Keto-enol Tautomerism of Aryldiketo Acids in Aqueous Solution: NMR Spectroscopy and Cyclic Voltammetry Study

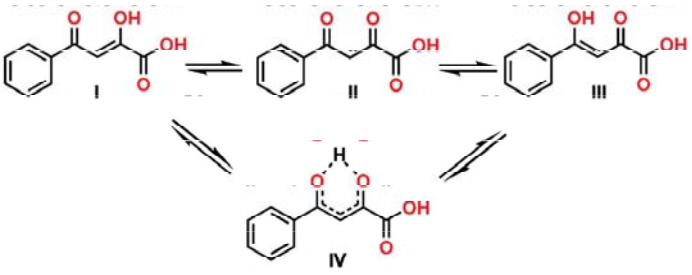
Tatjana Ž. Verbić✉, Mire Zloh*, Dalibor M. Stanković**, Milica N. Sentić, Dragan D. Manojlović, Ivan O. Juranić

Faculty of Chemistry, Univ. of Belgrade, Studentski Trg 16, Belgrade, Serbia, ✉tatjanad@chem.bg.ac.rs;
**The School of Pharmacy, Univ. of London, 29/39 Brunswick Square, London WC1N 1AX, United Kingdom;*
***Innovation Center of the Faculty of Chemistry, Univ. of Belgrade, Studentski Trg 16, Belgrade, Serbia.*

Introduction

б-Diketo acids and derivatives are the first, and so far the most successful group of compounds that selectively inhibit the integration of HIV-1 viral DNA in the host genome.[1-4] These compounds simultaneously exist in two enolate forms (**I** and **III**), conformationally locked by the pseudo-ring; and one diketo form (**II**) having two rotatable bonds responsible for the conformational flexibility (***Scheme 1***).[5-7]

Aryldiketo acids (ADK) act by functional sequestration of Mg^{2+} ion, an integral part of the active center of HIV-1 integrase (IN). This enzyme is responsible for integration of viral DNA in host genome. It was shown that ADK complexation ability depends on tautomeric form that is dominant in solution,[8] and that Mg^{2+} preferentially reacts with enolate form **I** (***Scheme 1***).[9] Furthermore, hydrolytic C–C bond cleavage of *б*-diketones by *б*-ketolases (mammals liver enzyme) is sensitive to tautomeric form in which a *б*-diketone is present in solution.[5,9,10]



Scheme 1. Tautomerization of 4-phenyl-2,4-dioxobutanoic acid (H- ADK) in aqueous solution.

Detailed study of keto-enol tautomerism in the set of eleven 4-alkyl- or 4-aryl-2,4-diketo acids revealed that in aprotic solvent ($CDCl_3$) enolate **I** is the predominant form (98%).[5] In aqueous media, tautomerization of aliphatic 2,4-diketo acids was discussed in the pH range 1.5–10. As aromatic 2,4-diketo acids (ADK) are much less soluble, their tautomerization was studied in solutions with pH ≥ 5.5, but three tautomeric forms (**I–III**) could not be distinguished due to the formation of pseudodienolate (**IV**) and fast interconversion between two enolate forms. No data on keto-enol tautomerism of ADK in aqueous solution with pH ≤ 5.5 was published so far.

Congeneric set of 3-, 4-, 3,4- and 2,5-phenyl substituted ADK was previously synthesized and studied within our research group.[11] During routine characterization, mass spectra of 3-alkyl substituted compounds showed significantly more intensive 2(M–1)+Na than 2M–1 peaks, opposite to other derivatives. UV-Vis and ^{1}H NMR spectroscopy revealed that of all studied ADK, just unsubstituted derivative (H- ADK, ***Scheme 1***), 3,4-di-Me-, 2,5-di-Me-, and *б*-naphtyl- ADK showed significant spectral changes upon the addition of Mg^{2+}, indicating their better complexation ability with Mg^{2+} ion.

The primary aim of this work was to study tautomeric equilibria of H-, 2,5-di-Me-, 3,4-di-Me-, and 4-Me- ADK (***Table 2***) in the aqueous solution within pH range 1–10. *б*-Naphtyl-ADK was omitted due to its insufficient solubility, and 4-Me- derivative was used as congener that showed no complexation ability with Mg^{2+} ion.[11]

Results and Discussion

ADK are diprotic acids weakly soluble in water, especially in acidic media where they are present in molecular (H_2A) form. Usable NMR spectra were obtained only for H- ADK after overnight signal acquisition (***Fig. 1***). Full structure-spectra assignments were achieved using COSY, HMQC, and HMBC spectra of H- ADK in CF_3COOD (***Table 1***), where H- ADK is present in its molecular (H_2A) form.

When dissolved in highly acidic media (CF_3COOD), H- ADK is in molecular (H_2A) form and the enolate **I** is predominant (***Fig. 1***). The singlet at 7.23 *ppm* in ^{1}H NMR spectrum is the signal of the vinyl group atom H_3 (***Table 1***). As the exchange rate between H_3 and D atom in CF_3COOD is fast, the integral of this signal is smaller than expected. Furthermore, the lack of a singlet around 4.5 *ppm* in ^{1}H NMR spectrum (not shown), as well as

a signal around 50 *ppm* in ^{13}C NMR spectrum which would correspond to >CH_2 group in diketo form **II,** confirm that diketo form does not exist. Signals at 7.33 *ppm* (*s*, 1H), 7.80 *ppm* (*t*, 1H), and 8.12 *ppm* (*d*, 2H), although very weak, confirm the existence of enolate **III**, but its concentration is negligible.

Table 1. ^{1}H and ^{13}C NMR chemical shifts of H- ADK in CF_3COOD (atom numeration as given in Table 2).

Atom	C_1	C_2	C_3	H_3	C_4	C_5	C_6	H_6	C_7	H_7	C_8	H_8
δ (*ppm*)	163.41	161.59	95.06	7.23[a]	189.46	129.85	124.05	7.96[b]	125.05	7.50[c]	131.13	7.64[d]

[a](*s*, 1H); [b](*d*, *J* = 7.37 Hz, 2H); [c](*t*, *J* = 7.90 Hz, 2H); [d](*t*, *J* = 7.47 Hz, 1H)

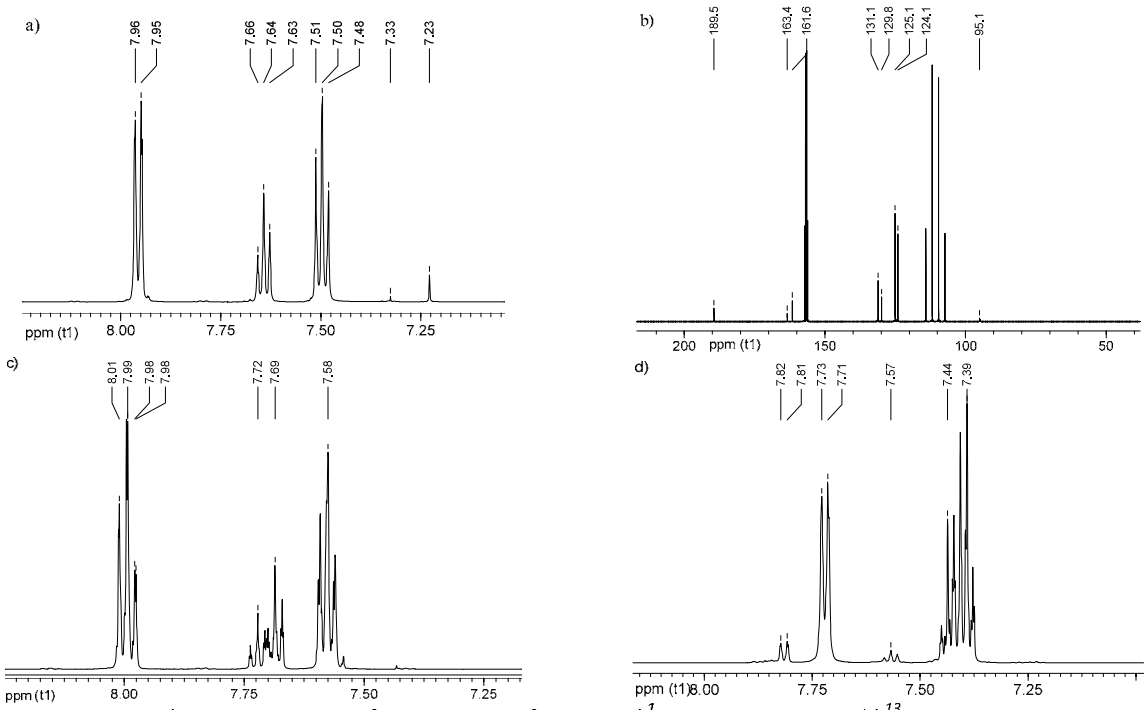


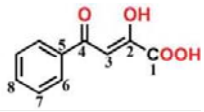
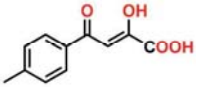
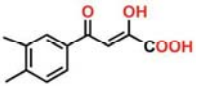
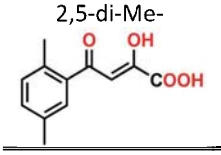
Figure 1. Characteristic parts of NMR spectra of H- ADK: a) ^{1}H NMR in CF_3COOD, b) ^{13}C NMR in CF_3COOD, c) ^{1}H NMR in D-acetate buffer pD 4.81, d) ^{1}H NMR in H-carbonate buffer pH 9.60.

When dissolved in highly acidic media (CF_3COOD), H- ADK is in molecular (H_2A) form and the enolate **I** is predominant (***Fig. 1***). The singlet at 7.23 *ppm* in ^{1}H NMR spectrum is the signal of the vinyl group atom H_3 (***Table 1***). As the exchange rate between H_3 and D atom in CF_3COOD is fast, the integral of this signal is smaller than expected. Furthermore, the lack of a singlet around 4.5 *ppm* in ^{1}H NMR spectrum (not shown), as well as a signal around 50 *ppm* in ^{13}C NMR spectrum which would correspond to >CH_2 group in diketo form **II,** confirm that diketo form does not exist. Signals at 7.33 *ppm* (*s*, 1H), 7.80 *ppm* (*t*, 1H), and 8.12 *ppm* (*d*, 2H), although very weak, confirm the existence of enolate **III**, but its concentration is negligible.

Due to long signal acquisition time, for all other solutions with different pH values, just ^{1}H NMR spectra were recorded. According to previously determined acidity constants (***Table 2***),[14] pD 4.8 (HA^- form) and pH 9.6 (A^{2-}) were chosen as solutions where just one form of ADK exists.

No signals of enolate **III,** and diketo form **II** were observed in ^{1}H NMR spectra in acetate (HA^- form) and carbonate (A^{2-}) buffers (***Fig. 1c*** and ***1d***). The reason for rather complicated structure of ^{1}H NMR spectrum of HA^- (***Fig. 1c***) is possible rotation around C_3–C_4 single bond and the presence of *Z* and *E* isomers of the –CH=CH(OH)– bond. In A^{2-} form *E* isomer is predominant, due to electrostatic repulsion between $–COO^-$ and $–CH=C–O^-$. Since π-electron delocalization occurs in keto-enol part of the molecule, the

Table 2. Spectrophotometrically determined pK_a values.[14]

Compound (ADK)	$pK_{a1} \pm SD$ $pK_{a2} \pm SD$
H- 	2.06±0.03 7.56±0.02
4-Me- 	2.22±0.05 7.99±0.02
3,4-di-Me- 	2.09±0.04 7.92±0.04
2,5-di-Me- 	2.39±0.04 7.23±0.04

distinction between tautomers is not possible. Thus, the origin of signals at 7.82 *ppm* (*d*) and 7.57 *ppm* (*t*) is still unsolved (***Fig. 1d***).

As no data about 4-Me-, 3,4-di-Me-, and 2,5-di-Me- ADK were obtained by NMR, cyclic voltammetry was used as a method with much lower detection limit than NMR spectroscopy. Behavior of all four studied ADK was monitored in solution where these ADK are present in H_2A (pH 1), HA^- (pH 5), or A^{2-} (pH 10) form (***Fig. 2***).

Under the electrochemical conditions used, H-, 3,4-di-Me-, and 4-Me- ADK in H_2A form (***Fig. 2a***) show sharp anodic peak, and no reversible cathodic peak. As pH value is raised, the –COOH group dissociate and, as expected, the peak is lowered and moved to lower oxidation potential (***Fig. 2b***). With the dissociation of –OH group (A^{2-} form) and π-electron delocalization, peaks are completely lost (***Fig. 2c***).

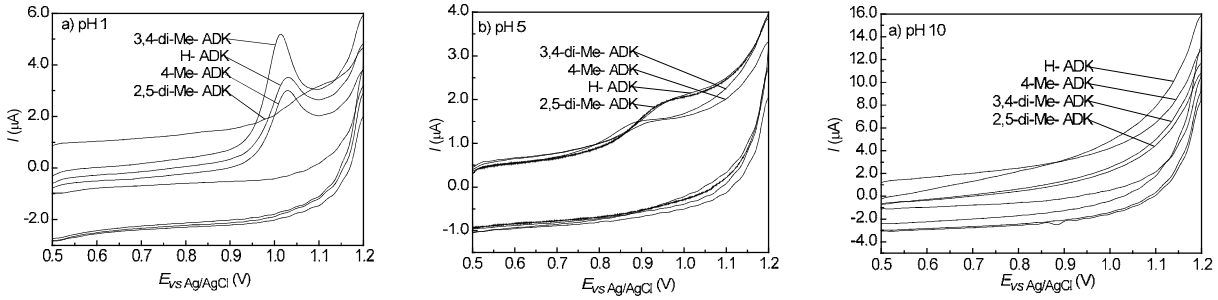


Figure 2. Cyclic voltammograms of studied ADK in aqueous solutions of different pH values.

The cyclic voltammogram (CV) of 2,5-di-Me- in H_2A form (***Fig. 2a***) is different than corresponding CVs of other ADK: characteristic peak is moved toward higher potentials and is not as sharp as peaks of other compounds. Still, the areas under all four observed peaks are the same, which implies that the electron exchange reaction is the same for all four ADK. Derivatization of CV at pH 1 shows that this wide peak consists of two overlapped peaks. This was confirmed in solutions with pH 2 and pH 3 (data not shown), where the mixture of H_2A and HA^- forms exists. If two different tautomers were present in solution, peaks in CV would be wide and overlapped. So, we may offer a possible explanation: in H_2A form, 2,5-di-Me- exists simultaneously in two tautomeric forms, probably as enolates **I** and **III** (***Scheme 1***). As the shape of all other peaks is the same, and as we have confirmed that the dominant form of H- ADK in solutions within pH range 1-10 is enolate **I**, we may conclude that, when in H_2A form, 3,4-di-Me- and 4-Me- ADK are also present as enolate **I**.

The conformation and electronic properties of studied ADK in their H_2A form, calculated using semiempirical molecular-orbital method, suggests possible reasons for different behaviour of 2,5-di-Me- ADK.

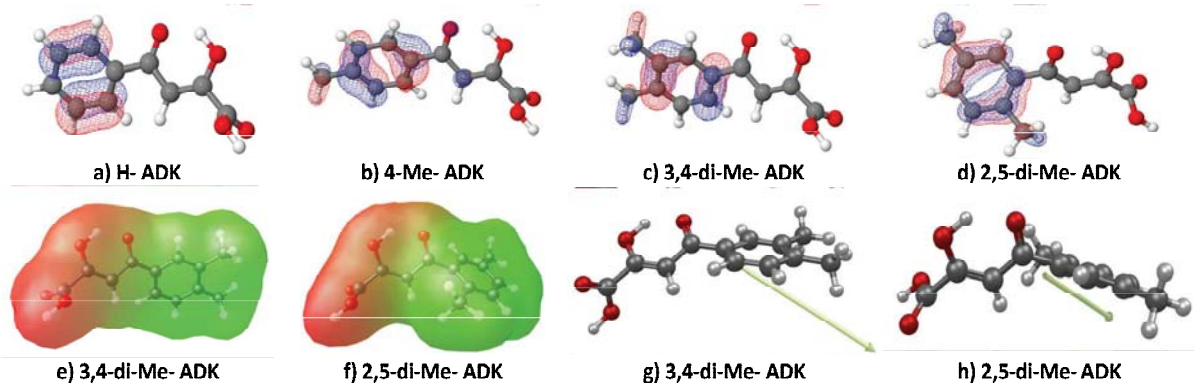


Figure 3. (a-d) HOMO of studied ADK; (e and *f) polar and apolar surface area for 3,4-di-Me- and 2,5-di-Me- derivatives;(g* and *h) dipole moments for 3,4-di-Me- and 2,5-di-Me- derivatives.*

Presence of *ortho*-alkyl substituent in 2,5-di-Me- ADK causes large torsion between aryl- and dioxo-carboxyl moiety. HOMO-s are located on phenyls in all studied congeners (***Fig. 3a - d***). Approach to electrode of the dioxo-carboxyl moiety, one most prone to electrochemical oxidation in 2,5-di-Me- derivative, is hindered in comparison to other compounds studied (shown by spatial arrangement of polar and apolar surface areas for 3,4-di-Me- and 2,5-di-Me- derivatives, ***Fig. 3e*** and ***3f***). Calculated heats of formation of all compounds in their molecular and radical cation forms show that 2,5-di-Me- derivative is the most stable one. Along with this, due

to its geometry, 2,5-di-Me- derivative has a lowest dipole in the set (***Fig. 3g*** and ***3h***), and consequently, can experience the attractive force in lesser extent than other derivatives when approaches the electrode.

Conclusion

Aryldiketo acids (ADK) complexation ability with Mg^{2+} ion (present in HIV-1 IN) and the hydrolytic C–C bond cleavage by *ß*-ketolases depend on predominant tautomeric form in which they are present at the reaction site. Thus, the keto-enol tautomerism of H-, 4-Me-, 2,5-di-Me-, and 3,4-di-Me- ADK in the aqueous solution within pH range 1–10 was studied by NMR spectroscopy and cyclic voltammetry. The NMR results for H- ADK showed that the predominant tautomeric form in the studied pH range is the enolate **I** (***Scheme 1***). The cyclic voltammetry data in acidic media (pH 1) suggest that the ratio of possible tautomeric forms for 4-Me- and 3,4-di-Me- ADK is similar to H- ADK, *i.e.* that enolate **I** is the dominant form in the solution, whilst two tautomeric forms of 2,5-di-Me- ADK (enolates **I** and **III**) may be present in solution. The higher oxidation potential observed for H_2A form of 2,5-di-Me- ADK at pH 1 may be due to non-planarity of molecule and steric hindrance imposed by *ortho* substituents.

Experimental

^1H, ^{13}C, COSY, HMQC, and HMBC NMR spectra were acquired using Bruker Avance 500/125 MHz NMR spectrometer at $t = 25 \pm 1$ ºC, and constant ionic strength $I = 0.1$ M ($NaNO_3$ was used to adjust the ionic strength). TSP was used as the internal standard for spectra calibration; chemical shifts (δ) are given in *ppm*. pH Values were measured using Corning 120 pH-meter equipped with Corning Ag/AgCl microelectrode and converted to pD according to relation: $pD = pH_{measured} + 0{,}4$.[12,13]
Cyclic voltammograms were recorded using CHI760B instrument (CH Instruments, USA). The cell was equipped with glassy carbon electrode and an accessory Pt electrode of larger surface (Model CHI221, cell top including Pt wire counter electrode) and Ag/AgCl reference electrode (Model CHI111). Scan speed 100 mV/s.
Initial 3D structures of compounds were generated by OMEGA,[15] and optimized on semiempirical molecular-orbital level using PM6 method,[16] implemented in MOPAC2009,[17] to root mean square gradient bellow 0.001 kcal/mol Å. Implicit solvation in water was used. Molecular orbitals were visualized by Jmol.[18]

Acknowledgment: The authors highly appreciate the help and comments provided by MSc Branko J. Drakulić, IChTM. The Ministry of Education and Science of Serbia supports this work. Grants 172035 and 172030.

Proučavanje keto-enolne tautomerije arildiketo kiselina u vodenoj sredini upotrebom NMR spektroskopije i ciklične voltametrije

*ß-Diketo kiseline i njihovi derivati su prva, i do sada najuspešnija, grupa jedinjenja koja selektivno inhibira proces integracije HIV-1 provirusne DNK u DNK ćelije domaćina. Poznato je da ß-diketoni u rastvoru podležu reakciji tautomerizacije i da afinitet ka kompleksiranju sa Mg^{2+} jonom (nalazi se u aktivnom centru HIV-1 integraze), kao i reakcija hidrolitičkog raskidanja C–C veze u diketo delu molekula ß-ketolazama (enzim jetre sisara) zavise od oblika u kom se ova jedinjenja nalaze na mestu dejstva. U ovom radu je, NMR spektroskopijom i cikličnom voltametrijom, proučavana keto-enolna tautomerija četiri jedinjenja iz grupe aril diketo kiselina (H-, 4-Me-, 2,5-di-Me- i 3,4-di-Me- ADK) u vodenom rastvoru u pH oblasti 1–10. Rezultati dobijeni NMR spektroskopijom pokazuju da se, u ispitivanoj pH oblasti, nesupstituisana ADK (H-) nalazi u obliku enola **I**, enolni oblik **III** je prisutan u tragovima, dok signali koji bi poticali od diketo oblika nisu vidljivi u spektrima. Prema cikličnim voltamogramima izveden je zaključak da se, kada su u molekulskom (H_2A) obliku, 4-Me i 3,4-di-Me- ADK u rastvoru nalaze takođe u obliku enola **I**, dok je u rastvoru 2,5-di-Me ADK prisutna smeša tautomera i to najverovatnije enola **I** i **III**. Izračunavanja dobijena semiempirijskim molekulsko-orbitalnim PM6 metodom nude moguće objašnjenje: veći torzioni ugao između aril grupe i ostatka molekula kod 2,5-di-Me ADK može biti razlog višeg potencijala na kome se ova ADK oksiduje, u poređenju sa ostalima.*

References

1. D. J. Hazuda, P. Felock, M. Witmer, A. Wolfe, K. Stillmock, J. A. Grobler, A. Espeseth, L. Gabryelski, W.
1. Schleif, C. Blau, M. D. Miller, *Science* **287** (2000) 646.
2. J. S. Wai, M. S. Egbertson, L. S. Payne, T. E. Fisher, M. W. Embrey, L. O. Tran, J. Y. Melamed, H. M. Langford, J. P. Guare Jr., L. Zhuang, V. E. Grey, J. P. Vacca, M. K. Holloway, A. M. Naylor-Olsen, D. J. Hazuda, P. J. Felock, A. L. Wolfe, K. A. Stillmock, W. A. Schleif, L. J. Gabryelski, S. D. Young, *J. Med. Chem.* **43** (2000) 4923.

3. H. G. Selnick, D. J. Hazuda, M. Egbertson, J. P. Guare, J. S. Wai, S. D. Young, D. L. Clark, J. C. Medina, *Patent* W09962513 A (1999) Merck & Co.Inc.
4. F. Toshio, Z. Tomokazu, *Patent* W099-JP1547 (1999) Shionogi & Co. Ltd.
5. L. Brecker, M. Pogorevc, H. Griengl, W. Steiner, T. Kappe, D. W. Ribbons, *New J. Chem.* **23** (1999) 437.
6. M. Sechi, A. Bacchi, M. Carcelli, C. Compari, E. Duce, E. Fisicaro, D. Rogolino, P. Gates, M. Derudas, L. Q. Al-Mawsawi, N. Neamati, *J. Med. Chem.* **49** (2006) 4248.
7. M. Huang, W. G. Richards, G. H. Grant, *J. Phys. Chem. A* **109** (2005) 5198.
8. M. Billamboz, F. Bailly, M. L. Barreca, L. De Luca, J. F. Mouscadet, C. Calmels, M. L. Andréola, M. Witvrouw, F. Christ, Z. Debyser, P. Cotelle, *J. Med. Chem.* **51** (2008) 7717.
9. V. O. Koz'minykh, E. N. Koz'minykh, *Pharm. Chem. J.* **38** (2004) 67.
10. M. Sechi, M. Derudas, R. Dallocchio, A. Dessi, A. Bacchi, L. Sannia, F. Carta, M. Palomba, O. Ragab, C. Chan, R. Shoemaker, S. Sei, R. Dayam, N. Neamati, *J. Med. Chem.* **47** (2004) 5298.
11. T. Ž. Verbić, B. J. Drakulić, M. Zloh, I. O. Juranić, *Lett. Org. Chem.* **5** (2008) 692.
12. P. K. Glasoe, F. A. Long, *J. Phys. Chem.* **64** (1960) 188.
13. K. Popov, H. Rönkkömäki, L. H. J. Lajunen, *Pure Appl. Chem.* **78** (2006) 663.
14. T. Ž. Verbić, B. J. Drakulić, M. F. Zloh, J. R. Pecelj, G. V. Popović, I. O. Juranić, *J. Serb Chem. Soc.* **72** (2007) 1201.
15. J. Boström, J. R. Greenwood, J. Gottfries, *J. Mol. Graphics Modell.* **21** (2003) 449. www.eyesopen.com
16. J. J. P. Stewart, *J. Mol. Mod.* **13** (2007) 1173.
17. J. J. P. Stewart, *J. Comput.-Aided Mol. Des.* **4** (1990) 1; J. J. P Stewart, *MOPAC2009*, Stewart Computational Chemistry, Colorado Springs, CO, 2009, http://OpenMOPAC.net
18. Jmol v11.2.10, an open-source Java viewer for chemical structures in 3D. http://www.jmol.org/