomokra végeztünk PM6 számításokat. (A számításokhoz a MOPAC 2016 programot (openmopac.net) használtuk.) Ellenőrizni akartuk, hogy a PM6 módszer visszaadja-e az atomok irodalomban közölt spinmultiplicitását. A számítási eredményeinket a 3. táblázatban tüntetjük fel az experimentális képződéshőkkel együtt.

Fém-oxid	Me	Z	ΔH_{Me} (kcal/mol) PM6	M_S	ΔH_{Me} (kcal/mol) Exp.
BeO	Be	4	76.96	1	77.44
MgO	Mg	12	35.00	1	35.16
Al ₂ O ₃	Al	13	74.07	2	78.87
SiO ₂	Si	14	108.38	3	107.55
CaO	Ca	20	42.60	1	42.50
Sc ₂ O ₃	Sc	21	69.42	2	90.30
TiO ₂	Ti	22	48.94	3	113.05
V ₂ O ₃	V	23	123.92	4	122.90
Cr ₂ O ₃	Cr	24	95.00	7	94.79
MnO ₂	Mn	25	67.70	6	67.09
Fe ₂ O ₃	Fe	26	99.30	5	99.50
CoO	Co	27	83.51	4	101.51
NiO	Ni	28	93.53	3	102.70
CuO	Cu	29	80.70	2	80.64
ZnO	Zn	30	31.17	1	31.17
Ga ₂ O ₃	Ga	31	65.40	2	65.01
GeO ₂	Ge	32	89.50	3	88.91
SrO	Sr	38	39.10	1	39.29
Y ₂ O ₃	Y	39	95.11	2	100.69
ZrO ₂	Zr	40	144.94	3	145.51
Nb ₂ O ₃	Nb	41	172.40	6	173.49
Mo ₂ O ₃	Mo	42	157.30	7	157.29

Fém-oxid	Me	Z	ΔH_{Me} (kcal/mol) PM6	M_S	ΔH_{Me} (kcal/mol) Exp.
TcO ₂	Tc	43	160.69	6	162.05
RuO ₂	Ru	44	182.54	5	153.61
Rh ₂ O ₃	Rh	45	127.17	2 (4)	133.10
PdO	Pd	46	90.00	1	90.39
Ag ₂ O	Ag	47	68.10	2	68.09
CdO	Cd	48	26.72	1	26.72
In ₂ O ₃	In	49	58.00	2	58.15
SnO ₂	Sn	50	72.20	3	71.99
Sb ₂ O ₃	Sb	51	63.20	4	62.69
BaO	Ba	56	58.76	1	43.02
La ₂ O ₃	La	57	75.39	2	103.01
HfO ₂	Hf	72	147.90	3	147.99
Ta ₂ O ₃	Ta	73	190.53	4	186.90
WO ₂	W	74	204.20	5	203.01
ReO ₂	Re	75	185.00	6	184.01
OsO ₂	Os	76	200.49	5	189.05
IrO ₂	Ir	77	148.43	4	159.01
PtO ₂	Pt	78	107.20	3	135.11
Au ₂ O ₃	Au	79	77.53	2	87.50
HgO	Hg	80	14.69	1	14.67
Tl ₂ O	Tl	81	5.76	2	43.55
PbO	Pb	82	46.62	3	46.65
Bi ₂ O ₃	Bi	83	50.10	4	49.50

3. táblázat: Számítási eredmények és kísérleti adatok [40] a fématomokra (kcal/mol).

Látható, hogy összesen 9 esetben szingulett az alapállapot, a többinél a spinmultiplicitás nyílthéjú atomi alapállapotra utal. A ródium esetén a PM6 módszerrel meghatározott spinmultiplicitás ($M_S = 2$) eltér a kísérletitől ($M_S = 4$). A tallium és a titán esetében tapasztalható a legnagyobb különbség a számított és az experimentális értékek között. Összességében elmondható, hogy az atomok standard képződéshője valamilyen módon be van építve a PM6 módszerbe.

Ezek után a Leszczynski és mtsai. által felállított nano-QSAR-modell (1. Modell) vizsgálata következett. A felállított modelljük adatait már a 2. táblázatban

bemutattuk. Az általuk – PM6 szemiempirikus módszerrel – fémionokra számított standard képződéshők jóságát akartuk megvizsgálni. A számítási eredményeinket, illetve az ionok experimentális standard képződéshőit a 4. táblázatban mutatjuk be.

Fém-oxid	Me ⁿ⁺	pMEC ₅₀	M _S	ΔH_{Me^+} Exp.	ΔH_{Me^+} 1. Modell	ΔH_{Me^+} PM6
Al ₂ O ₃	Al(III)	2.49	1	1307.12	1187.83	1187.83
Bi ₂ O ₃	Bi(III)	2.82	1	1192.12	1137.40	1137.40
CoO	Co(II)	3.51	4	677.22	601.80	589.30
Cr ₂ O ₃	Cr(III)	2.51	4	1344.95	1268.70	1250.18
CuO	Cu(II)	3.20	2	726.77	706.25	706.25
Fe ₂ O ₃	Fe(III)	2.29	6	1361.88	1408.29	1277.12
In ₂ O ₃	In(III)	2.81	1	1273.13	1271.13	1271.13
La ₂ O ₃	La(III)	2.87	1	928.88	1017.22	1017.22
NiO	Ni(II)	3.45	3	697.86	596.70	562.81
Sb ₂ O ₃	Sb(III)	2.64	1	1228.13	1233.06	1233.06
SiO ₂	Si(IV)	2.20	1	2485.84	1686.38	1686.38
SnO ₂	Sn(IV)	2.01	1	2221.55	1717.32	1717.32
TiO ₂	Ti(IV)	1.74	1	2215.31	1575.73	1575.73
V ₂ O ₃	V(III)	3.14	3	1291.50	1097.73	1054.30
Y ₂ O ₃	Y(III)	2.87	1	999.16	837.15	837.15
ZnO	Zn(II)	3.45	1	662.07	662.44	662.44
ZrO ₂	Zr(IV)	2.15	1	1922.65	1357.66	1357.66

4. táblázat: PM6 számítási eredmények és kísérleti adatok [40] a fémionokra (kcal/mol).

A képződéshő dimenziója minden esetben kcal/mol, az EC₅₀ koncentrációé mol/L. A 4. oszlopban az általunk a PM6 számítások során beállított spinmultiplicitások láthatók. Megállapítható, hogy Leszczynski és mtsai. (1. Modell) 5 ion esetén rosszul választották meg az alapállapot spinmultiplicitását: Co(II): dublett, Cr(III): dublett, Fe(III): dublett, Ni(II): szingulett, V(III): szingulett. Ebből adódóan a mi értékeinknél nagyobb képződéshő értékeket kaptak. Az eltérések mértéket az 5. táblázat utolsó oszlopában tüntettük fel; ezeknél az ionoknál „X” karakter szerepel. A PM6 képződéshőket összevetve a kísérleti értékekkel, elmondható, hogy 4 ion esetén az

eltérés nagyobb, mint 500 kcal/mol. Ezek a következők: Si(IV), Sn(IV), Ti(IV) és Zr(IV). Azaz az összes négyszeres pozitív töltésű ion képződéshője kb. 500-800 kcal/mol-lal kisebb a kísérleti értéknél. A Si(IV) esetén a legnagyobb az eltérés: 799.46 kcal/mol. Az 5. táblázat utolsó oszlopában ezeknél az ionoknál a „†” szimbólum szerepel az eltérés mértékével együtt.

Fém-oxid	Me ⁿ⁺	ΔH_{Me^+} Exp.	ΔH_{Me^+} 1. Modell	ΔH_{Me^+} PM6	Megjegyzés
CoO	Co(II)	677.22	601.80	589.30	X: 12.50
Cr ₂ O ₃	Cr(III)	1344.95	1268.70	1250.18	X: 18.52
Fe ₂ O ₃	Fe(III)	1361.88	1408.29	1277.12	X: 131.17
NiO	Ni(II)	697.86	596.70	562.81	X: 33.89
SiO ₂	Si(IV)	2485.84	1686.38	1686.38	†: 799.46
SnO ₂	Sn(IV)	2221.55	1717.32	1717.32	†: 504.23
TiO ₂	Ti(IV)	2215.31	1575.73	1575.73	†: 639.58
V ₂ O ₃	V(III)	1291.50	1097.73	1054.30	X: 43.43
ZrO ₂	Zr(IV)	1922.65	1357.66	1357.66	†: 564.99

5. táblázat: Az 1. Modell PM6 értékeinek hibái (X:számítási hiba, †: fizikailag irreális).

Megállapítható, hogy az 1. Modell tanítókészletében 4 ion (Fe(III), Si(IV), Sn(IV), Ti(IV)) deskriptorának értéke valamilyen oknál fogva (számítási vagy fizikai) irreális. Az eredeti validáló készlet 7 ionjából 5 ion (Co(II), Cr(II), Ni(II), V(III), Zr(IV)) deskriptorának az értéke hibával terhelt. Így a Leszczynski és mtsai. által felállított nano-QSAR-modell fizikai realitása megkérdőjelezhető, ugyanis 9 adat hibával terhelt, ami több mint 50%-át jelenti a modell alapjául szolgáló adathalmaznak. Ezek a hibák azt mutatják, hogy a PM6 standard képződéshőket le kell cserélni experimentális értékekre, vagy magasabb szintű kvantumkémiai módszer által számoltakra.

Végezzünk lineáris regressziókat a pMEC₅₀ értékek és a fémionokra vonatkozó, általunk számított PM6 képződéshők, illetve a rendelkezésünkre álló kísérleti standard képződéshők között. A 6. táblázatban tüntetjük fel az eredeti tanítókészletre (1. Modell) végzett kétparaméteres lineáris regressziók ($y_i = ax_i + b$) statisztikai adatait [16].

Módszer	m	\hat{a}	\hat{b}	r_{xy}	s_r	n
1. Modell	10	-0.50 ± 0.17	2.59 ± 0.16	-0.9221	0.2207	Ti
ΔH_{Me^+} PM6	10	-0.49 ± 0.18	2.59 ± 0.17	-0.9128	0.2329	0
ΔH_{Me^+} Exp.	10	-0.48 ± 0.19	2.59 ± 0.18	-0.9006	0.2479	Si

6. táblázat: Az eredeti tanítókészletre végzett kétparaméteres lineáris regressziók statisztikai mutatói.

A 6. táblázat első sora megadja a Leszczynski és mtsai. által kapott összefüggést ($-\lg(EC_{50}) = 2.59 - 0.50 \Delta H_{Me^+}$). Az utolsó oszlopban találhatóak a kiugró pontok: Si(IV) illetve Ti(IV). Távolítsuk el a Si(IV) iont a tanítókészletből, vagyis így kapunk egy redukált tanítókészletet, melyben összesen 9 ion szerepel. Végezzük el erre a készletre is a lineáris regressziókat a fenti modellek esetén. A statisztikai adatokat a 7. táblázatban tüntettük fel.

Módszer	m	\hat{a}	\hat{b}	r_{xy}	s_r	n
1. Modell	9	-0.51 ± 0.18	2.63 ± 0.17	-0.9296	0.2175	0
ΔH_{Me^+} PM6	9	-0.51 ± 0.19	2.63 ± 0.18	-0.9225	0.2277	0
ΔH_{Me^+} Exp.	9	-0.52 ± 0.15	2.63 ± 0.14	-0.9506	0.1831	0

7. táblázat: A redukált tanítókészletre végzett kétparaméteres lineáris regressziók statisztikai mutatói.

Itt már látható, hogy egyetlen egy kiugrópontunk sincs, vagyis a Si(IV) ionhoz tartozó EC_{50} koncentráció valamilyen oknál fogva hibásnak minősül, így szükséges az eltávolítása. Az is megfigyelhető, hogy az experimentális értékek jobban teljesítenek (3. sor), mint a PM6 számítási eredmények.

Ezt követően Wu és mtsai. által felállított nano-QSAR-modell (2. Modell) jóságát vizsgáltuk. Az általuk talált összefüggés a következő:

$$-\lg(EC_{50}) = 4.412 - 0.121 Z / r - 0.001 \Delta H_{Me^+}, \quad (54)$$

ahol a Z/r az egyes fémionok polarizációs ereje (Z az ionok töltése és r a rádiuszuk).

A szerzők a MOPAC 2012-es programcsomagot használták a PM6 standard képződéshők számítására. A 8. táblázatban mutatjuk be Wu és mtsai. adatait (2. Modell), az általunk számítottakat, illetve az experimentális képződéshőket is. A képződéshő dimenziója minden esetben kcal/mol itt is. A számításaink során a MOPAC 2016 programot használtuk. A táblázat 4. oszlopában található a megfelelő fémion – PM6

módszer által adódó – spinmultiplicitása. A Si(IV) ion a szerzők számára is kiugrópontnak adódott, így a teljes adathalmazuk 16 fém-oxidot tartalmazott, a maradék 35 toxicitására predikciót hajtottak végre (becslési adathalmaz).

Fém-oxid	ΔH_{Me^+} 2. Modell	ΔH_{Me^+} PM6	M_S	Δ	ΔH_{Me^+} Exp.
Al ₂ O ₃	1187.83	1187.83	1	0.00	1307.12
Bi ₂ O ₃	1137.40	1137.40	1	0.00	1192.12
CoO	594.59	589.30	4	-5.29	677.22
Cr ₂ O ₃	1266.62	1250.18	4	-16.44	1344.95
CuO	713.74	706.25	2	-7.49	726.77
Fe ₂ O ₃	1363.40	1277.12	6	-86.28	1361.88
In ₂ O ₃	1271.13	1271.13	1	0.00	1273.13
La ₂ O ₃	1017.22	1017.22	1	0.00	928.88
NiO	596.88	562.81	3	-34.07	697.86
Sb ₂ O ₃	1233.06	1233.06	1	0.00	1228.13
SnO ₂	1717.32	1717.32	1	0.00	2221.55
TiO ₂	1575.73	1575.73	1	0.00	2215.31
V ₂ O ₃	1097.73	1054.30	3	-43.43	1291.50
Y ₂ O ₃	837.15	837.15	1	0.00	999.16
ZnO	662.44	662.44	1	0.00	662.07
ZrO ₂	1357.66	1357.66	1	0.00	1922.65
BeO	657.34	657.34	1	0.00	712.39
MgO	543.10	543.10	1	0.00	558.21
CaO	407.29	407.29	1	0.00	457.24
SrO	413.87	413.87	1	0.00	424.98
BaO	388.71	388.71	1	0.00	393.9
Sc ₂ O ₃	874.35	874.35	1	0.00	1107.68
HfO ₂	1555.34	1555.34	1	0.00	1957.21
Nb ₂ O ₃	1312.47	1197.87	3	-114.60	1229.64
Ta ₂ O ₃	1368.16	1304.71	3	-63.45	-
Mo ₂ O ₃	1306.84	1322.06	2	15.22	1319.14
WO ₂	2005.58	2000.59	3	-4.99	-
MnO ₂	1602.05	1601.92	2	-0.13	2556.29

Fém-oxid	ΔH_{Me^+} 2. Modell	ΔH_{Me^+} PM6	M_S	Δ	ΔH_{Me^+} Exp.
TcO ₂	2114.98	1881.79	2	-233.19	-
ReO ₂	1780.23	1780.28	4	0.05	-
RuO ₂	1674.84	1402.19	3	-272.65	-
OsO ₂	1757.52	1704.35	5	-53.17	-
Rh ₂ O ₃	960.20	937.78	5	-22.43	1438.3
IrO ₂	1410.10	1450.00	6	39.90	-
PdO	657.34	620.93	3	-36.41	730.71
PtO ₂	2165.13	2081.33	5	-83.80	-
Ag ₂ O	233.51	233.51	1	0.00	242.8
Au ₂ O ₃	1289.31	1270.77	3	-18.54	-
CdO	548.22	548.22	1	0.00	624.04
HgO	679.86	679.86	1	0.00	687.91
Ga ₂ O ₃	1384.15	1384.15	1	0.00	1385
Tl ₂ O	207.75	207.75	1	0.00	184.41
GeO ₂	1857.61	1857.61	1	0.00	2481.93
PbO	499.19	499.19	1	0.00	564.34
PbO ₂	1994.02	1994.02	1	0.00	2276.75
Co ₃ O ₄	811.12	799.39	-	-11.73	1192.24
PtO	813.94	785.63	3	-28.31	769.78
Au ₂ O	314.86	320.55	1	5.69	300.25
Tl ₂ O ₃	1341.37	1341.37	1 v 3	0.00	1343.39
Mn ₂ O ₃	1017.99	995.07	3	-22.92	1375.59
Co ₂ O ₃	919.39	904.43	5	-14.96	1449.75
SiO ₂	-	1686.38	1	-	2485.84

8. táblázat: Számítási eredmények és kísérleti adatok [40] a fémionokra (kcal/mol).

Sajnos 9 ion esetén nem áll rendelkezésünkre experimentális képződéshő. Érdekes módon a PM6 módszer a Tl(III) ion szingulett és triplett állapotára ugyanakkora képződéshő értéket adott.

A Co₃O₄ fém-oxid nanorészecske deszkriptorát az alábbi összefüggés alapján számították:

$$\Delta H_{Me^+} = (\Delta H_{Co^{2+}} + 2 \Delta H_{Co^{3+}}) / 3. \quad (55)$$

A szerzők a következő fémionok esetén a spinmultiplicitásokat hibásan szingulettnek állították be: V(III), Nb(III), Ta(III), W(IV), Ru(IV), Os(IV). Összesen 17 fémionnál a számításukat nem tudtuk reprodukálni: a közölt képződéshők irreálisak. A következő 17 fémionról van szó: Cu(II), Fe(III), Co(II), Ni(II), Cr(III), Mo(III), Mn(IV), Tc(IV), Rh(III), Ir(IV), Pd(II), Pt(IV), Au(III), Pt(II), Au(I), Mn(III), Co(III). A számított PM6 deszkriptorok értékei között az alábbi három fémion esetén találtuk a legnagyobb különbségeket: Ru(IV), Tc(IV), Nb(III). A különbségek rendre: 272.65, 233.19, illetve 114.60 kcal/mol. Tehát megállapítható, hogy a nano-QSAR-modellben összesen 23 fém-oxid nanorészecske deszkriptora valamilyen oknál fogva hibás. Ezek alapján a Wu és mtsai. által felállított modell jósága megkérdőjelezhető.

Itt is egyértelműen látszik, hogy az általunk helyesen számolt PM6 képződéshők jelentősen eltérnek az experimentális értékektől. Vagyis szükséges a modell deszkriptorainak reálisabbakkal való helyettesítése.

4.2 Fématomok ionizációs energiája, mint deszkriptor

Ismeretes, hogy kvantumkémiai módszerekkel sokkal pontosabban tudunk energiakülönbségeket számítani, mint diszkrét energiaértékeket. Azt is tudjuk, hogy a fém-oxid nanorészecskék toxikusságáért nem a fématomok felelősek, hanem a vízben belőlük kioldódó fémionok. Ezért a következő ötletünk támadt: a fémionok standard képződéshői helyett használjuk a fématomoknak a megfelelő fémionokhoz vezető ionizációs energiáit deszkriptorként.

Az ionizációs energiák kvantumkémiai számításához a Gaussian 09 program esetén a G2 módszert és a CCSD(T) módszert alkalmaztuk növekvő bázisokkal [43]. A kísérleti ionizációs energiákat és a számítási eredményeinket a 9. táblázat mutatja be. Az CCSD(T)/aug-cc-pVnZ és a G2 módszerek sajnos csak a kriptonig alkalmazhatók, így összesen csak 10 fémionra tudunk ionizációs energiát számolni ezekkel a módszerekkel. Ellenben a CCSD(T)/def2-QZVPP módszer egészen a radonig használható, azaz mind a 17 fémionra tudtuk alkalmazni. Látható, hogy a G2 módszerrel számolt energiák több esetben közelebb vannak az experimentális értékhez, mint a CCSD(T)/aug-cc-pV5Z-vel számoltak. Azonban a Cu(II) G2 ionizációs energiájánál megfigyelhető egy jelentős (44.44 kcal/mol) energiakülönbség a kísérleti értékhez képest. Az ionizációs energia dimenziója minden esetben kcal/mol.

Me ⁿ⁺	IE Exp.	IE G2	IE CCSD(T)/ def2-QZVPP	IE CCSD(T)/aug-cc-pVnZ			
				n=2	n=3	n=4	n=5
Al(III)	1228.25	1215.55	1213.60	1210.07	1212.56	1213.50	1213.75
Bi(III)	1142.62	-	1109.12	-	-	-	-
Co(II)	575.71	562.60	556.73	534.46	547.20	559.15	560.06
Cr(III)	1250.16	1229.58	1227.83	1220.54	1227.55	1232.00	1232.50
Cu(II)	646.13	601.69	603.47	584.15	599.09	600.47	641.32
Fe(III)	1262.38	1223.50	1212.10	1169.90	1215.70	1216.02	1222.64
In(III)	1214.98	-	1197.73	-	-	-	-
La(III)	825.87	-	777.26	-	-	-	-
Ni(II)	595.16	583.49	582.14	564.74	587.48	584.53	590.92
Sb(III)	1165.44	-	1158.79	-	-	-	-
Si(IV)	2378.29	2358.56	2357.93	2345.59	2355.97	2357.81	2358.32
Sn(IV)	2149.56	-	2126.30	-	-	-	-
Ti(IV)	2102.26	2059.19	2058.77	2093.65	2054.73	2058.15	2058.53
V(III)	1168.60	1142.40	1142.18	1139.47	1141.68	1139.57	1128.43
Y(III)	898.47	-	863.70	-	-	-	-
Zn(II)	630.90	620.24	615.46	606.10	615.13	616.83	617.46
Zr(IV)	1777.14	-	1728.59	-	-	-	-

9. táblázat: A fém-oxid nanorészecskéket felépítő atomok ionjaihoz tartozó kísérleti illetve számított ionizációs energiák (kcal/mol).

Végezzünk egyparaméteres lineáris regressziókat a számított és a kísérletileg meghatározott ionizációs energiák között ($y_i = ax_i$)! A regressziók statisztikai mutatói a 10. táblázatban láthatóak.

Módszer	m	\hat{a}	r_{xy}	s_r	n
CCSD(T)/aug-cc-pVDZ	10	0.9771 ± 0.01759	0.9993	32.64	Fe
CCSD(T)/aug-cc-pVTZ	10	0.9800 ± 0.009345	0.9998	17.33	Si,Cu
CCSD(T)/aug-cc-pVQZ	10	0.9813 ± 0.008857	0.9998	16.43	Si,Cu
CCSD(T)/aug-cc-pV5Z	10	0.9830 ± 0.006716	0.9998	12.46	Si
G2	10	0.9825 ± 0.008079	0.9998	14.99	Si,Cu
CCSD(T)/def2-QZVPP	17	0.9803 ± 0.006552	0.9997	17.21	La

10. táblázat: Fémionokra számított ionizációs energiák és a kísérleti értékek közötti egyparaméteres lineáris regressziók statisztikai mutatói.

A 10. táblázat 4. sora mutatja, hogy a legjobb illesztést a CCSD(T)/aug-cc-pV5Z módszerrel számolt ionizációs energiák adják. Megállapítható az is, hogy a CCSD(T)/def2-QZVPP módszer számítási igénye, illetve alkalmazási lehetősége (radonig alkalmazható) jóval kedvezőbb bármelyik másik, a 10. táblázatban felsorolt módszernél.

Hogy az összes (17) fémionra a lehető legpontosabban meghatározzuk az ionizációs energiát, komolyabb kvantumkémiai apparátus és hardver szükséges. Ezeket a nagy pontosságú kvantumkémiai számításokat Dr. Csontos József és mtsai. (MTA-BME Kvantumkémiai Kutatócsoport) végzik együttműködésben velünk. A munka még folyamatban van, így csak egy köztes állapotról tudunk itt beszámolni. A számítások a def2-TZVPP és def2-QZVPP [44,45], illetve a bizmut esetén a SARC (*segmented all-electron relativistically contracted*) bázisokra [46] épülnek. A def2-QZVPP bázissal számolt Hartree-Fock energiáról feltételezhető, hogy az már jól megegyezik a végtelen bázisra extrapolált értékkel. A def2-TZVPP és a def2-QZVPP bázisokkal számolt MP2, CCSD, CCSD(T), CCSDT, CCSDT(Q) additív energiakorrekciónak pedig végtelen bázisra való extrapoláció (*complete basis set, CBS*) végezhető egy kétparaméteres inverz köbös formula (*inverse cubic formula*) [47] segítségével. A harmadik és a negyedik periódus elemeinél a relativisztikus hatás másodrendű Douglas-Kroll-Hess (DKH) Hamilton-operátor [48,49] segítségével figyelembe vehető. A többi elemre a Stuttgart-Dresden effektív törzspotenciálok módszerével (*effective core potentials, ECP*) [50,51] számítható a relativisztikus hozzájárulás. A számítások az MRCC és az ORCA kvantumkémiai programcsomagokkal végezhetők el. Csontos és mtsai. által számított aktuális ionizációs energiaértékek és az abból felépített modell az RQC (*Relativistic Quantum Chemistry*) jelölést kapja. A 11. táblázatban tüntettük fel az experimentális, az RQC, illetve a többi fontosabb – általunk számított – ionizációs energiákat. A táblázat 5. oszlopában az általunk helyesen számított PM6 ionizációs energiák ($IE = \Delta H_{Me^+} - \Delta H_{Me}$) is megtalálhatók. (Az ionizációs energia dimenziója minden esetben kcal/mol, az EC₅₀ koncentrációé pedig mol/L.) Mint ahogy az várható volt: a kísérleti értékekhez a legközelebb az RQC energiák vannak. Ez alól kivételt képez a Ni(II) ion, mivel a CCSD(T)/aug-cc-pV5Z módszerrel számolt ionizációs energia 2.92 kcal/mol-al közelebb van az experimentális értékhez.

Me ⁿ⁺	pMEC ₅₀	IE Exp.	IE RQC	IE PM6	IE CCSD(T)/	
					def2-QZVPP	aug-cc-pV5Z
Al(III)	2.49	1228.25	1225.35	1113.76	1213.60	1213.75
Bi(III)	2.82	1142.62	1113.72	1087.30	1109.12	-
Co(II)	3.51	575.71	569.72	505.79	556.73	560.06
Cr(III)	2.51	1250.16	1255.73	1155.18	1227.83	1232.50
Cu(II)	3.20	646.13	649.26	625.55	603.47	641.32
Fe(III)	2.29	1262.38	1255.20	1177.82	1212.10	1222.64
In(III)	2.81	1214.98	1207.86	1213.13	1197.73	-
La(III)	2.87	825.87	821.18	941.83	777.26	-
Ni(II)	3.45	595.16	588.00	469.28	582.14	590.92
Sb(III)	2.64	1165.44	1163.06	1169.86	1158.79	-
Si(IV)	2.20	2378.29	2372.66	1578.00	2357.93	2358.32
Sn(IV)	2.01	2149.56	2138.29	1645.12	2126.30	-
Ti(IV)	1.74	2102.26	2098.69	1526.79	2058.77	2058.53
V(III)	3.14	1168.60	1165.99	930.38	1142.18	1128.43
Y(III)	2.87	898.47	893.86	742.04	863.70	-
Zn(II)	3.45	630.90	623.31	631.27	615.46	617.46
Zr(IV)	2.15	1777.14	1768.81	1212.72	1728.59	-

11. táblázat: A fémionokhoz tartozó kísérleti és számított ionizációs energiák (kcal/mol).

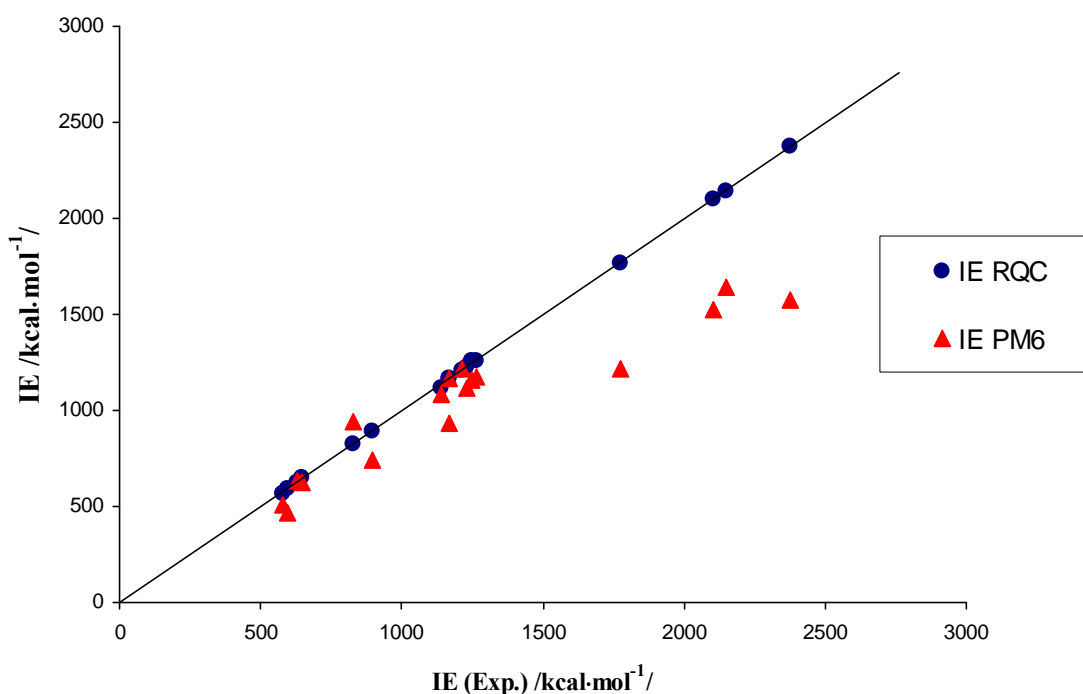
Sajnos, itt is látható, hogy a PM6 módszerrel számított energiaértékek megbízhatatlanok: van ahol szinte teljesen megegyezik a kísérleti értékkel (Zn(II)), van ahol meg a különbség több mint 500 kcal/mol (Si(IV)).

Végezzünk most is egyparaméteres lineáris regressziókat a számított (PM6, RQC) és a kísérleti ionizációs energiák között ($y_i = ax_i$)! A regressziók statisztikai mutatóit a 12. táblázatban tüntettük fel.

Módszer	m	\hat{a}	r_{xy}	S_r	n
IE PM6	17	0.804 ± 0.067	0.9373	177.07	Si
IE RQC	17	0.996 ± 0.003	0.9999	7.37	Bi

12. táblázat: Fémionokra számított ionizációs energiák és a kísérleti értékek közötti egyparaméteres lineáris regressziók statisztikai mutatói.

Az RQC értékek nagyon jó egyezést mutatnak a kísérleti értékekkel! Egy kiugró pont azonban itt is van (Bi). Ez várható volt, mivel a Bi(III) ion esetén kaptuk a legnagyobb különbséget (28.90 kcal/mol) az RQC és a kísérleti értékek között. A 2. ábrán grafikusán is bemutatjuk a 12. táblázat adatait.



2.ábra: A PM6 és az RQC ionizációs energiák a kísérleti értékek függvényében.

A 2. ábrán látható, hogy amíg az RQC értékek jól illeszkednek az $y = x$ egyenesre, addig a PM6 értékek esetén minél nagyobb a meghatározandó ionizációs energia értéke, annál pontatlanabb a számítás. Vagyis most már teljes bizonyossággal kijelenthetjük, hogy a nano-QSAR-modell felállításához nem alkalmazható a Leszczynski és mtsai., illetve Wu és mtsai. által használt PM6 szemiempirikus kvantumkémiai módszer a deskriptor értékeinek a meghatározására. A PM6 módszer helyett mi számos másikat javasolunk, melyek segítségével a következőkben felállítjuk a lehető legmegbízhatóbb modellt.

Eddig a lineáris regressziók végrehajtására a közönséges legkisebb négyzetek módszerét alkalmaztuk, azaz feltételeztük, hogy a független változó értéke pontos, a függő változó értékei meg ugyanakkora hibával rendelkeznek. Válasszuk most független változónak a kis hibával rendelkező ionizációs energiát! Függő változónak pedig a $pMEC_{50}$ értéket megfelelő súlyozással. Feltételezhető, hogy minél toxikusabb az adott fém-oxid nanorészecske, annál kisebb a mérési hiba; vagyis az (56) egyenlet jól alkalmazható a súlyfaktorok számítására.

$$w_i = \frac{(pMEC_{50})_i^2}{\sum_{k=1}^m (pMEC_{50})_k^2}, \quad (56)$$

ahol m az EC_{50} adatok száma.

Korábban már megmutattuk, hogy a SiO_4 EC_{50} értéke egyértelműen kiugró pont, vagyis összesen 16 iont vizsgálunk a továbbiakban. Végezzünk kétparaméteres lineáris regressziókat a $pMEC_{50}$ és az experimentális IE értékek, illetve a számított IE értékek között először a közönséges legkisebb négyzetek módszerével (*ordinary linear least-squares method*, OLLS), majd a súlyozott legkisebb négyzetek módszerével (*weighted linear least-squares method*, WLLS) az (56) súlyfaktorokkal! A 13. táblázatban tüntettük fel az illesztések statisztikai adatait. A 16 ionra történő regressziók esetén a V(III) ion minden esetben kiugrópontnak bizonyult, így eltávolítottuk.

A 15 ionra történő lineáris regresszióknál először a közönséges lineáris legkisebb négyzetek módszerét alkalmaztuk. Ezeknél csak a CCSD(T)/def2-QZVPP módszer esetén találtunk kiugrópontot (Fe(III)). Ezt azonban nem távolítottuk el, mivel később mikor már a súlyozott legkisebb négyzetek módszerével végeztünk regressziókat, akkor a Fe(III) nem bizonyult kiugrópontnak. A 13. táblázat utolsó két sorában található az általunk felállított legjobb modellek:

$$-\lg(EC_{50}) = 2.731 - 0.549 \Delta H_{Me^+}, \quad (57)$$

$$-\lg(EC_{50}) = 2.730 - 0.550 \Delta H_{Me^+}. \quad (58)$$

Ez a kettő modell jóval pontosabb, mint az eredeti 1. Modell; a maradék szórás értéke körülbelül 0.044, míg az eredeti 1. Modell esetén 0.2207 volt.

Módszer	m	\hat{a}	\hat{b}	r_{xy}	s_r	n
CCSD(T)/ def2-QZVPP OLLS	16	-0.490 ± 0.055	2.747 ± 0.053	-0.9228	0.2117	V
IE Exp. OLLS	16	-0.492 ± 0.053	2.747 ± 0.051	-0.9278	0.2050	V
IE RQC OLLS	16	-0.493 ± 0.052	2.747 ± 0.051	-0.9290	0.2033	V
CCSD(T)/ def2-QZVPP OLLS	15	-0.507 ± 0.050	2.721 ± 0.048	-0.9421	0.1875	Fe
IE Exp. OLLS	15	-0.510 ± 0.048	2.721 ± 0.046	-0.9469	0.1798	-
IE RQC OLLS	15	-0.511 ± 0.047	2.721 ± 0.046	-0.9485	0.1771	-
CCSD(T)/ def2-QZVPP WLLS	15	-0.545 ± 0.057	2.733 ± 0.050	-0.9470	0.0465	-
IE Exp. WLLS	15	-0.549 ± 0.053	2.731 ± 0.048	-0.9521	0.0442	-
IE RQC WLLS	15	-0.550 ± 0.052	2.730 ± 0.047	-0.9538	0.0435	-

13. táblázat: Az ionizációs energiák és a toxicitási értékek közötti kétparaméteres lineáris regressziók statisztikai mutatói.

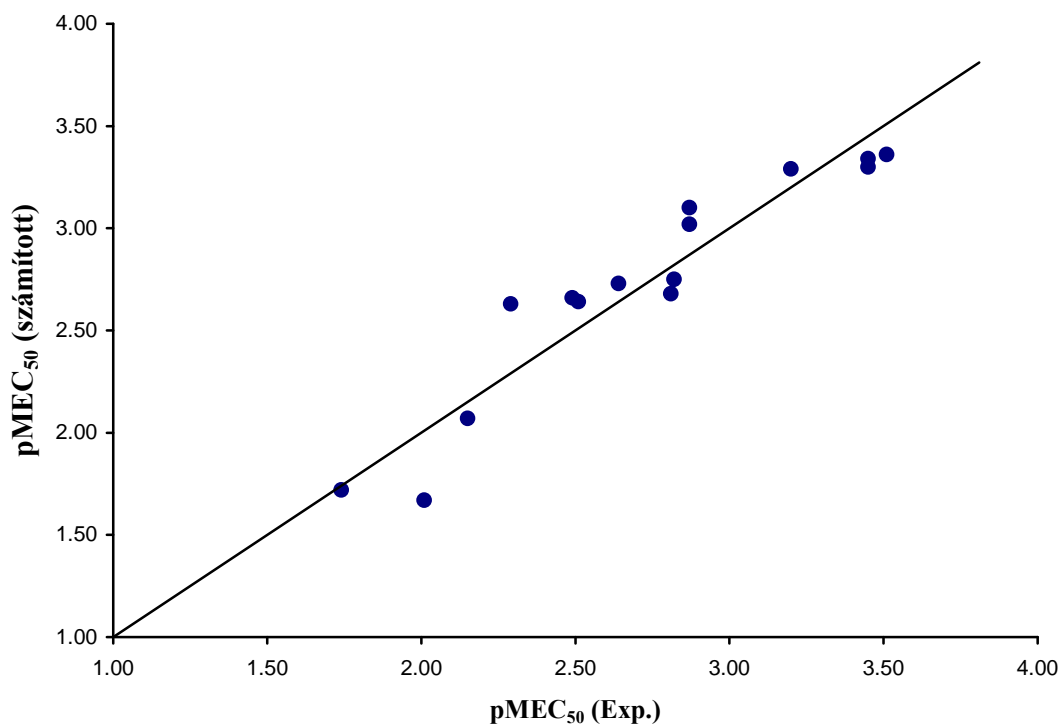
4.3 Fém-oxid nanorészecskék toxicitásának becslése

A fém-oxid nanorészecskék toxicitásának becslésére a 13. táblázat 8. sorában szereplő modellt alkalmazzuk. Ennek az az oka, hogy csak azok az RQC ionizációs energiák állnak rendelkezésünkre, amelyek a modell felállításához voltak szükségesek. Ellenben a fématomok kísérleti ionizációs energiái általában hozzáférhetőek az irodalomban [40]. Nevezzük el az alkalmazni kívánt modellünket – (57) egyenlet – a már megszokott módon (3. Modell). A 14. táblázatban bemutatjuk a kísérleti és a 3. Modell által számított toxicitási adatokat.

Fém-oxid	IE Exp.	pMEC ₅₀ Exp.	pMEC ₅₀ 3. Modell	Δ
Al ₂ O ₃	1228.25	2.49	2.66	0.17
Bi ₂ O ₃	1142.62	2.82	2.75	-0.07
CoO	575.71	3.51	3.36	-0.15
Cr ₂ O ₃	1250.16	2.51	2.64	0.13
CuO	646.13	3.20	3.29	0.09
Fe ₂ O ₃	1262.38	2.29	2.63	0.34
In ₂ O ₃	1214.98	2.81	2.68	-0.13
La ₂ O ₃	825.87	2.87	3.10	0.23
NiO	595.16	3.45	3.34	-0.11
Sb ₂ O ₃	1165.44	2.64	2.73	0.09
SnO ₂	2149.56	2.01	1.67	-0.34
TiO ₂	2102.26	1.74	1.72	-0.02
Y ₂ O ₃	898.47	2.87	3.02	0.15
ZnO	630.9	3.45	3.30	-0.15
ZrO ₂	1777.14	2.15	2.07	-0.08

14. táblázat: A 3. Modell alapjául szolgáló fém-oxid nanorészecskék kísérleti és számított toxicitási adatai (mol/L).

Látható, hogy legnagyobb különbség a toxicitási értékek között a Fe₂O₃ és a SnO₂ fém-oxidok esetén van (± 0.34 kcal/mol). A 3. ábrán grafikusán is bemutatjuk a kapcsolatot a kísérleti és a számított értékek között.



3.ábra: A kísérleti és a számított toxicitási adatok közötti kapcsolat a 3. Modell tanítókészlete esetén.

Ahhoz, hogy további fém-oxid nanorészecskék kólibaktériumokra való toxicitását megbecsüljük – melyekre nem áll rendelkezésre kísérleti adat – két fontos szempontot kell figyelembe venni. Először is, hogy csak azon fém-oxid nanorészecskék toxicitására tudunk predikciót végrehajtani, melyek estén a nanorészecskéket alkotó fématomok ionjaihoz tartozó kísérleti ionizációs energiák az irodalomban hozzáférhetőek. Másodsor figyelembe kell venni a 3. Modell tanítókészletének az értelmezési (kémiai) tartományát; csak azon nanorészecskékre tudunk becslést mondani, melyekhez tartozó ionizációs energia 575.71 és 2149.56 kcal/mol, illetve a pMEC₅₀ érték pedig 2.01 és 3.51 között van. Vagyis csak interpolációt szabad végrehajtanunk, extrapolációt nem. A 15. táblázatban foglaltuk össze a 3. Modell alapján az összes lehetséges fém-oxid nanorészecske becsült toxicitását kólibaktériumokra nézve. A fent említett két feltétel mellett összesen 16 fém-oxid nanorészecske toxicitását tudtuk megbecsülni a 3. Modell alapján.

A Co₃O₄ nanorészecske deskriptorát a már bemutatott (55) formula alapján számítottuk:

$$IE_{Me^+} = (IE_{Co^{2+}} + 2 IE_{Co^{3+}}) / 3. \quad (59)$$

Fém-oxid	Me ⁿ⁺	IE Exp.	pMEC ₅₀ 3. Modell
CdO	Cd(II)	597.32	3.34
PtO	Pt(II)	634.67	3.30
BeO	Be(II)	634.95	3.30
PdO	Pd(II)	640.32	3.29
HgO	Hg(II)	673.24	3.26
Sc ₂ O ₃	Sc(III)	1017.38	2.89
Nb ₂ O ₃	Nb(III)	1056.15	2.85
Mo ₂ O ₃	Mo(III)	1161.85	2.73
V ₂ O ₃	V(III)	1168.60	2.73
Tl ₂ O ₃	Tl(III)	1299.84	2.59
Rh ₂ O ₃	Rh(III)	1305.20	2.58
Mn ₂ O ₃	Mn(III)	1308.50	2.58
Ga ₂ O ₃	Ga(III)	1319.99	2.56
Co ₂ O ₃	Co(III)	1348.24	2.53
Co ₃ O ₄	Co(III) és Co(II)	1599.58	2.26
HfO ₂	Hf(IV)	1809.22	2.04

15. táblázat: A kémiai tartományba eső összes fém-oxid nanorészecske toxicitásának becsült értéke a 3. Modell alapján.

Sajnos a SiO₂ nanorészecskéhez tartozó ionizációs energia értéke a modell tanítókészletén kívül esik, így annak toxicitására nem tudunk megfelelő becslést tenni. Ellenben a V₂O₃ nanorészecskére tudunk, a becsült érték 2.73, míg a kísérleti érték: 3.14. Reméljük, hogy további experimentális toxicitási adatok is rendelkezésünkre fognak majd állni, melyekkel ellenőrizni tudjuk modellünk teljesítőkéességét.

Számítási eredményeinkkel sikerült igazolnunk, hogy a Leszczynski és mtsai., illetve a Wu és mtsai. által felállított nano-QSAR-modellekben szereplő PM6 standard képződéshők hibásak. Ezt a deszkriptort kísérleti, illetve elméleti ionizációs energiákra cserélve felállítottunk egy pontosabb QSAR-modellt. Az új modell segítségével további 16 fém-oxid nanorészecske toxicitását sikerült megbecsülnünk.

5. Összefoglalás

Leszczynski és mtsai. 2011-ben felállítottak egy nano-QSAR-modellt fém-oxid nanorészecskék kólibaktériumokra vonatkozó citotoxicitásának leírására. A szerzők a modell felállításához a fém-oxid nanorészecskéket alkotó fématomok ionjainak PM6 standard képződéshőjét választották deskriptorként. A toxikus hatásért a nanorészecskékből vízben kioldódó fémionokat tartották felelősnek. Wu és mtsai. a modellt kiegészítve – összesen 51 fém-oxid nanorészecskére – egy pontosabb összefüggést állítottak fel. Számukra kettő deskriptor bizonyult megfelelőnek: a fémionok PM6 standard képződéshője és az ionok polarizációs ereje. Munkánk során e két nano-QSAR-modell megalapozottságát vizsgáltuk és sikerült egy fizikailag reálisabbat felállítanunk.

Kvantumkémiai és statisztikai módszerek segítségével megállapítottuk, hogy a Leszczynski és mtsai. által felállított nano-QSAR-modell 17 fémionja közül 9 ionra a szerzők által közölt PM6 standard képződéshők hibásak. A modell Wu és mtsai. által továbbfejlesztett változatát is megvizsgáltuk. Kiderült, hogy a PM6 standard képződéshők számítása során 6 fémionra a spinmultiplicitást hibásan állították be. További 17 fémion esetén nem sikerült reprodukálni az általuk közölt adatokat.

Vizsgálataink azt mutatják, hogy lehetséges a PM6 standard képződéshöket fizikailag reálisabb kísérleti vagy nagy pontosságú kvantumkémiai ionizációs energiákra cserélni. Az ionizációs energia kvantumkémiai számítására elsősor a G2 és a CCSD(T) módszereket alkalmaztuk. Az experimentális értékekhez legközelebb a CCSD(T)/aug-cc-pV5Z módszerrel számoltak voltak. Dr. Csontos József és mtsai. – együttműködésben velünk – nagy pontosságú kvantumkémiai számításokat végeztek és végeznek: meghatározták a modell felépítéséhez szükséges összes fémionra vonatkozó ionizációs energiát. Ezek a számítási eredmények mutatták a legjobb egyezést a kísérleti adatokkal.

Az experimentális ionizációs energiákat felhasználva egy megbízhatóbb, pontosabb nano-QSAR-modellt sikerült felállítanunk, mellyel megbecsülhető a modell kémiai tartományába eső összes fém-oxid nanorészecske citotoxicitása kólibaktériumra nézve: 16 fém-oxid nanorészecskére becslést is adtunk.

6. Felhasznált irodalom

- [1] Shevchenko, V. Y., Madison, A. E., Shudegov, V. E.: "*The Structural Diversity of the Nanoworld.*" *Glass Physics and Chemistry*, 29, 577 (2003).
- [2] Shevchenko, V. Y.: "*Structural chemistry of the nanoworld - A new page of inorganic chemistry.*" *Glass Physics and Chemistry*, 37, 467 (2011).
- [3] Gajewicz, A., Rasulev, B., Dinadayalane, T. C., Urbaszek, P., Puzyn, T., Leszczynska, D., Leszczynski, J.: "*Advancing risk assessment of engineered nanomaterials: application of computational approaches.*" *Advanced drug delivery reviews*, 64, 1663 (2012).
- [4] Artilles, M. S., Rout, C. S., Fisher, T. S.: "*Graphene-based hybrid materials and devices for biosensing.*" *Advanced drug delivery reviews*, 63, 1352 (2011).
- [5] Puzyn, T., Gajewicz, A., Leszczynska, D., Leszczynski, J. (2010). *Nanomaterials - the next great challenge for QSAR modelers. Recent advances in QSAR studies*, Springer: 383.
- [6] Venkataraman, S., Hedrick, J. L., Ong, Z. Y., Yang, C., Ee, P. L. R., Hammond, P. T., Yang, Y. Y.: "*The effects of polymeric nanostructure shape on drug delivery.*" *Advanced drug delivery reviews*, 63, 1228 (2011).
- [7] Hou, W.-C., Westerhoff, P., Posner, J. D.: "*Biological accumulation of engineered nanomaterials: a review of current knowledge.*" *Environmental Science: Processes & Impacts*, 15, 103 (2013).
- [8] Luyts, K., Napierska, D., Nemery, B., Hoet, P. H. M.: "*How physico-chemical characteristics of nanoparticles cause their toxicity: complex and unresolved interrelations.*" *Environmental Science: Processes & Impacts*, 15, 23 (2013).
- [9] Clark, K. A., White, R. H., Silbergeld, E. K.: "*Predictive models for nanotoxicology: current challenges and future opportunities.*" *Regulatory Toxicology and Pharmacology*, 59, 361 (2011).
- [10] Gilbertson, L. M., Zimmerman, J. B., Plata, D. L., Hutchison, J. E., Anastas, P. T.: "*Designing nanomaterials to maximize performance and minimize undesirable implications guided by the Principles of Green Chemistry.*" *Chemical Society Reviews*, 44, 5758 (2015).
- [11] Basiuk, V. A., Basiuk, E. V., Eds.: *Green Processes for Nanotechnology: From Inorganic to Bioinspired Nanomaterials*, Springer (2015).
- [12] Puzyn, T., Leszczynska, D., Leszczynski, J.: "*Toward the development of "Nano-QSAR": advances and challenges.*" *Small*, 5, 2494 (2009).
- [13] Dreher, K. L.: "*Health and Environmental Impact of Nanotechnology: Toxicological Assessment of Manufactured Nanoparticles.*" *Toxicological Sciences*, 77, 3 (2004).
- [14] Kar, S., Roy, K.: "*Predictive toxicology using QSAR: a perspective.*" *Journal of the Indian Chemical Society*, 87, 1455 (2010).
- [15] Cronin, M. T. D., Madden, J. C., Eds.: *In Silico Toxicology: Principles and Applications*. Cambridge, The Royal Society of Chemistry (2010).
- [16] Tasi, G.: *Számítógépes toxikológia*. Szeged, JATEPress (2014).
- [17] Szabó, G., Ed. *Sejtbiológia*. Budapest, Medicina Könyvkiadó (2009).
- [18] Rappoport, J., Preece, J., Chipman, K., Argatov, I., Davies, A., Dyson, L., Dyson, R., Lang, G., Johnson, A., Mayaud, L.: "*How do manufactured nanoparticles enter cells?*" Report on a problem studied at the UK Mathematics-in-Medicine Study Group. Reading. (2011).

- [19] Mahmoudi, M., Azadmanesh, K., Shokrgozar, M. A., Journeay, W. S., Laurent, S.: "Effect of Nanoparticles on the Cell Life Cycle." *Chemical Reviews*, 111, 3407 (2011).
- [20] Puzyn, T., Rasulev, B., Gajewicz, A., Hu, X., Dasari, T. P., Michalkova, A., Hwang, H.-M., Toropov, A., Leszczynska, D., Leszczynski, J.: "Using nano-QSAR to predict the cytotoxicity of metal oxide nanoparticles." *Nat Nano*, 6, 175 (2011).
- [21] Sizochenko, N., Gajewicz, A., Leszczynski, J., Puzyn, T.: "Causation or only correlation? Application of causal inference graphs for evaluating causality in nano-QSAR models." *Nanoscale*, 8, 7203 (2016).
- [22] Toropov, A. A., Toropova, A. P., Benfenati, E., Gini, G., Puzyn, T., Leszczynska, D., Leszczynski, J.: "Novel application of the CORAL software to model cytotoxicity of metal oxide nanoparticles to bacteria *Escherichia coli*." *Chemosphere*, 89, 1098 (2012).
- [23] Stewart, J. J. P.: "Optimization of parameters for semiempirical methods V: Modification of NDDO approximations and application to 70 elements." *Journal of Molecular Modeling*, 13, 1173 (2007).
- [24] Carbó-Dorca, R., Besalu, E.: "Construction of coherent nano quantitative structure-properties relationships (nano-QSPR) models and catastrophe theory." *SAR and QSAR in Environmental Research*, 22, 661 (2011).
- [25] Mu, Y., Wu, F., Zhao, Q., Ji, R., Qie, Y., Zhou, Y., Hu, Y., Pang, C., Hristozov, D., Giesy, J. P.: "Predicting toxic potencies of metal oxide nanoparticles by means of nano-QSARs." *Nanotoxicology*, 10, 1207 (2016).
- [26] Martin, Y. C.: "Hansch analysis 50 years on." *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 2, 435 (2012).
- [27] Selassie, C., Verma, R. P., Abraham, D. J. (2003). *History of Quantitative Structure–Activity Relationships*. Burger's Medicinal Chemistry and Drug Discovery, John Wiley & Sons, Inc.
- [28] Hansch, C., Leo, A., Hoekman, D. H.: *Exploring QSAR: fundamentals and applications in chemistry and biology*, American Chemical Society Washington, DC (1995).
- [29] Hansch, C., Leo, A., Hoekman, D. H.: *Exploring QSAR: Volume 2: Hydrophobic, Electronic, and Steric Constants*, American Chemical Society Washington, DC (1995).
- [30] Veszprémi, T., Fehér, M.: *A kvantumkémia alapjai és alkalmazása*. Budapest, Műszaki Könyvkiadó (2002).
- [31] Tasi, G., Mayer, I.: "An extension of the virial theorem for general wave functions." *Chemical Physics Letters*, 449, 221 (2007).
- [32] Boys, S. F.: "Integrals Over Primitive Gaussians." *Proc. R. Soc. London. Ser. A*, 200, 532 (1950).
- [33] Ostlund, N. S., Szabo, A.: *Modern quantum chemistry: introduction to advanced electronic structure theory*, Dover Publications Inc New edition edn (1996).
- [34] Raghavachari, K., Trucks, G. W., Pople, J. A., Head-Gordon, M.: "A fifth-order perturbation comparison of electron correlation theories." *Chemical Physics Letters*, 157, 479 (1989).
- [35] Thiel, W.: "Semiempirical quantum-chemical methods." *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 4, 145 (2014).
- [36] Stewart, J. J. P.: "MOPAC: a semiempirical molecular orbital program." *Journal of computer-aided molecular design*, 4, 1 (1990).

- [37] Stewart, J. J. P.: "*Optimization of parameters for semiempirical methods VI: more modifications to the NDDO approximations and re-optimization of parameters.*" *Journal of Molecular Modeling*, 19, 1 (2013).
- [38] Bohr, N.: "*I. On the constitution of atoms and molecules.*" *Philosophical Magazine Series* 6, 26, 1 (1913).
- [39] Hedvig, P.: *Experimental quantum chemistry*. New York, Academic Press (1975).
- [40] Haynes, W.: *CRC Handbook of Chemistry and Physics, 96th Edition (Internet Version 2016)*, CRC Press/Taylor and Francis (2016).
- [41] Tasi, G.: *Számítógépes kémia*. Szeged, JATEPress (2010).
- [42] Tasi, G.: *Matematikai kémia*. Szeged, JATEPress (2014).
- [43] Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Mennucci, B., Petersson, G. A., Nakatsuji, H., Caricato, M., Li, X., Hratchian, H. P., Izmaylov, A. F., Bloino, J., Zheng, G., Sonnenberg, J. L., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., A. Montgomery Jr., J. A., Peralta, J. E., Ogliaro, F., Bearpark, M., Heyd, J. J., Brothers, E., Kudin, K. N., Staroverov, V. N., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A., Burant, J. C., Iyengar, S. S., Tomasi, J., Cossi, M., Rega, N., Millam, J. M., Klene, M., Knox, J. E., Cross, J. B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R. E., Yazyev, O., Austin, A. J., Cammi, R., Pomelli, C., Ochterski, J. W., Martin, R. L., Morokuma, K., Zakrzewski, V. G., Voth, G. A., Salvador, P., Dannenberg, J. J., Dapprich, S., Daniels, A. D., Farkas, Ö., Foresman, J. B., Ortiz, J. V., Cioslowski, J., Fox, D. J. (2009). *Gaussian 09, Revision C.01*. Wallingford CT, Gaussian, Inc.
- [44] Schafer, A., Horn, H., Ahlrichs, R.: "*Fully optimized contracted Gaussian basis sets for atoms Li to Kr.*" *The Journal of Chemical Physics*, 97, 2571 (1992).
- [45] Weigend, F., Ahlrichs, R.: "*Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy.*" *Physical Chemistry Chemical Physics*, 7, 3297 (2005).
- [46] Pantazis, D. A., Neese, F.: "*All-electron scalar relativistic basis sets for the 6p elements.*" *Theoretical Chemistry Accounts*, 131, 1292 (2012).
- [47] Helgaker, T., Klopper, W., Koch, H., Noga, J.: "*Basis-set convergence of correlated calculations on water.*" *The Journal of chemical physics*, 106, 9639 (1997).
- [48] Douglas, M., Kroll, N. M.: "*Quantum electrodynamical corrections to the fine structure of helium.*" *Annals of Physics*, 82, 89 (1974).
- [49] Hess, B. A.: "*Relativistic electronic-structure calculations employing a two-component no-pair formalism with external-field projection operators.*" *Physical Review A*, 33, 3742 (1986).
- [50] Andrae, D., Haeussermann, U., Dolg, M., Stoll, H., Preuss, H.: "*Energy-adjusted ab initio pseudopotentials for the second and third row transition elements.*" *Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta)*, 77, 123 (1990).
- [51] Metz, B., Stoll, H., Dolg, M.: "*Small-core multiconfiguration-Dirac-Hartree-Fock-adjusted pseudopotentials for post-d main group elements: application to PbH and PbO.*" *The Journal of Chemical Physics*, 113, 2563 (2000).

7. Köszönetnyilvánítás

Szeretném megköszönni témavezetőimnek, Dr. Kónya Zoltán professzor úrnak és Dr. Tasi Gyula egyetemi docens úrnak a munkám során nyújtott magas színvonalú szakmai vezetést. Támogatásuk és aktív segítségük nélkül ez a munka nem készülhetett volna el.

8. Nyilatkozat

Alulírott Tasi Domonkos Attila, vegyész MSc szakos hallgató, kijelentem, hogy a diplomadolgozatban foglaltak saját munkám eredményei, és csak a hivatkozott forrásokat (szakirodalom, eszközök, stb.) használtam fel.

Tudomásul veszem azt, hogy diplomamunkámat a Szegedi Tudományegyetem könyvtárában, a kölcsönözhető könyvek között helyezik el.

2017. 05. 08.

.....

Szegedi Tudományegyetem
Természettudományi és Informatikai Kar