

## Zárójelentés

Fogarasi Géza:

Szisztematikus kvantumkémiail számítások biomolekulák szerkezetének és energetikájának vizsgálatára

OTKA T 043628, 2003-2007.

A pályázat eredetileg a 2003-2006. időszakra szólt. Erősen csúszott azonban a forrás megnyitása (2003. október), majd maga a kutatás is (1. alább). Ezért 1 éves hosszabbítást kértem és kaptam, s így a tényleges tevékenység inkább 2004-2007-ben folyt.

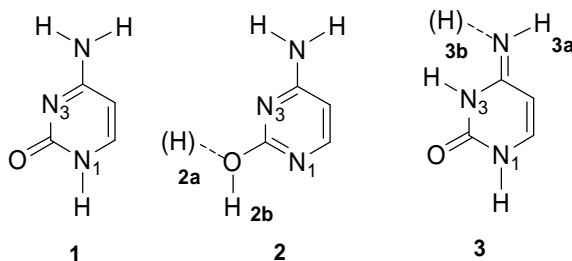
Előre kell bocsátanom, hogy a munka a vártnál lassabban haladt, s a tervezettnél sokkal szűkebb területre korlátozódott. Ennek két, egymással összefüggő oka volt: 1. A hagyományos kvantumkémiail számítások egy-egy statikus molekulaszervezetre vonatkoznak; pl. esetünkben különböző izomereket-tautomereket egyedi molekulaként számolnak. (Az atomi mozgások legfeljebb a harmonikus, kisamplitúdójú rezgések figyelembe vételében jelennek meg.) A legújabb időben vált csak lehetővé, hogy egy kémiai átalakulást annak folymatában, dinamikailag vizsgáljunk. Miután egy ilyen számítógép-program kifejlesztésébe magam is bekapcsolódtam [1] s a programhoz így hozzájutottam [2], a *citozin* tervezett tautomerizációjának vizsgálatában áttértem ilyen, dinamikai számításokra. Ez a munka összesen mintegy 1 millió (!) komplett ab initio számítást igényelt (összeadva több hónap tiszta gépidőt fogyasztott), s nagyon lelassította az előrehaladást. 2. Ugyanakkor az eredményt annyira újnak és érdekesnek találtam, hogy a legszélesebb olvasókörrrel bíró, legrangosabb folyóiratoknál próbáltam meg közzétenni. Először a *Science* -nél, majd a *Nature Chemical Biology*-nál próbálkoztam, mindkét esetben sikertelenül. Nem sikerült publikálnom a *J. Am. Chem. Soc.*-nál sem. (Az elutasítás mindhárom esetben arra hivatkozott, hogy a téma mégsem eléggé általános.) Ezekkel a próbálkozásokkal - a kéziratot mindig lényegesen megváltoztatva - több mint egy év eltelt; végül 2007 őszén a *Chem. Phys.* -hez nyújtottam be, s a napokban jelent meg - egyelőre az elektronikus formája (Közl. Jegyzék 12.) Megjegyzem, hogy egész tudományos pályafutásom alatt ilyen problémával nem találkoztam, valószínűleg elragadott a téma, s a léccet túl magasra tettem ...A végeredmény az, hogy a pályázat fő irányvonalát jelentő biomolekulák területén egyedül a citozin dinamikai vizsgálatáról tudok beszámolni - bár azt remélem, kiderül, hogy ez valóban nagyon új, érdekes eredmény.

### 1. Tautomerizáció

#### 1a) A citozin tautomerizációja (Közl. Jegyzék 4, 6, 12)

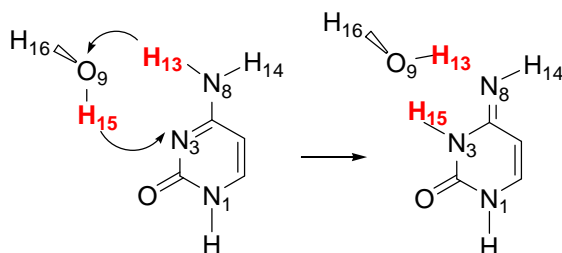
Jól ismert, hogy a citozin egyike a négy nukleotid bázisnak, melyek a DNS felépítésében részt vesznek. A kettős spirál szerkezetét a bázismolekulák között kialakuló hidrogénhidak határozzák meg. A bázisok jellegzetessége, hogy a heteroatomhoz kötött protonok helye megváltozhat (tautoméria; protonvándorlás), ami alapvetően megváltoztatja a lehetséges H-híd-kapcsolódásokat. A spontán genetikai mutációk egyik lehetséges okaként már Watson és Crick a ritka tautomerek megjelenését nevezte meg [3]. A citozinnak sok elvileg lehetséges tautomerje közül 3 olyan van, melyek energetikailag egy nagyon szűk, 1-2 kcal/mol szélességű tartományba esnek; e három szerkezetet az 1. ábra mutatja. **1** a "kanonikus" forma, ez van a DNS-ben természetes körülmények között. Szabad állapotban, az izolált molekula legstabilabb szerkezete a **2b** - minden kísérleti (spektroszkópiail) eredmény és a kvantumkémiail számítások (kivéve a DFT-t!) megegyeznek ebben.

Külön érdekességet jelent a **3a**, imino-forma: ennek lehetséges jelenlétére több spektroszkópai munka utal, de általában ezt tekintik a kis valószínűségű, "ritka" formának. Saját korábbi számításaim az energia, és a szabadentalpia alapján is, nagyobb stabilitást jósolnak [4].



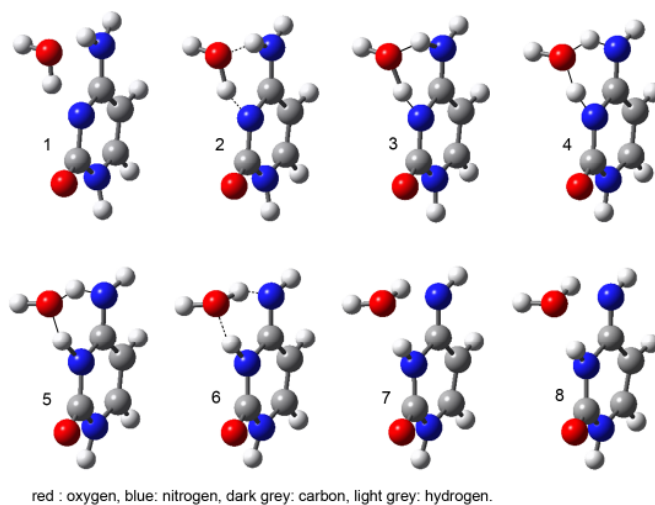
1. ábra. A citozin három, alacsony energiájú tautomerje. A kevésbé stabil rotamereket szaggatott vonallal és zárójelbe tett H-atommal jeleztük.

A DNS-ben az **1** tautomer az  $N_1$  atomnál glükózidos kötéssel kapcsolódik a cukorhoz, így **2**-be nem alakulhat át. Ezért a szóbjövő tautomerizáció a 2. ábra szerinti **1**  $\rightarrow$  **3** transzformáció lehet. Ezt vizsgáltam a projekt keretében azzal a céllal, hogy a tautomerizáció folyamatát, részletes mechanizmusát kövessük. Ennek érdekében molekuladinamikai számításokat végeztem.



2. ábra. A citozin víz által közvetített tautomerizációja az amino-oxo formából az imino-oxo formába.

A módszer [1] keveri a kvantummechanikai, illetve klasszikus elméletet, s a Born-Oppenheimer közelítést alkalmazza: tetszőleges kiindulási geometriában az elektronok Schrödinger-egyenletéből (az adott kvantumkémiai közelítésben) meghatározzuk az energiát és a magokra ható erőt; a magoknak az adott hőmérséklettel összhangban levő, véletlenszerű induló sebességeket adunk, s elmozdulásukat a klasszikus mechanika törvényei szerint számoljuk. A trajektória ezután úgy adódik, hogy minden egyes pontban az elektron-Schrödinger-egyenletet mindig újra, a hullámfüggvény teljes konvergenciájáig megoldjuk — szemben az elterjedt Car-Parrinello módszerrel (CP) [5], mely a hullámfüggvényt csak "propagáltatja". A kvantumkémiai számításokat tehát "menet közben" (*on the fly*) végzi a program. Szemben a CP-vel, a rendszer itt pontosan a Born-Oppenheimer hiperfelületen halad. Az egyes trajektóriák az induló sebességekben különböznek. A számítások volumene rendkívül nagy volt, több mint 300 trajektóriát futtattam, melyek egyenként 3000-5000 pontból álltak, 1 fs felbontással. Ez összesen kb. egy millió komplett kvantumkémiai számítást (energia és deriváltak) jelentett, ami rekordszámba mehet - bár persze az ab initio számítások érthetően csak a viszonylag szerény, B3LYP/3-21G szinten készültek. (A számításokat 6-10, dualprocesszoros PC-n végeztem, parallel üzemmódban.)



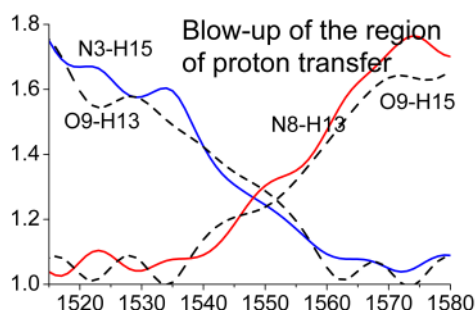
3. ábra. A tautomerizációs trajektóriából vett pillanatképek az átmenetről: az 1 fs-os felbontással végzett szimuláció minden ötödik geometriáját mutatjuk (35 fs időtartam).

Az elszigetelt, szabad citozin molekula esetében tautomerizáció nem várható, mert az átmenet gátja - kvantumkémiai számítások szerint - igen magas, 30-40 kcal mol<sup>-1</sup>. Mintegy 50 trajektóriát mégis lefuttattunk, mesterségesen magas hőmérsékleteken. Érdekes megfigyelés volt, hogy pl. , ~2000 K-en (!) az amino-csoport még stabil marad, miközben már a CH-kötések és a gyűrű is kezd felszakadni. A továbbiakban a víz által közvetített tautomerizációt modelleztük, olyan formában, hogy egy vízmolekulát tettünk a rendszerbe, a 2. ábrán látható módon. E konfiguráció környékén a víz lokális minimumban tud kötődni; az induló geometria ennek közelében volt, s 350 K hőmérsékletnek megfelelő véletlen sebességekkel indítottuk a trajektóriát. (A citozin és a víz kölcsönhatását korábban már külön tanulmányoztuk, s erről pl. előadásokon számoltam be (Közl. Jegyzék 1, 2, 5 6)). A több száz trajektória közül egyetlen sikeres esemény volt, melyet a 3. ábrán reprodukáltam. Érdekes az a megfigyelés, hogy ha a rendszer egyszer az átmeneti állapot közelébe jut - ehhez a víz nem csak közel, hanem megfelelő orientációban kell legyen - az átalakulás 15 - 20 fs alatt lejátszódik (2-6. lépés az ábrán). (Összehasonlításképpen: egy X-H nyújtási molekularezgés periódusideje ~10 fs.)

Mivel egyetlen sikeres eseményt találtunk, természetesen statisztikát (termodinamikát) nem tudunk csinálni. De információt kaptunk az (egyik lehetséges) átmenet mechanizmusáról. A trajektória részletes vizsgálata szerint, a megfigyelt tautomerizáció egy koncertált (összehangolt) — köztitermék nélküli —, szinkron folyamat, melyet a részt vevő H-atomok mozgásának igen erős csatoltsága jellemez. Ezt illusztrálja a 4. ábra: a két releváns X-H kötés, O<sub>9</sub>-H<sub>15</sub> és N<sub>8</sub>-H<sub>13</sub> egyszerre kezdenek megnyúlni 1530-1540 fs táján, s mozgásuk végig parallel; s ahogy ezek szakadnak, alakul ki a két új kötés, N<sub>3</sub>-H<sub>15</sub> és O<sub>9</sub>-H<sub>13</sub>. A két hidrogén mozgása végig teljesen szinkronizált.

A citozin-víz komplex egyike lehet a legnagyobb rendszereknek, melyre részletes ab initio dinamika készült. Az eredmény azonban erősen támadható amiatt, hogy az alkalmazott módszer (DFT) és különösen a nagyon kicsi bázisfüggvény-készlet durva közelítést jelent csak. Itt jegyzem meg: a dinamikában használt B3LYP/3-21G szinten az átmenet gátja mindössze 6 kcal/mol — ezért volt egyáltalán esély a tautomerizáció megfigyelésére —, míg irodalmi és saját számításaim szerint pontos módszerek szerint ennek kb. háromszorosa, 18-19 kcal/mol. Ezért tesztszámításokat végeztem a következő módon. Az eredeti trajektóriából az átmeneti állapot két oldalán 200 konfigurációt kivettem, s ezen pontokban két magasabb szintű közelítéssel is (B3LYP/6-31G\*\*); ill.

MP2/aug-cc-pVDZ) újraszámoltam az energiát: a publikált közleményben (Publ. Jegyz. 12) bemutatom, hogy — megfelelő energiaeltolással — a két pontosabb számítás jól követi az eredeti energiafüggvényt. Joggal remélhetjük tehát, hogy a számított trajektória a valóságot tükrözi.



4. ábra. Az X-H távolságok változása a tautomerizációs átmenet körül: x-tengelyen az idő fs-ban, y-tengelyen a távolság 100 pm egységben.

Összefoglalva: a citozin tautomerizációját a víz jelentősen megkönnyíti; az atomi elmozdulások követése szerint a folyamat egy koncertált, szinkron reakciómechanizmus szerint, gyorsan (10-20 fs) játszódik le, s a két részt vevő hidrogén igen erős csatolása jellemzi.

#### 1b) Tipikus tautomer molekulapárok relatív energiája (tesztszámítások) (Közl. Jegyz. 11)

Fenti és korábbi munkáim arra a tapasztalatra vezettek, hogy a tautomerizáció kvantumkémiai leírása — szemben például a sokkal megbízhatóbban számolható konformációs változásokkal — nagyon kényes feladat. Ez azzal értelmezhető, hogy a tautomerek elektronszerkezeti teljesen különböznek, lényegében külön molekulákat jelentenek. A következőkben vázolt munkában ezért szisztematikusan igyekeztem azt felmérni, hogy tautomer-párok relatív energiáját mennyire lehet megbízhatóan számolni. Négy molekulapárra végeztem számításokat: acetaldehid - vinilalkohol, acetaldimin - vinilamin, formamid - formamidsav, malonaldehid diketo-, ill. enol formája. Az 1. táblázatban példaképp az acetaldimin - vinilamin párt mutatom be. Látható, hogy a 3-21G bázis — mely ugyan nagyon kicsi, de számos kérdésre, pl. konformációs energiákra már értelmes választ ad — itt még kvalitatíve is rossz eredményre vezet. Ennél sokkal nyugtalanítóbb, hogy már magasnak tekintett szintek között is durva különbségek adódnak: pl. az MP2/6-311++G(3d,3p)-hez f ill. d polarizációs függvényeket adva (MP2/6-311++G(3df,3pd)), az energiakülönbség 1.49-ről 0.66 kcal/mol-ra csökken. Úgy látszik, a *coupled cluster* számítások kevésbé érzékenyek a bázisra, a PVTZ és az aug-PVTZ között már alig van különbség, de a biztos kijelentéshez még el kellene érni a PVQZ szintet, amit egyelőre nem sikerült. Mindenesetre az irodalomban közölt rengeteg eredménnyel szemben kritikusak kell legyünk: ahhoz, hogy a kis, 1-2 kcal/mol különbséget akár csak 0.5 kcal/mol pontossággal le tudjuk írni, legalább a CCSD(T)/PVTZ szintig kell elmenni. Az előzetes eredményeket konferencián már bemutattam (Közl. Jegyz. 11), a teljes munkát néhány kiegészítés után a jövőben tervezem publikálni.

**1. táblázat.** Keto-enol relatív energiák:  
az acetaldimin - vinilamin pár

	Acetaldimine E+132 au	Vinylamine E+132 au	$\Delta E$ kcalmol <sup>-1</sup>
<b>RHF</b>			
/3-21G	-0.320805	-0.326443	<b>-3.54</b>
/6-31G(d,p)	-1.080903	-1.075154	<b>3.61</b>
/6-311G(d,p)	-1.107668	-1.103987	<b>2.31</b>
/6-311++G(2d,2p)	-1.118126	-1.115103	<b>1.90</b>
/6-311++G(3d,3p)	-1.119075	-1.115368	<b>2.33</b>
/6-311++G(3df,3pd)	-1.122762	-1.120161	<b>1.63</b>
<b>B3LYP</b>			
/3-21G	-1.207746	-1.209913	<b>-1.36</b>
/6-31G(d,p)	-1.958724	-1.954299	<b>2.78</b>
/6-311G(d,p)	-1.991017	-1.989799	<b>0.76</b>
/6-311++G(2d,2p)	-2.001092	-2.001065	<b>0.02</b>
/6-311++G(3d,3p)	-2.001669	-2.001153	<b>0.32</b>
/6-311++G(3df,3pd)	-2.005619	-2.006049	<b>-0.27</b>
<b>MP2</b>			
/3-21G	-0.612823	-0.609018	<b>2.39</b>
/6-31G(d,p)	-1.529613	-1.521994	<b>4.78</b>
/6-311G(d,p)	-1.576867	-1.571923	<b>3.10</b>
/6-311++G(2d,2p)	-1.616024	-1.614151	<b>1.17</b>
/6-311++G(3d,3p)	-1.622827	-1.620453	<b>1.49</b>
/6-311++G(3df,3pd)	-1.670529	-1.669476	<b>0.66</b>
/PVTZ	-1.667207	-1.665153	<b>1.29</b>
/aug-PVTZ	-1.678035	-1.676821	<b>0.76</b>
/PVQZ//aug-PVTZ	-1.713596	-1.712580	<b>0.65</b>
/aug-PVQZ//aug-PVTZ			
<b>CCSD//MP2<sup>1</sup></b>			
/PVTZ	-1.740236	-1.737541	<b>1.69</b>
/aug-PVTZ	-1.758417	-1.755714	<b>1.70</b>
/PVQZ	-1.820783	-1.818900	<b>1.18</b>
<b>CCSD(T)/MP2<sup>1</sup></b>			
/PVTZ	-1.765141	-1.762240	<b>1.82</b>
/aug-PVTZ	-1.784369	-1.781576	<b>1.75</b>
/PVQZ	-1.848375		<b>---</b>

<sup>1</sup>A coupled cluster energiaszámítások az MP2/6-311++G(3df,3pd) geometriában történtek.

## 2. Elméleti fejlesztések

### 2a) Molekuladinamika [1] (Közl. Jegyzék 3.)

Pulay ötletéből egy olyan módszert dolgoztunk ki, mely molekuladinamikai számításokban drámaian felgyorsíthatná a kvantumkémiai számításokat. E kutatás nagyobbik része még a pályázat előtt folyt, de ide sorolható a publikálási munka jelentős része, s főleg: további próbálkozásaim a publikáció után.

A "Fock mátrix dinamika" néven leközölt eljárás lényege az, hogy az SCF iterációs eljárást egy adott lépésben olyan Fock mátrix-szal indítjuk, melyet néhány előző lépés (4-7) Fock mátrixaiból extrapoláltunk. Több Hartree-Fock számítás példáján mutattuk be, hogy az SCF konvergencia így általában 2-3 lépésben elérhető. A módszer azonban még nem elegendően stabil, néha váratlanul hosszabb iterációra van szükség, s ezt a hibát később sem tudtam kijavítani.

## 2b) Elektronkorreláció (Közl. Jegyzék 13.)

A kvantumkémiai módszerek fejlesztése területén Szalay Péter publikált egy legújabb eredményt, a *Multireference (MR) Averaged Coupled Pair Functional (ACPF)* és a *MR Averaged Quadratic Coupled Cluster (AQCC)* módszerrel kapcsolatban. Ezek az elektronkorreláció magas szintű leírását adó módszerek, melyek kifejlesztése jelentős részben az ő nevéhez fűződik. A most leírt új verzió kiküszöböli az eredeti megfogalmazás azon hibáját, melyet a közelfekvő elektronállapotok esetében mutatott. A projekt témaköre szempontjából, a módszer különösen a gerjesztett állapotokra tervezett vizsgálatainkhoz lehet fontos.

## 3. Infravörös spektrumok kvantumkémiai számítása (Közl. Jegyzék 7.)

E munkát Pongor Gábor és munkatársai végezték, a csoportunkban régebben kidolgozott SQM-módszer alapján. A vizsgálat tárgya egy benzimidazol-származék reakciója volt egy olyan szililező szerrel (BSTFA), mely két tautomer formában létezhet, s ennek megfelelően a szililezés eredménye két különböző termék lehet. Az SQM-mel számított és a mért infravörös spektrum összehasonlítása alapján a terméket egyértelműen azonosítani lehetett. A munka nyilván csak nagyon lazán kapcsolódik a projekthez, de az eredményeknek azon része, mely az elméleti és a kísérleti spektrum összehasonlítását vizsgálja, felhasználható lehet biomolekulákra is.

## Hivatkozások

- [1] P. Pulay, G. Fogarasi, Chem. Phys. Lett. 386 (2004) 272-278
- [2] PQS version 2.4, Parallel Quantum Solutions, 2013 Green Acres Road, Fayetteville, AR 72703. Private version from Peter Pulay.
- [3] J.D. Watson, F.H.C. Crick, Nature 171 (1953) 964-967.
- [4] G. Fogarasi, J. Phys. Chem. A, 106 (2002) 1381-1390.

## Jelentősebb eltérések a költségtervtől (T43628 pályázat, Fogarasi Géza)

**A.** Két alkalommal kértem az OTKA-tól engedélyt **keretátcsoportosításra**, ezen leveleim megfelelő részét idézem:

### **A1.** 2004. december 1.

A külföldi konferenciákra tervezett évi 250 eFt-ot ez évtől, tehát a 2004-2006. évekre, 400 eFt-ra szeretném felemelni, keretátcsoportosítással. 2004-ben 100 eFt-ot az 1. (személyi juttatások) rovaton levő maradvány, 50 eFt-ot a 3.3. (egyéb költség) rovat fedez. Hasonló átcsoportosítást tervezek a későbbi évekre is.

### **A2.** 2005. augusztus 30.

a) A külföldi konferenciákra tervezett keretet sajnos jelentősen túlléptem — annak ellenére, hogy az eredeti, évi 250 eFt-ot a 2004-2006. évekre már tavalyi kérésem alapján (2004. dec. 1.), átcsoportosítással 400 eFt-ra emeltem. A túllépésnek az az oka, hogy az idén két nagyon drága konferencián (Délafrika, USA) vettem részt előadóként. Ez szakmailag világosan indokolható, mert mindkettő nagyon színvonalas, elismert konferencia-sorozatok része (*WATOC 2005, Cape Town*, ill. *Gordon Research Conference, Bates College, Maine*).

Ennek alapján a következő átcsoportosítást szeretném tenni:

1.6. ( személyi )	→	3.1. (külf. konf. )	:	320 eFt
1.7. ( napidíj )	→	3.1. (külf. konf. )	:	100 eFt

Összesen 1. → 3.1. **420 eFt**

b) A nagyon alultervezett készletbeszerzési keretem bővítésére kérem továbbá a következő átcsoportosítás engedélyezését:

2. ( munkaadói járulék )	→	3.2. (készletbeszerzés )	:	<b>100 eFt</b>
--------------------------	---	--------------------------	---	----------------

Az **A1.** átcsoportosításra vonatkozó engedélyező levelet megkaptam. Az **A2.**-re nincs nálam válaszlevél, de remélhetőleg erre is megvan az engedély. Ha nem így lenne, ezúton kérem az utólagos engedélyezést.

**B.** Külön kérvényeim alapján, az alábbi **speciális kiadásokra** kaptam engedélyt:

B1. Internethozzáférés otthonról, havi kb. 9 eFt értékben.

B2. Egy speciális ergonomikus szék vásárlása 48 eFt értékben.

B3. Számítógépterem klímaberendezésének cseréje (saját pályázatom mellett, Szalay Péter 47182 és Császár Attila 47185 pályázatok terhére), összesen 550-600 eFt értékben.

Budapest, 2008-03-02.

Fogarasi Géza