

RSC Advances



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. This *Accepted Manuscript* will be replaced by the edited, formatted and paginated article as soon as this is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

Thermodynamic and Kinetic Investigation of Monoketo-Aldehyde-Peroxyhemiacetal-(MKA), a Stereolabile Degradation Product of Dihydroartemisinin.

Cite this: DOI: 10.1039/x0xx00000x

Received 00th January 2012,
Accepted 00th January 2012

DOI: 10.1039/x0xx00000x

www.rsc.org/

D. Kotoni,^a M. Piras,^{a,d} W. Cabri,^b F. Giorgi,^b A. Mazzanti,^c M. Pierini,^{*a} M. Quaglia,^b C. Villani,^a F. Gasparrini.^{*a}

Peroxyhemiacetal MonoKeto-Aldehyde (MKA) is a DiHydroArtemisinin (DHA) derivative, still endowed with significant (on a ng-scale) in vitro antimalarial activity and neurotoxicity lower than that displayed by its precursor. Through literature it is known that MKA may be formed under essentially physiological conditions (organic-aqueous environments buffered at pH 7.4) or prepared by acid decomposition of DHA in water to afford yields that, however, do not reach 50%. A more convenient procedure of preparation is here reported at room temperature. Owing to its hemiacetal nature MKA was expected to occur in solution as a mixture of α and β epimers. In the present study this, in fact, has been demonstrated through dynamic chromatography measurements, which allowed to highlight that for MKA the $\beta \rightleftharpoons \alpha$ equilibration is about 3 times slower than for DHA. In consideration of the propensity of MKA to originate by DHA in mild both acid and base conditions, its formation appears possible during the stages of production and storage of its parent drug and also widely expected in physiological environment. Thus, the improvement of knowledge about this product of DHA degradation could have useful impact in the rational development of new methods for the analysis of purity of DHA in its pharmaceutical formulations, as well as in taking into account the biological activity expressible in vivo by the mixture of α and β epimers of MKA, whose equilibrium composition is a function of the met specific biological environment. In this context, in our study we elucidated the stereo-structure of the epimers of MKA through NMR measurements and performed a comprehensive thermodynamic investigation of the process that governs the related interconversion through Linear Solvation Energy Relationships (LSER) approach. A convincing rationalization of the whole findings has then been achieved through support from molecular modeling calculations.

Introduction

Malaria is a mosquito-borne infectious disease caused by eukaryotic protists of the genus *Plasmodium*, responsible for approximately 350-500 million cases of malaria worldwide, which, according to the World Health Organization 2010 World Malaria Report,¹ is killing around 781,000 people each year, a number corresponding to 2.23% of deaths worldwide. The majority of deaths are young children in Sub-Saharan Africa.² In the past years, the combination of artemisinin derivatives

with other conventional antimalarial drugs (artemisinin-based combination therapies or ACTs) was developed and is now the treatment of choice for the most lethal forms of malaria.³ Artemisinin (Qinghaosu, **1**, Figure 1) is a sesquiterpene lactone endoperoxide isolated from *Artemisia annua* L. that Chinese herbalists have traditionally used to treat malaria.^{4,5}

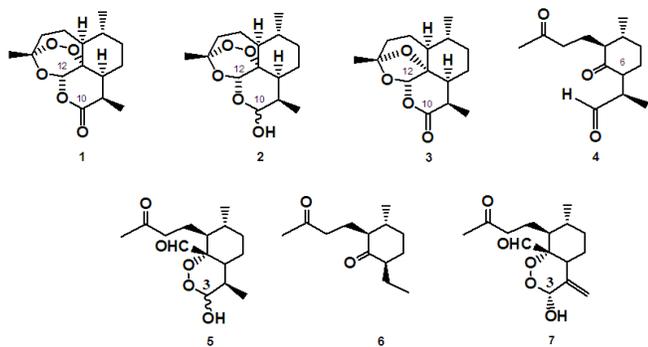


Figure 1. Chemical structures of artemisinin (**1**), dihydroartemisinin (**2**, **DHA**), 2-desoxyartemisinin (**3**), diketoaldehyde (**4**, **DKA**), monoketo-aldehyde (**5**, **MKA**), a new dicarbonyl compound (**6**) and an α -**MKA-ene** (**7**).

The peroxide group present in the 1,2,4-trioxane ring of artemisinin (**1**) and in its derivatives confers them antimalarial activity, but also makes artemisinins more difficult to handle in a drug-development setting compared to other conventional antimalarial drugs, as it is both a chemically and thermally sensitive group.⁶ Reduction of **1** by sodium borohydride in methanol⁷ produced dihydroartemisinin (**DHA**, **2**, Figure 1), which is also the main *in vivo* metabolite of artemisinin and provides improved antimalarial potency and a major elimination route.⁸ Conversion of the lactone carbonyl group at C-10 of **1** into the hydroxyl (hemiacetal) group in **2** yields a new stereochemical labile centre in the molecule, which, in turn, provides two hemiacetal interconverting epimers, namely, **2 α** and **2 β** . The **2 β** \rightleftharpoons **2 α** interconversion of **DHA** has been thoroughly investigated by our group from a thermodynamic⁹ and kinetic¹⁰ point of view, using both experimental data and theoretical calculations.

Given the relative instability of the artemisinin structure, the decomposition of both **1** and **2** under a variety of conditions has been frequent object of study starting from the 1980s. Artemisinin has been reported to be labile under acidic or basic treatment, but unexpectedly stable in neutral solvents when heated up to 150°C.¹¹ However, considerable changes were detected when heating was extended to 190°C for ten minutes,¹² with no prevalent product being formed (isolation yields ranging between 4 and 12%). In contrast, thermolysis of **2** at 190°C for three minutes produced mainly 2-desoxyartemisinin (compound **3**, Figure 1, 30% yield) and the diketoaldehyde **4** (**DKA**, Figure 1, 50% yield) as a temperature dependent mixture of stereoisomers, assumed to be epimeric at C-6.¹³ Twenty years later, these same compounds were recovered, albeit with lower yields, by heating solid **2** at 100 °C for 14 h under a nitrogen atmosphere.^{14a} In this case other minor compounds were observed, such as the already known peroxyhemiacetal **5** (Figure 1, 7% yield) and a hitherto unknown dicarbonyl compound **6** (Figure 1, 2% yield), also independently identified by another research group,¹⁵ and most likely formed by oxidation and subsequent decarboxylation of

its putative precursor **4**. Haynes and collaborators reported that comparable yields were observed when performing the reaction under air. Higher yields of **5** (48%) were obtained by decomposition of **2** under aqueous acidic solution.^{14a} Interestingly, **5** has also been obtained from **2** in much more mild conditions, under N-benzyl-1,4-dihydronicotinamide control in a MeCN/pH 7.4 buffer mixture employed as the solvent, that is to say, in an environment not too far from the physiological.^{14b} Therefore, this strongly suggest that **5** is formed *in vivo* as the result of **DHA** degradation. Peroxyhemiacetal **5** (hereafter named monoketo-aldehyde, **MKA**, for the presence of one ketone and one aldehyde group in its structure, in contrast with **DKA**), was isolated for the first time in 1993 by the group of professor Baker through exposure of the ethyl ether analogue of **2** (called β -arteether) to simulated stomach acid, and was reported to have significant antimalarial activity in an *in vitro* assay (IC₅₀ in the 0.99–2.2 ng/mL range).¹⁶ In their work, Baker and co-workers compare **MKA 5** with a reference compound **7**,¹⁷ bearing an olefinic moiety in position C-4. Such compound **7** has been crystallized and appears to have an *R* configuration at C3 in the solid state, that is, just the opposite of what it is found for **2** which, in the solid state, displays only the *S* configuration at the same hemiacetalic carbon. In 1995, Acton and Roth also obtained **5** in a 25 % yield through acid decomposition of both β -arteether and **DHA**.¹⁸ **MKA** not only is an active (on a ng-scale) *in-vitro* derivative of **DHA** (and probably an *in-vivo* metabolite^{14b}) but also displays a lower neurotoxicity when compared to its precursor.¹⁹ No further information is present in literature about the preferred configuration at C3 of **5** in the solid state, as it has never been crystallized. Being a hemiacetal like **2**, **MKA** was also expected to occur in solution as a mixture of epimers (at position C-3), namely, **5 α** and **5 β** (Figure 2).

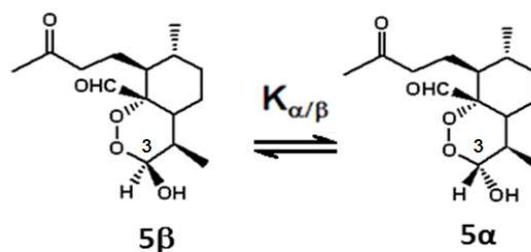


Figure 2. Chemical structure of the interconverting epimers of **5**: the **5 α** -epimer bears the hydroxyl group in the axial position, *anti* to the aldehyde group, whereas the **5 β** -epimer possesses an equatorial hydroxyl group, *syn* to the aldehyde group.

As with **DHA**, the interconversion of the two **MKA** epimers does, in fact, occur in solution and takes place on a chromatographic time scale; for this reason, separation of the two species under non-interconverting conditions must be a prerequisite of any analytical method aimed at quantitating the drug in active ingredients, pharmaceutical formulations, and biological fluids. In addition, in view of the biological activity

expressible in vivo by the mixture of α and β epimers, also of relevant importance appears to get information about the position assumed by the $5\beta \rightleftharpoons 5\alpha$ equilibrium as a function of the biological environment in which **5** is formed. In this perspective, a study aimed at better understanding the kinetic, thermodynamic, and mechanistic features of the $5\beta \rightleftharpoons 5\alpha$ equilibration is therefore strongly advisable. According to this, and basing the adopted methodology on the guidelines already set during the exhaustive investigation performed on the $2\beta \rightleftharpoons 2\alpha$ epimerization,^{9,10} we have carried out a thorough and modular study on the epimerization process of **5**, of which will be here presented in detail the thermodynamic aspects. Such findings have been achieved by resorting to in-depth linear solvation energy relationships (LSER)²⁰ involving the change of standard free energy related to the process $5\beta \rightleftharpoons 5\alpha$ as a function of the used media (i.e. ${}^x\Delta G^\circ_{\alpha-\beta}$ values, with x denoting the solvent). The ${}^x\Delta G^\circ_{\alpha-\beta}$ amounts (and so, also the relevant epimerization equilibrium constants, ${}^xK_{\alpha/\beta}$) have been determined within a wide number of solvents, pure and in mixture, characterized by very different proticity and polarity, the permittivity ranging from 80.2 to 4.8. In this context, a new descriptor of solvent hydrogen-bond donor (HBD) capacity, named S_N^{OH} , has been defined and used in substitution of the widely employed α -scale of solvent HBD acidities,²¹ which demonstrated modest sensitivity in monitoring the acid properties of the alcoholic solvents used in this study. The rationalization of the whole obtained data has also been supported by relevant molecular modeling calculations. As a further and final goal, a conclusive answer to the limited knowledge about the stereo-structure of **5** was also provided, basing the characterization of 5β and 5α by means of NMR measurements supported by theoretical data coming from DFT calculations.

Results and discussion

Preparation and characterization of MKA. In our initial studies, dihydroartemisinin **2** was submitted to pyrolysis (120 °C for 1h) in order to obtain a series of by-products reported in literature, such as **DKA 4** and **MKA 5**. **MKA** formation under thermal conditions is likely to proceed via unzipping triggered by protonation of the peroxide and re-closure by reaction of the pendant hydroperoxide with the aldehyde just formed.^{14a} The proton sources for the thermal rearrangement can presumably be formic acid produced as the thermolysis proceeds. The purification of the reaction mixture, however, proved rather tricky and **MKA** could not be obtained in high yields (13%, even after repeated semi-preparative HPLC experiments). At this point, a new preparation process was devised: **5** was obtained directly from **2** by adding a catalytic amount of NaOH to a solution of **2** in a H₂O:MeCN (30:70; v/v) mixture, at room temperature. Purification of **5** was, in this case, quite easily obtained through preparative normal-phase HPLC (NP-HPLC) by collecting the relevant epimers in a single cumulative peak, with a 72 % yield and with a higher degree of purity than the

one obtained following pyrolysis of **2**. Once obtained in an almost pure state (95-99% purity), **5** was submitted to reversed-phase HPLC (RP-HPLC) on a Symmetry C18 column using a mobile phase made of 10 mM NaH₂PO₄/MeCN/MeOH (55/35/10; v/v) with a ${}^w\text{pH} = 4.76$ and a ${}^s\text{pH} = 5.61$. The preliminary chromatographic runs were performed at room temperature ($T = 25^\circ\text{C}$). A typical chromatogram obtained is reported in Figure 3 (chromatogram on the left), where the UV trace was registered at 214 nm. As it can be observed, two well resolved chromatographic peaks, corresponding to the two epimers, were obtained, with satisfactory retention ($k_1 = 5.73$, $k_2 = 8.13$), selectivity ($\alpha = 1.42$) and resolution ($R_s = 11.9$). However, a visible plateau could be observed as the temperature was raised, as can be seen in the right chromatogram (dynamic chromatogram) of Figure 3 ($T = 50^\circ\text{C}$).

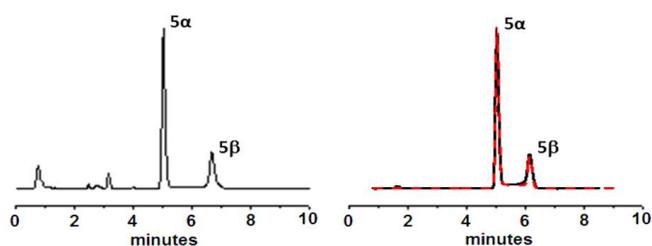


Figure 3. Chromatogram of purified **MKA 5**. Column: Symmetry C18, 3.5 μm (100 mm \times 4.6 mm I.D.). Eluent: 10mM NaH₂PO₄/MeCN/MeOH (55/35/10; v/v) with a ${}^w\text{pH} = 4.76$ and a ${}^s\text{pH} = 5.61$. Flow: 1.00 mL/min. UV detection at 214 nm. Left: $T = 25^\circ\text{C}$. Right: $T = 50^\circ\text{C}$ (red trace refers to the simulated chromatogram).

This behavior closely reminds that of **2**, on which an in-depth kinetic study of the related epimerization process has been reported.¹⁰ The above dynamic chromatogram has been successfully simulated by the stochastic model (for details see Experimental Section) and the obtained results indicated that, under the same operative conditions of pH and temperature, the $5\beta \rightleftharpoons 5\alpha$ interconversion is about 2.5 times slower than that involving the epimers of **2**. The determined relevant rate constants are $k_{\beta \rightarrow \alpha} = 9.2 \times 10^{-4} \text{ s}^{-1}$ and $k_{\alpha \rightarrow \beta} = 3.9 \times 10^{-4} \text{ s}^{-1}$, corresponding to free energy activation barriers of 23.5 and 24.0 kcal mol⁻¹, respectively (for the α / β assignments, see below). At this point, a stereochemical characterization of compound **5** became fundamental in order to assign the α - and β - configuration to the epimers. ¹H-NMR and ¹³C-NMR, with special reference to NOE experiments, were used to the purpose. The comparison of the physico-chemical and spectroscopic properties of **MKA** with data present in literature, allowed its identification as a previously described product, obtained by exposure of the ethyl ether analogue of **2** (β -arteether) to simulated stomach acid,¹⁶ and, more recently, by heating solid **2** at 100 °C for 14 h under nitrogen.^{14a} The relative configuration of the two epimers at C-3 was determined by means of NOE spectra, performed in CD₃CN (for NMR assignments see Electronic Supporting Information, ESI†). These outcomes, integrated with J -coupling analysis of

the hemiacetalic hydrogen in **5 α** and **5 β** and relevant DFT calculations,²² allowed a reliable assignment of the (*R*)-configuration (epimer **5 α** , hydroxyl group in the axial position) to the C-3 center of **MKA** in the major isomer, and the (*S*)-configuration at the same carbon atom (epimer **5 β** , hydroxyl group in the equatorial position) to the minor epimer (for a detailed description see the relevant section reported in ESI†). Experimental determinations of equilibrium constants of epimerization, ${}^xK_{\alpha\beta}$, in several media, *x*, (see about next subsection) also indicated that the energies of the two epimers are quite similar, with their differences ranging from 0.9 to 0.3 kcal/mol as a function of the considered solvent (in all cases **5 β** is the less stable of the two epimers), therefore accounting for the ratio observed in solution. A full characterization of **5** is reported in ESI†, including elemental analysis, optical rotation data, NMR and mass spectra.

Determination of ${}^xK_{\alpha\beta}$ values as a function of change of solvent *x*.

As already stressed in the introduction, at present time no information is available in literature on the thermodynamics of the equilibrium **5 β** ⇌ **5 α** . To shed light on this matter, we planned an in-depth analysis of the differential energy stability of the pair of stereoisomers **5 β** and **5 α** as a function of the solvent. Consequently a large number of solvents (8 pure and 10 as binary mixtures, see Table 1), displaying wide spanning values of permittivity ϵ and proticity α -scale (with the ϵ and α -scale varying from 4.8 to 80.2 and from 0.00 to 1.17, respectively), have been selected to secure the establishment of the widest typology of solute-solvent interactions. Accordingly, the equilibrium constants of the **5 β** ⇌ **5 α** isomerization within the 18 above solvents, (i.e. the quantities ${}^xK_{\alpha\beta}$, with *x* denoting the solvent), have been determined by either ¹H-NMR or cryo-HPLC-UV measurements (details are given in Experimental Section) and then collected in Table 1.

LSER investigations.

Starting from such a variety of thermodynamic data, it becomes possible to perform suitable LSER analyses, useful to distinguish the several contributions of solute-solvent interactions responsible for the changes of relative amounts of the epimers **5 β** and **5 α** in response to a change of solvent. Accordingly, the following seven descriptors of solvent effects have been taken into consideration: the permittivity (ϵ) and cohesive pressure (δ^2)²³ of the solvent (suitable to take into account energetic effects due to shielding of charges and generation of solvent cavities in which host the molecules of the epimers, respectively), the *solvent-to-solute* hydrogen-bond donor (α -scale,²¹ hereafter symbolized as α_s , the acceptor number AN²⁴ and the new parameter S_N^{OH} defined in this work for the first time) and acceptor (β -scale,²¹ hereafter symbolized as β_s) parameters and, finally, the solvatochromic descriptor $E_T(30)$,²⁵ able to monitor, although in a poorly specialized way, dispersive, electrostatic and acid-base interactions.

Table 1. Experimental data of ${}^xK_{\alpha\beta}$, and related $\Delta G^\circ_{\alpha\beta}$ values, as a function of change of solvent and descriptors of solvent effect employed in LSER analyses.

	$K_{\alpha\beta}$	$\Delta G^\circ_{\alpha\beta}$	δ^2	ϵ	$E_T(30)$	β_s	α_s	AN	S_N^{OH}
H ₂ O	4.41	-0.879	552.3	80.2	63.1	0.47	1.17	0.548	1.00
10/90 MeOH/H ₂ O	4.35	-0.87	509.8	70.70	62.2	0.489	1.151	0.535	0.92
10/90 ACN/H ₂ O	4.06	-0.83	499.2	69.39	61.5	0.463	1.072	0.512	0.88
30/70 MeOH/H ₂ O	3.70	-0.78	429.9	56.86	60.0	0.527	1.113	0.508	0.72
30/70 ACN/H ₂ O	3.08	-0.67	401.0	56.01	58.4	0.449	0.876	0.440	0.72
50/50 ACN/H ₂ O	2.70	-0.59	313.6	48.04	57.1	0.435	0.680	0.369	0.60
50/50 MeOH/H ₂ O	2.98	-0.65	356.8	47.27	58.3	0.565	1.075	0.482	0.60
DMSO	1.60	-0.28	167.2	46.7	45.1	0.76	0.00	0.193	0.00
70/30 ACN/H ₂ O	2.15	-0.45	237.0	42.76	56.8	0.421	0.484	0.297	0.51
70/30 MeOH/H ₂ O	2.91	-0.63	290.6	40.23	57.1	0.603	1.037	0.455	0.47
90/10 ACN/H ₂ O	2.08	-0.43	171.0	39.00	54.6	0.407	0.288	0.225	0.31
ACN	2.07	-0.43	142.1	37.5	45.6	0.40	0.19	0.189	0.00
90/10 MeOH/H ₂ O	2.05	-0.43	231.1	34.85	56.1	0.641	0.999	0.428	0.39
MeOH	1.98	-0.40	203.9	32.6	55.4	0.66	0.98	0.415	0.35
EtOH	1.76	-0.33	166.9	24.6	51.9	0.75	0.86	0.379	0.26
CH ₂ Cl ₂	2.77	-0.60	98.6	8.9	40.7	0.10	0.13	0.204	0.00
THF	2.17	-0.459	90.6	7.6	37.4	0.55	0.00	0.080	0.00
CHCl ₃	3.35	-0.72	84.8	4.8	39.1	0.10	0.20	0.231	0.00

Inclusion of these descriptors within proper multiparameter equations of the type reported below (equation 1) allows the performance of linear regression analyses which supply the required information on the differential effect exercised by the solvent on the epimers **5 β** and **5 α** :

$${}^x\Delta G^\circ_{\alpha-\beta} = {}^{gas}\Delta G^\circ_{\alpha-\beta} + \sum (f_i D_i)^x = {}^{gas}\Delta G^\circ_{\alpha-\beta} + {}^x\Delta G^{\circ S}_{\alpha-\beta} \quad (1)$$

In equation 1, ${}^x\Delta G^\circ_{\alpha-\beta}$ is the change of standard free energy (ΔG°) related to the **5 β** ⇌ **5 α** epimerization in solvent *x*, ${}^{gas}\Delta G^\circ_{\alpha-\beta}$ is the constant obtained by regression, expressing ΔG° in absence of solvent and, finally, D_i represents the generic *i* descriptor of solute-solvent interactions (which assumes a specific value in each solvent *x*), f_i being its regression coefficient. Therefore, the sum $\sum (f_i D_i)^x$ corresponds to the whole contribution given by solvent *x* to the quantity ${}^x\Delta G^\circ_{\alpha-\beta}$, which hereafter will be denoted by the term ${}^x\Delta G^{\circ S}_{\alpha-\beta}$. Thus, when the equation used for the regression can effectively reproduce the experimental ${}^x\Delta G^\circ_{\alpha-\beta}$ data, the value assumed by each coefficient f_i will give a quantitative estimation of the role played by the specific solute-solvent interaction monitored by the related descriptor D_i . The detailed procedure followed in performing the LSER study has been extensively reported in ESI†.

Definition of S_N^{OH} , a new descriptor of hydrogen-bond donor capacity.

In the course of the LSER investigation it has been pointed out that, although the role played by the *solvent-to-solute* hydrogen-bond donor capacity is important in determining the values assumed by the ${}^xK_{\alpha\beta}$ quantities (see LSER findings collected in Table S1 of ESI†), neither α_s nor AN have proved to act as fully satisfactory descriptors for this kind of interaction, as they displayed a quite modest sensitivity in discriminating highly protic solvents (i.e. water and the alcoholic solvents, pure or in mixtures, used in the present study, see bars *a* and *c* in Figure 4, and relevant section reported in ESI†). To overcome this limit, we resorted to the definition of the new descriptor S_N^{OH} , based on the UV-visible absorption properties of the organic salt formed by the acid-base reaction between 2-nitrocyclohexanone (NCE), a very acidic ketone ($pK_a=6.0^{26a,b}$), and 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU), which hereafter will be denoted through the acronym NCEDBU. The design of such molecular probe was suggested by previous studies^{26c-g} in which some of us highlighted that NCE may act as an effective probe of aspecific as well as lone pair donor-acceptor and hydrogen bonding solute-solvent interactions. According to these properties, we have been driven to use the ionized form of NCE (i.e. its enolate anion, NCE^-) as a possible probe of specific HBD interactions, while DBU was selected as the base suitable to deprotonate NCE, in order to afford the large and apolar cation $DBUH^+$ as the counter-ion of NCE^- . Consistently with this choice, the resulting structure of the obtained salt has proven to possess a good solubility within the widest typology of solvents (from the most to the less polar ones). Moreover, the UV-visible spectral properties of NCEDBU evidenced high sensitivity in discriminating polarity and proticity of very different solvents, pure or in mixture (Figure S2 and Table S2 in ESI†), also if these are endowed with high and quite similar values of permittivity, as indeed is the case of the media investigated in this study (ESI† contains an extensive description of the adopted procedure with which the new descriptor of hydrogen-bond donor capacity S_N^{OH} has been successfully defined). As it may be easily seen in the graph of Figure 4 and Table 1, compared to α_s and AN, S_N^{OH} is much more able to discriminate the protic capabilities of solvents like water and alcohols, especially in the case of water/alcohol mixtures. Therefore, regressions based on four-descriptors equations were repeated by substituting α_s or AN with S_N^{OH} . As expected, in this case all correlations run with better values of R^2 . The best one of this set, which is also the best in absolute sense (entry 41, $R^2 = 0.9624$ in Table S1 of ESI†), has δ^2 , β_s , ϵ and S_N^{OH} as the four descriptors of solvent effect (equations 2a-b), all resulting widely significant at statistical level:

$${}^x\Delta G^\circ_{\alpha-\beta} = {}^{gas}\Delta G^\circ_{\alpha-\beta} + f_1 \times \delta^2 + f_2 \times \beta_s + f_3 \times \epsilon + f_4 \times S_N^{OH} \quad (2a)$$

$${}^x\Delta G^\circ_{\alpha-\beta} = -5.05 \times 10^{-1} - 2.72 \times 10^{-3} \times \delta^2 + 4.52 \times 10^{-1} \times \beta_s + 7.20 \times 10^{-3} \times \epsilon + 3.54 \times 10^{-1} \times S_N^{OH} \quad (2b)$$

A very important aspect related to this ensemble of descriptors is that they are all highly specialized in monitoring different but complementary kind of solute-solvent interactions. Therefore, the dissection of single contributions to ${}^x\Delta G^{\circ S}_{\alpha-\beta}$ (and then also to ${}^x\Delta G^\circ_{\alpha-\beta}$) becomes in this way much easier to be performed.

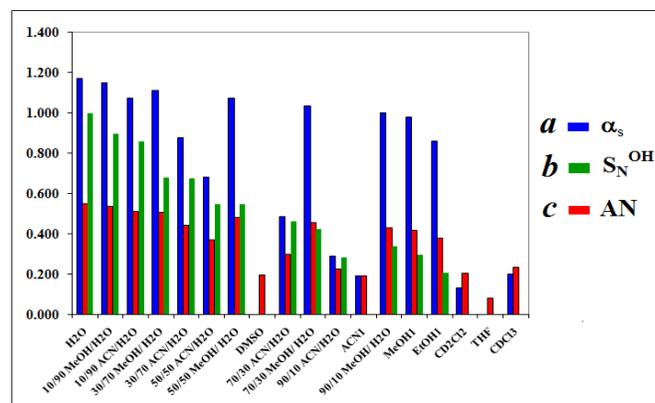


Figure 4. Trend of the values assumed by α_s (blue bars), AN (red bars) and S_N^{OH} (green bars) in the 18 solvents used in the LSER analyses

Rationalization of ${}^xK_{\alpha\beta}$ data through findings coming from LSER analysis.

From the LSER analysis based on equation (2b) we can draw several important considerations. The first one is that, globally, the solvent plays a decisive role in determining the final amount of ${}^x\Delta G^\circ_{\alpha-\beta}$ through the quantity ${}^x\Delta G^{\circ S}_{\alpha-\beta}$. For 15 out of the 18 considered media (see Table 1), the contribution afforded to ${}^x\Delta G^\circ_{\alpha-\beta}$ by ${}^x\Delta G^{\circ S}_{\alpha-\beta}$ is in fact greater than 60% (Figure S5 in ESI†). The second remark is that only one among the four monitored solute-solvent interactions (that one taken into account by δ^2) acts in favor of the epimer **5a** (see equation 2b and Figure S5 in ESI†). The parameter δ^2 is in fact the only one with negative coefficient and so also the only one contributing to a reduction of ${}^x\Delta G^\circ_{\alpha-\beta}$. This indication could be interpreted by admitting that, because of the different stereo-structure, the molar volume of **5a** is smaller than that of **5b**, so that, compared to **5b**, **5a** would require less energy to be hosted in a cavity of the solvent of suitable volume. On the contrary, the positive coefficients related to the descriptors ϵ , β_s and S_N^{OH} point to the establishment of aspecific and specific (HBA and HBD) electrostatic forces all playing in favor of the **5b** epimer. More in particular, it may be seen (Figure 5) that, on the whole, this energy input of electrostatic nature is overcome by the one favorable to **5a** when the epimerization equilibrium is established in solvents either strongly apolar (as in the case of chloroform and dichloromethane, which cannot be involved into the establishment of H-bonds of appreciable strength as acceptor or donating partners) or strongly polar and protic, endowed with a permittivity greater than 47 (in these solvents, indeed, the cohesive pressure reaches very high values, thus providing a particularly great advantage in hosting **5a** in a solvent cavity compared to its bulkier epimer **5b**).

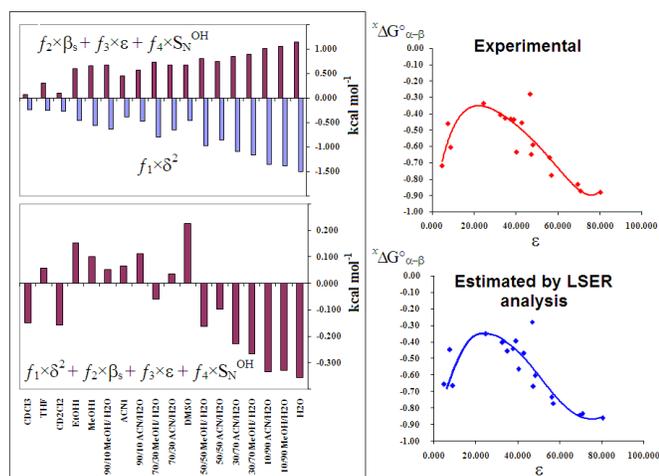


Figure 5. Additive contributions of the solute-solvent interactions monitored by the relevant descriptors to the quantity ${}^x\Delta G_{\alpha-\beta}^{\text{S}}$ as a function of solvent change. Top left graph: trend of the solute-solvent interactions favorable to the epimer **5 β** in the positive range of energies and to the epimer **5 α** in the negative range. Bottom left graph: trend of the overall solvent contribution to ${}^x\Delta G_{\alpha-\beta}^{\text{S}}$. Right graphs: experimental and calculated trend of ${}^x\Delta G_{\alpha-\beta}^{\text{S}}$ as a function of solvent permittivity.

Thus, on the whole, for eight out of the eighteen considered solvents a positive ${}^x\Delta G_{\alpha-\beta}^{\text{S}}$ was calculated, so that in these media the **5 β** epimer would be dominant, in absence of the prevailing and negative contribution of ${}^{\text{gas}}\Delta G_{\alpha-\beta}^{\text{O}}$. Interestingly, the trend of the computed ${}^x\Delta G_{\alpha-\beta}^{\text{S}}$ (and therefore also by the related ${}^x\Delta G_{\alpha-\beta}^{\text{O}}$) on increasing of permittivity reproduces faithfully the experimental trend, as it is visible in Figure 5. A more detailed analysis also put in evidence that HBA interactions progressively reduce their importance when going from low to high permittivity values, while the opposite is observed for the HBD interactions (Figure S6). On the whole, however, passing from solvents of low permittivity to solvents with high permittivity, the combined action of all type of H-bonds leads to a slowly decreasing, positive contribution to the ${}^x\Delta G_{\alpha-\beta}^{\text{S}}$ amount. In summary, from a purely quantitative point of view, the findings coming from the LSER analysis could be rationalized by affirming that the cavitation effects are those always playing the most important role, with an average contribution to ${}^x\Delta G_{\alpha-\beta}^{\text{S}}$ amounting to 53% (in no one of the considered solvents it becomes smaller than 40%). Next, these are followed by the effects deriving from the establishment of aspecific electrostatic forces modulated by the solvent's permittivity (an average value of 21%), and finally by contributions arising from HBA and HBD interactions (average values of 18% and 9%, respectively). Therefore, it seems plausible that the achieved equilibrium **5 β** \rightleftharpoons **5 α** is essentially driven by the differential molar volume, while only in a more fragmented way, although significant, by dipolarity, acidity and lone pair acceptor capacity displayed by the couple of stereoisomers **5 β** and **5 α** .

Rationalization of the LSER findings by molecular modeling calculations.

With the aim of gaining an independent support to these reasonable and interesting conclusions we resorted to relevant theoretical calculations. Structures of **5 β** and **5 α** have been modeled by suitable integration of standard procedures based on molecular mechanics (MM), Hartree-Fock (HF) and Density Functional Theory (DFT) calculations (for details see Experimental Section). For each of the epimers **5 α** and **5 β** we obtained two main conformations, named **5 α_1** , **5 α_2** , **5 β_1** and **5 β_2** , respectively (Figure S7). In order to estimate the differential acidity of the couple of epimers of **5**, the same DFT calculations were also extended to the conjugated bases derived by the above four conformations, hereafter denoted as **5 α_1^-** , **5 α_2^-** , **5 β_1^-** and **5 β_2^-** , respectively. All the computed energy stability data have been collected in Tables S3 (absolute data) and S5 (relative data) of ESI†. Starting from the above geometries, in addition to the aforementioned acidities, we also assessed the molar volumes and the dipolarities of both **5 α** and **5 β** (see Table S5 in ESI†). Inspection of all these findings afforded a clear rationalization of the results achieved by the LSER analysis:

- the estimation of the molar volumes indicated for **5 β** a slightly larger value ($\Delta V_{\alpha-\beta} = -0.0066 \text{ cm}^3 \text{ mol}^{-1}$), which can logically account for the negative sign and magnitude that the f_i coefficient of δ^2 achieved by regression (f_{δ^2} , equation 2b). This latter should just correspond to the differential molar volume existing between **5 α** and **5 β** and, indeed, its value ($f_{\delta^2} = \Delta V_{\alpha-\beta} = -0.0027 \text{ cm}^3 \text{ mol}^{-1}$) is in good agreement with the calculated one;
- the results of calculations collected in Table S5 indicate that in all the pure organic solvents the computed average dipolarity is greater for **5 α** than for **5 β** (a finding that agrees with the elution order found for the epimers in the performed RP-HPLC separations reported in Figure 3, in which the mobile phase has properties intermediate between those of water and acetonitrile in their pure state, suggesting that, also in the limit composition H₂O/MeCN 65/35 v/v, the assessed dipolarity is in favor of the α epimer: $5\alpha_{\text{Dipole}} = 4.49$ debye; $5\beta_{\text{Dipole}} = 4.37$ debye), but also that, upon increasing of solvent permittivity, the dipolarity of **5 β** grow faster than that of **5 α** , up to become a bit larger in pure water. This is in fact what is pointed out by the amounts that the difference $\Delta\text{Dipole}_{5\alpha-5\beta} = 5\alpha_{\text{Dipole}} - 5\beta_{\text{Dipole}}$ progressively assumes as a consequence of following change of solvents: chloroform \rightarrow dichloromethane \rightarrow acetonitrile \rightarrow water (0.77, 0.64, 0.40 and -0.03 being the respective values of $\Delta\text{Dipole}_{5\alpha-5\beta}$). Such a behavior is consistent with the positive sign found by LSER analysis for the f_i coefficient related to the descriptor ϵ (see equation 2b);
- assessment of the differential acidity expressed by the couple of epimers **5 α** and **5 β** (i.e. the $\Delta pK_i^{\alpha-\beta}$ quantity) was based on the thermodynamic cycle reported in Scheme 1. From its inspection it may be seen that ionization equilibria of both **5 α** and **5 β** (K_i^{α} and K_i^{β} being the respective ionization constants) as well as the epimerization equilibria either between **5 α** and **5 β** or

their conjugated bases $5\alpha^-$ and $5\beta^-$ (whose equilibrium constants were expressed through the symbols $K_{\alpha/\beta}$ and $K_{\alpha/\beta}^-$, respectively) have been taken into consideration. The equilibrium constants $K_{\alpha/\beta}$ and $K_{\alpha/\beta}^-$ were then assessed in four solvents endowed with very different permittivity values (water, $\epsilon = 80.2$; acetonitrile, $\epsilon = 37.5$; dichloromethane, $\epsilon = 8.9$; chloroform, $\epsilon = 4.8$) by resorting to the aforementioned molecular modeling calculations and, by them, the required quantity $\Delta pK_i^{\alpha-\beta}$ has been eventually estimated according to equation 3:¹⁰

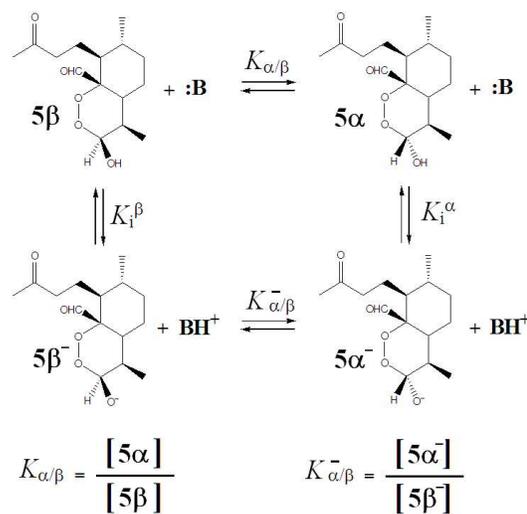
$$\Delta pK_i^{\alpha-\beta} = \log(K_{\alpha/\beta}) - \log(K_{\alpha/\beta}^-) \quad (3)$$

All the obtained values indicated a significantly greater acidity of 5β (${}^W\Delta pK_i^{\alpha-\beta} = 2.07$, ${}^{ACN}\Delta pK_i^{\alpha-\beta} = 2.25$, ${}^{DCM}\Delta pK_i^{\alpha-\beta} = 2.12$, ${}^{Chl}\Delta pK_i^{\alpha-\beta} = 2.25$, the superscripts W, ACN, DCM and Chl referring to water, acetonitrile, dichloromethane and chloroform, respectively), suggesting that this epimer may give rise to more effective solvent-HBA interactions than 5α can, irrespective of the solvent in which such a couple of stereoisomers is solubilized. This is exactly what may also be inferred by the positive sign that the LSER analysis afforded for the f_i coefficient related to the descriptor β_s . Assessment of the differential acidity was also repeated at a higher level of theory (GGA-BLYP/QZ4P, large core) in water, in order to exclude any possible dependence of the result by the employed basis set. This is indeed what we found, as the obtained new value (${}^W\Delta pK_i^{\alpha-\beta} = 2.03$) was perfectly consistent with the one achieved in the same solvent at the lower level of theory (${}^W\Delta pK_i^{\alpha-\beta} = 2.07$). Interestingly, an indirect indication of the reliability of the performed assessments is provided by the good agreement found between experimental and calculated values of $K_{\alpha/\beta}$ constants (i.e. ${}^{exp}K_{\alpha/\beta}$ and ${}^{cal}K_{\alpha/\beta}$, respectively), whose ratios $R = {}^{exp}K_{\alpha/\beta}/{}^{cal}K_{\alpha/\beta}$, or their inverse $1/R$, were always smaller than 2 (${}^{Chl}R=1.2$, ${}^{DCM}R=1.2$, ${}^{ACN}R=0.7$, ${}^WR=1.8$ from the lowest level and ${}^WR=1.9$ from the higher level of DFT calculations). This corresponds to the maximum and minimum absolute errors of 0.34 (water) and 0.09 (dichloromethane and chloroform) kcal mol⁻¹ existing between experimental and calculated $\Delta G_{\alpha/\beta}^\circ$.

Conclusions

A more convenient synthetic procedure has been proposed to prepare **MKA** (**5**) from the **DHA** precursor in higher yields. For the first time the two epimeric forms 5α and 5β of **MKA** have been successfully resolved by a RP-HPLC technique, also allowing the determination of the activation barriers of the forward and backward epimerization process at 50°C by dynamic-chromatography approach. The activation barriers were found quite lower than those measured for the epimers of **DHA** in the same conditions of pH and temperature. The stereochemical structures of 5α and 5β have been completely elucidated through a comprehensive NMR study, also supported by relevant DFT calculations. In-depth information about the factors influencing the thermodynamic epimerization

constant related to the equilibrium $5\beta \rightleftharpoons 5\alpha$ as a function of the solubilizing media has been achieved resorting to extensive LSER analyses.



Scheme 1. Thermodynamic cycle expressing the interrelations existing between the ionization and epimerization equilibria of 5α and 5β .

As an indirect but really important result of this study, a new descriptor of solvent hydrogen-bond donor capacity named S_N^{OH} has been defined, allowing a more sensible differentiation of the HBD properties exhibited by aqueous and/or alcoholic solvent mixtures (that is, by strongly protic media) when compared to that achievable using the widely known and employed α_s and AN descriptors. Eventually, a rationalization of the factors modulating the position assumed by the equilibrium $5\beta \rightleftharpoons 5\alpha$ has been drawn by interpreting the LSER findings on the base of indications suggested by molecular modeling calculations. The identified factors, whose effects have also been singularly quantified, are the differential molar volume, dipolarity, acidity and lone pair acceptor capacity displayed by the couple of stereoisomers 5β and 5α . In view of a further and logical extension of this study we are currently engaged in an in-depth investigation of the mechanistic aspects involved in the epimerization process $5\beta \rightleftharpoons 5\alpha$, including an extensive determination of the forward and backward rate constants governing the interconversion under different conditions of pH and temperature. We believe that a proper knowledge of the combined thermodynamic and kinetic features of the $5\beta \rightleftharpoons 5\alpha$ equilibration might proactively contribute in promoting the rational development of new robust analytical methods for quantitation of **DHA**, considering the demonstrated potential formation of **MKA** during the stages of production and storage of its parent drug. In addition, the elucidated thermodynamic aspects of the $5\beta \rightleftharpoons 5\alpha$ process could also supply information on the dependence that the not negligible *in vivo* biological activity of 5α and 5β in mixture can manifest as a function of the actual composition in different biological compartments.

Acknowledgements

We thank financial support from PRIN 2010, grant no. 2010N3T9M4_008.

Experimental

Apparatus. Analytical liquid chromatography was performed using an HPLC separation module coupled with a Photodiode Array Detector. Low temperature variable temperature HPLC (DHPLC) was performed by using a lab-made thermally insulated container cooled by the expansion of liquid carbon dioxide (CO₂). Flow of liquid CO₂ and column temperature were regulated by a solenoid valve, thermocouple, and electric controller. A column oven was used for variable temperature HPLC (DHPLC) above 25°C. Temperature variations after thermal equilibration were within ± 0.2 °C. Semi-preparative LC separations were carried out on a preparative chromatographic system, equipped with a 5 mL loop injector, and a differential refractometer (RI) detector. All DHPLC experiments were performed on a C18 Symmetry, 3.5 µm (75 mm × 4.6 mm I.D.) column, with different binary and ternary hydro-organic mobile phases delivered at 1.0 mL/min and UV detection at 214 nm. ESI-MS and high resolution mass spectra were acquired on an Orbitrap mass spectrometer; samples were made up in acetonitrile.

Dihydroartemisinin (DHA, 2) samples were supplied by sigma-tau S.p.A., Pomezia (Italy). Monoketo-aldehyde (**MKA, 5**) was obtained in two different ways: by thermal degradation of **2** or by basic rearrangement of **2**.

Thermal Degradation: A screw-cap vial containing crystalline **2** (1.0 g; 3.5 mmol) was placed in an aluminum block pre-heated to 120 °C and left for 1h. Periodic withdrawals of the reaction mixture were made for the HPLC monitoring, under the following normal-phase conditions: column Reprosil Si 120, 3 µm (150 mm × 4.6 mm I.D.), *n*-hexane-ethanol 95:5 (v/v), flow-rate 1.0 mL/min, and UV detection at 214 nm. An evaporative light scattering detection (ELSD) system was also used (P = 45 psi; T = 45 °C; spray gun scale: 70%; gain = 9). The brown pyrolysate obtained (0.82 g) was allowed to cool to room temperature and dissolved in 10 mL of *n*-hexane-ethanol-dichloromethane 72:8:20 (v/v/v). Afterwards, it was submitted to semi-preparative HPLC, under the following normal-phase conditions: *n*-hexane-ethanol 95:5 (v/v), flow-rate 20.0 mL/min, and UV detection at 220 nm. A further purification of the last eluting fraction (0.11 g) gave **MKA (5)** as one major product (13%, 99% purity). An axially compressed column (500 mm × 20 mm I.D.) packed with LiChrosorb Si 100, 10 µm silica gel was used for the purification of the brown pyrolysate.

Rearrangement of DHA under basic conditions: 0.351 g of pure crystallized **DHA** were added to 42 mL of a MeCN/H₂O (70/30) mixture at room temperature. A catalytic amount of NaOH was then added (concentration in the final solution = 2 mM) and the reaction was complete after 1.5 h. Reversed-phase HPLC was used to monitor the reaction. RP-HPLC was performed on a C18 Symmetry, 3.5 µm (100 mm × 4.6 mm

I.D.) column, with a mobile phase composed of water-acetonitrile-methanol 55:35:10 (v/v/v) buffered using 10 mM sodium phosphate ($\epsilon_{\text{H}_2\text{O}} = 4.76$; $\epsilon_{\text{MeCN}} = 5.61$) delivered at 1.0 mL/min, and UV detection at 214 nm (T = 5°C). The mixture was diluted with water, extracted with CH₂Cl₂ (3 × 60 ml), dried over anhydrous Na₂SO₄, filtered, and concentrated *in vacuo*. We obtained 0.322 g of a pale yellow oil. Afterwards, 0.245 g of the crude oil obtained were dissolved in a mixture of 4.75 mL of *n*-hexane, 0.25 mL of EtOH and 0.4 mL of dichloromethane (concentration = 45 mg/mL) and were purified by NP preparative chromatography, obtaining 0.18 g of pure **MKA**. Yield = 72%; 99 % purity.

High resolution and accuracy mass spectrum: ESI-MS (pos.): *m/z* found 307.1514 ([*M* + Na]⁺), C₁₅H₂₄O₅Na requires 307.1516 (monoisotopic), Δ*m* = 0.7 ppm. [α]_D²⁰ = -166 (c = 2.0; CH₃CN). Anal. Calcd (%) for C₁₅H₂₄O₅: C, 63.36; H, 8.51. Found: C, 63.51; H, 8.59. For the fully spectroscopic characterization of **5**, see the ESI†.

Synthesis of NCEDBU.

To a solution of 2-nitrocyclohexanone (0.150 g, 1.05 mmol) in diethyl ether (2 mL) at room temperature and under magnetic stirring a solution of 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU, 0.170 g, 1.11 mmol) in diethyl ether (2mL) was added dropwise. A pale yellow precipitate instantly appeared and the suspension was stirred for few minutes. Then, the solid was filtered and rinsed four times with Et₂O (20 ml). The product was dried under vacuum affording 231 mg of a pale yellow solid (74.5% yield).

¹H NMR (400MHz, CDCl₃) δ 3.44-3.39 (m, 6H), 2.81-2.69 (m, 2H), 2.73 (t, J = 5.9 Hz, 2H), 2.25 (t, J = 5.9 Hz, 2H), 2.01-1.95 (m, 2H), 1.71-1.64 (m, 10H); ¹³C NMR (101 MHz, CDCl₃) δ 186.6, 165.9, 120.9, 54.2, 48.6, 39.5, 38.3, 32.4, 29.1, 28.2, 26.9, 24.0, 23.5, 23.0, 19.7; FT-IR (KBr cm⁻¹) 3300 (NH), 1644 (C=N), 1616 (C=C), 1550 (NO₂); ESI-MS (+) *m/z* 153.1, ESI-MS (-) *m/z* 141.8. All relevant spectra are supplied in ESI†.

NMR Measurements and related Calculations. NMR spectra were recorded in CD₃CN at T = 25 °C on a 600 MHz spectrometer equipped with a triple resonance indirect probe. The standard pulse sequence and phase cycling were used for gradient-enhanced COSY,²⁷ HSQC²⁸ and HMBC²⁹ spectra. NOE spectra were acquired in CD₃CN at 25°C by means of the DPFGE-NOE sequence³⁰, using a mixing time of 2.0 s and two “rsnob” 50 Hz-wide selective pulses. Geometry optimization were carried out at the B3LYP/6-31G(d) level by means of the Gaussian 03 series of programs.²² The reported energy values are not ZPE corrected. Harmonic vibrational frequencies were calculated for all the stationary points. For each optimized ground state the frequency analysis showed the absence of imaginary frequencies. *J*-coupling calculations were obtained with the GIAO method at the B3LYP/6-311++G(2d,p)//B3LYP/6-31G(d) level, and by including the Fermi contact term.

Determination of S_N^{OH} descriptor by UV-Vis spectrometry.

The UV spectra of NCEDBU were recorded in different media (pure solvents and binary mixtures of fully miscible solvents) at 25°C and in a range of wavelengths from 500 to 200 nm. A general procedure for the preparation of all the solutions analyzed is the follow: 5 mg of NCEDBU were dissolved in 1ml of the appropriate solvent or binary mixture to obtain a 1.7×10^{-2} M stock solution. 100 μ L of the stock solution were transferred in a 5 mL graduated flask partially filled with the same media which was then gradually added to reach 5 mL volume. The final solution was 3.4×10^{-4} M. Because of the low solubility of the NCEDBU salt in hexane and cyclohexane, a stock solution of NCEDBU in CHCl_3 1.7×10^{-2} M was prepared and 25 μ L were transferred in hexane (3 mL) and cyclohexane (3 mL) giving rise to the formation of clear solutions having a concentration of 1.4×10^{-4} M. 3 mL of the final freshly made solutions were placed in a 3.5 mL capacity quartz cuvette and submitted to UV-Vis spectrometry analysis. An extensive description of the procedure adopted to define S_N^{OH} , inclusive of all the relevant UV data, has been reported in ESI† (subsection “Definition of the new descriptor S_N^{OH} ”, Figures S2-S3 and Table S2).

Determination of ET_{30} descriptor by UV-Vis spectrometry.

The ET_{30} descriptor was determined acquiring the UV spectra of Betaine 30 [2,6-Diphenyl-4-(2,4,6-triphenylpyridinio)-phenolate] in different media (pure solvents and binary mixtures of fully miscible solvents) at 25°C and in a range of wavelengths from 800 to 200 nm. The ET_{30} values were calculated by the equation: $ET_{30} (\text{kcal mol}^{-1}) = h c N / \lambda_{\text{max}} (\text{nm}) = 28591 / \lambda_{\text{max}} (\text{nm})$ where:

h = Planck constant, c = speed of light in vacuum, N = Avogadro constant and λ_{max} = wavelength of maximum absorbance of Betaine 30 acquired in a particular media.

A general procedure for the preparation of all the solutions analyzed is the followin: 10 mg of Betaine 30 were dissolved in 1 mL of the appropriate solvent or binary mixture to obtain a 1.8×10^{-2} M stock solution. 25 μ L of the stock solution were transferred in a 3 mL graduated flask partially filled with the same media which was gradually added to reach 3ml volume. The molarity of the final solution was 1.5×10^{-4} M. 3 mL of the freshly prepared samples were transferred to a 3.5 ml capacity quartz cuvette and submitted to UV-Vis spectrometry analysis.

Simulation of Dynamic Chromatograms. Simulation of the variable-temperature experimental chromatogram was performed by using the lab-made,³¹⁻³⁷ computer program Auto DHPLC y2k which implements both stochastic³⁸ and theoretical plate models³⁹ and may take into suitable account asymmetry of peaks. The program functionality was validated on several first-order isomerizations (both enantiomerization and non-enantiomerization) by comparing DHPLC results with those obtained by DNMR technique³¹⁻³⁵ or by classical method.³⁷ Simulation was carried out by the stochastic model, taking into account peak asymmetry and driving the optimization in automatic fashion by minimizing the RMS of

the differences existing between experimental and simulated chromatogram profiles.

LSER Analyses. Linear regression analyses, as well as F-test and t-test statistical evaluations were performed by the dedicated functions implemented in the Microsoft Office Excel 2003 program. Probability related to the t-Student distribution in the t-test was set to 0.05. Permittivity values of all solvent mixtures were calculated according to the Kirkwood’s theory.⁴⁰

Molecular Modeling Calculations. Geometries of $5\alpha_1$, $5\alpha_2$, $5\beta_1$ and $5\beta_2$ were modeled within three steps. First, a conformational search, performed in vacuum by molecular mechanics calculations based on the force field MMFF implemented in Spartan v. 1.2.0, was carried out on the structures of 5α and 5β . More in particular, in an energy window of 3 kcal mol⁻¹ the conformational search afforded 25 conformers for 5α and 19 conformers for 5β . Starting from these ones, on the base of geometric similitude criteria (RMS = 1.0) they were selected two main geometries for 5α and three for 5β , representative of 85% and 89% of their respective whole Boltzmann population (Bp). As the second step, such geometries were then optimized in vacuum at the ab initio HF/3-21G level of theory through the Spartan program, which provided two final conformations for both 5α (named $5\alpha_1$ and $5\alpha_2$, Figure S7a) and 5β (named $5\beta_1$ and $5\beta_2$, Figure S7b), as two of the three original conformers of 5β collapsed in a couple of very similar geometries. As the third and conclusive step, these latter four structures have been further optimized at the SCF level through DFT calculations run with the Amsterdam Density Functional (ADF) package v. 2007.01. (GGA-BLYP/QZ4P large core was the employed method) which were also integrated with the effect of water simulated through the COnductor like Screening MOdel (COSMO) procedure, with the cavity defined according to the Solvent Excluding Surface (SES) algorithm. Starting from the achieved structures of $5\alpha_1$, $5\alpha_2$, $5\beta_1$ and $5\beta_2$ they were also modeled, by deprotonation of the hemiacetal OH group, their conjugated bases $5\alpha_1^-$, $5\alpha_2^-$, $5\beta_1^-$ and $5\beta_2^-$, which were again optimized at the GGA-BLYP/QZ4P large core level of theory. Afterwards, in order to take theoretical information on the effect played by the solvent on the acid and dipolar properties of 5α and 5β , all the obtained geometries $5\alpha_1$, $5\alpha_2$, $5\beta_1$, $5\beta_2$, $5\alpha_1^-$, $5\alpha_2^-$, $5\beta_1^-$ and $5\beta_2^-$ were submitted to a new re-optimization, performed at a slightly lowest level of theory (the GGA-BLYP/DZP medium core being the employed method) by simulating the presence of water, acetonitrile, dichloromethane or chloroform through the COSMO procedure. The re-optimization at the same level of theory but in vacuum was also extended to the only neutral species $5\alpha_1$, $5\alpha_2$, $5\beta_1$ and $5\beta_2$. Molar volumes of the geometries optimized by the above DFT calculations were computed through the Spartan program. All energies and properties of the isomers were computed as values weighed on the Boltzmann populations assessed for the couple of conformations of each epimer. Values of $\Delta pK_1^{\alpha-\beta}$ were calculated taking also into

account the entropy contributions arising from the couple of conformations $5\alpha_1/5\alpha_2$ and $5\beta_1/5\beta_2$.

Notes and references

^a Dipartimento di Chimica e Tecnologie del Farmaco, Sapienza Università di Roma, P. le Aldo Moro 5, 00185 Roma, Italy. Fax: +3949912780; E-mail: marco.pierini@uniroma1.it; francesco.gasparrini@uniroma1.it.

^b Analytical Development, R&D Department, Sigma-Tau S.p.A., Via Pontina km 30.400, 00040 Pomezia, Italy.

^c Dipartimento di Chimica Industriale "T. Montanari", Università di Bologna, Viale Risorgimento 4, 40136 Bologna, Italy.

^d Kosterlitz Centre for Therapeutics, Institute of Medical Sciences, School of Medical Sciences, University of Aberdeen, Foresterhill, Aberdeen, AB25 2ZD, Scotland UK.

† Electronic Supplementary Information (ESI) available: Table and discussion of LSER correlations with relevant results. Definition of the new descriptor S_N^{OH} and relevant Table of experimental UV data. Tables collecting the absolute/relative stability energies of the calculated structures of 5α , 5β and of their conjugated bases, as well as the acid properties relevant to the epimers. ¹H and ¹³C NMR characterization of **MKA** (compound **5**), including COSY,HSQC and HMBC spectra. Spectral data relevant to **NCEDBU**. Cartesian coordinates of all the optimized structures discussed in the text.

See DOI: 10.1039/b000000x/

References

- WHO world malaria report. World Health Organization, 2010. http://www.who.int/malaria/world_malaria_report_2010/en/index.html
- R. W. Snow, C. A. Guerra, A.M. Noor, H. Y. Myint and S. I. Hay, *Nature*, 2005, **434**, 214.
- E. A. Ashley and N.J. White, *Curr. Opin. Infect. Dis.* 2005, **18**, 531.
- D. L. Klayman, *Science* 1985, **228**, 1049.
- D. Chaturvedi, A. Goswami, P. P. Saikia, N. C. Barua and P. G. Rao, *Chem. Soc. Rev.* 2010, **39**, 435.
- F. H. Jansen, *Malaria Journal*, 2010, **9**, 212.
- Y. Li, P.-L. Yu, Y. -X. Chen, L. -Q. Li, Y. -Z. Gai, D. -S. Wang and Y. -P. Zheng, *Chin. Sci. Bull.*, 1979, **24**, 667.
- (a) P. M. O'Neill and G.H. Posner, *J. Med. Chem.* 2004, **47**, 2945; (b) G. Bez, B. Kalita, P. Sarmah, N. C. Barua and D. K. Dutta, *Curr. Org. Chem.*, 2003, **7**, 1231; (c) G. H. Posner, I.-H. Paik, W. Chang, K. Borstnik, S. Sinishtaj, A. S. Rosenthal and T. A. Shapiro, *J. Med. Chem.*, 2007, **50**, 2516; (d) Q. G. Li, J. O. Peggins, L. L. Fleckenstein, K. Masonic, M. H. Heiffer and T. G. Brewer, *J. Pharm. Pharmacol.*, 1998, **50**, 173.
- W. Cabri, I. D'Acquarica, P. Simone, M. Di Iorio, M. Di Mattia, F. Gasparrini, F. Giorgi, A. Mazzanti, M. Pierini, M. Quaglia and C. Villani, *J. Org. Chem.*, 2011, **76**, 1751.
- (a) W. Cabri, I. D'Acquarica, P. Simone, M. Di Iorio, M. Di Mattia, F. Gasparrini, F. Giorgi, A. Mazzanti, M. Pierini, M. Quaglia and C. Villani, *J. Org. Chem.*, 2011, **76**, 4831; (b) I. D'Acquarica, F. Gasparrini, D. Kotoni, M. Pierini, C. Villani, W. Cabri, M. Di Mattia and F. Giorgi, *Molecules*, 2010, **15**, 1309; (c) W. Cabri, A. Ciogli, I. D'Acquarica, M. Di Mattia, B. Galletti, F. Gasparrini, F. Giorgi, S. Lalli, M. Pierini and P. Simone, *J. Chromatogr. B*, 2008, **875**, 180.
- M. Zeng, L. Li, S. Chen, C. Li, X. Liang, M. Chen and J. Clardy, *Tetrahedron*, 1983, **39**, 2941.
- A. J. Lin, D. L. Klayman, J. M. Hoch, J. V. Silverton and C. F. George, *J. Org. Chem.*, 1985, **50**, 4504.
- A. J. Lin, A. D. Theoharides and D. L. Klayman, *Tetrahedron*, 1986, **42**, 2181.
- (a) R. K. Haynes, H. -W. Chan, C. -M. Lung, N. -C. Ng, H. -N. Wong, L. Y. Shek, I. D. Williams, A. Cartwright and M. F. Gomes, *ChemMedChem*, 2007, **2**, 1448; (b) R. K. Haynes, H. -W. Chan, H. -N. Wong, K. -Y. Li, W. -K. Wu, K. -M. Fan, H. H. Y. Sung, I. D. Williams, D. Prosperi, S. Melato, P. Coghi and D. Monti, *ChemMedChem.*, 2010, **5**, 1282.
- L. Dhooghe, S. Van Miert, H. Jansen, A. Vlietinck and L. Pieters, *Pharmazie*, 2007, **62**, 900.
- J. K. Baker, J. D. McChesney and H. T. Chi, *Pharm. Res.*, 1993, **10**, 662.
- F. S. El-Ferally, A. Ayalp and M. A. Al-Yahya, *J. Nat. Products.*, 1990, **53**, 920.
- N. Acton and R. J. Roth, *Heterocycles* 1995, **41**, 95.
- D. L. Wesche, M. A. De Coster, F. C. Tortella and T. G. Brewer, *Antimicrob. Agents. Chemother.*, 1994, **38**, 1813.
- (a) M. J. Kamlet and R. W. Taft, *J. Chem. Soc., Perkin Trans. II*, 1979, 337; (b) M. J. Kamlet, R. M. Doherty, M. H. Abraham, Y. Marcus and R. W. Taft, *J. Phys. Chem.*, 1988, **92**, 5244.
- (a) C. Laurence, P. Nicolet, M. T. Dalati, J. -L. M. Abboud and R. J. Notario, *Phys. Chem.*, 1994, **98**, 5807; (b) Y. Marcus: The Properties of Solvents, Wiley, Chichester, 1998, Chapter 4, p. 131; (c) Y. Marcus, *Chem. Soc. Rev.*, 1993, **22**, 409; (d) R. Lungwitz, M. Friedrichb, W. Linertc and S. Spange, *New J. Chem.* 2008, **32**, 1493.
- Gaussian 03, Revision E.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, Jr., J. A. Montgomery, T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez and J. A. Pople, Gaussian, Inc., Wallingford CT, 2004.
- J. H. Hildebrand, J. M. Prausnitz and R. L. Scott: Regular and Related Solutions, Van Nostrand-Reinhold, Princeton, 1970.
- (a) U. Mayer, V. Gutmann and W. Gerger, *Monatsh. Chem.*, 1975, **106**, 1235; (b) U. Mayer, W. Gerger and V. Gutmann, *Monatsh. Chem.*, 1977, **108**, 489; (c) V. Gutmann, *Electrochimica Acta*, 1976, **21**, 661; (d) V. Gutmann, *CHEMTECH*, 1977, **7**, 255; (e) A. J.

- Parker, U. Mayer, R. Schmid and V. J. Gutmann, *Org. Chem.*, 1978, **43**, 1843; (f) R. J. Schmid, *Sol. Chem.*, 1983, **12**, 135; (g) R. Schmid and N. V. Sapunov: Non-Formal Kinetics in Search for Chemical Reaction Pathways, Verlag Chemie, Weinheim, 1982; (h) R. J. Schmid, *Sol. Chem.*, 1983, **12**, 135.
- 25 (a) L. P. Hammett, *J. Am. Chem. Soc.*, 1937, **59**, 96; (b) L. P. Hammett, *Trans. Faraday Soc.*, 1938, **34**, 156; (c) M. S. Greenberg, R. L. Bodner and A. I. Popov, *J. Phys. Chem.*, 1973, **77**, 2449.
- 26 (a) G. Angelini, P. De Maria, A. Fontana, M. Pierini and G. Siani, *J. Org. Chem.*, 2007, **72**, 4039; (b) G. Angelini, C. Coletti, P. De Maria, R. Ballini, C. Gasbarri, A. Fontana, M. Pierini and G. Siani, *J. Org. Chem.*, 2012, **77**, 899; (c) R. Ballini, S. Cerritelli, A. Fontana, P. De Maria, M. Pierini and G. Siani, *Eur. J. Org. Chem.*, 2000, 1641; (d) G. Angelini, C. Chiappe, P. De Maria, A. Fontana, F. Gasparrini, D. Pieraccini, M. Pierini and G. Siani, *J. Org. Chem.*, 2005, **70**, 8193; (e) F. Gasparrini, M. Pierini, C. Villani, P. De Maria, A. Fontana and R. Ballini, *J. Org. Chem.*, 2003, **68**, 3173; (f) G. Siani, G. Angelini, P. De Maria, A. Fontana and M. Pierini, *Org. Biomol. Chem.*, 2008, **6**, 74, 4236; (g) W. Cabri, I. D'Acquarica, P. Simone, M. Di Iorio, M. Di Mattia, F. Gasparrini, F. Giorgi, A. Mazzanti, M. Pierini, M. Quaglia and C. Villani, *J. Org. Chem.*, 2011, **76**, 1751.
- 27 (a) R. E. Hurd, *J. Magn. Reson.*, 1990, **87**, 422; (b) M. von Kielin, C. T. W. Moonen, A. van der Toorn and P. C. M. van Zijl, *J. Magn. Reson.*, 1991, **93**, 423; (c) W. F. Reynolds and R. G. Enriquez, *J. Nat. Prod.*, 2002, **65**, 221.
- 28 (a) S. A. Bradley and K. Krishnamurthy, *Magn. Res. Chem.*, 2005, **43**, 117; (b) G. Kontaxis, J. Stonehouse, E. D. Laue and J. J. Keeler, *Magn. Reson. Ser. A*, 1994, **111**, 70.
- 29 (a) R. E. Hurd and B. K. John, *J. Magn. Reson.*, 1991, **91**, 648; (b) W. Willker, D. Leibfritz, R. Kerssebaum and W. Bermel, *Magn. Res. Chem.*, 1993, **31**, 287.
- 30 (a) K. Stott, J. Stonehouse, J. Keeler, T. -L. Hwand and A. J. Shaka, *J. Am. Chem. Soc.*, 1995, **117**, 4199; (b) K. Stott, J. Keeler, Q. N. Van and A. J. Shaka, *J. Magn. Reson.*, 1997, **125**, 302; (c) Q. N. Van, E. M. Smith and A. J. Shaka, *J. Magn. Reson.*, 1999, **141**, 191.
- 31 F. Gasparrini, L. Lunazzi, A. Mazzanti, M. Pierini, K. M. Pietrusiewicz and C. Villani, *J. Am. Chem. Soc.*, 2000, **122**, 4776.
- 32 F. Gasparrini, I. D'Acquarica, M. Pierini and C. Villani, *J. Sep. Sci.*, 2001, **24**, 941.
- 33 C. Dell'Erba, F. Gasparrini, S. Grilli, L. Lunazzi, A. Mazzanti, M. Novi, M. Pierini, C. Tavani and C. Villani, *J. Org. Chem.*, 2002, **67**, 1663.
- 34 F. Gasparrini, S. Grilli, R. Leardini, L. Lunazzi, A. Mazzanti, D. Nanni, M. Pierini and M. Pinamonti, *J. Org. Chem.*, 2002, **67**, 3089.
- 35 A. Dalla Cort, F. Gasparrini, L. Lunazzi, L. Mandolini, A. Mazzanti, C. Pasquini, M. Pierini, R. Rompietti and L. Schiaffino, *J. Org. Chem.*, 2005, **70**, 8877.
- 36 A. Ciogli, A. Dalla Cort, F. Gasparrini, L. Lunazzi, L. Mandolini, A. Mazzanti, C. Pasquini, M. Pierini, L. Schiaffino and F. YaftehMihan, *J. Org. Chem.*, 2008, **73**, 6108.
- 37 (a) R. Cirilli, R. Ferretti, F. La Torre, D. Secci, A. Bolasco, S. Carradori and M. Pierini, *J. Chromatogr. A*, 2007, **1172**, 160; (b) R. Cirilli, R. Costi, R. Di Santo, F. La Torre, M. Pierini and G. Siani, *Anal. Chem.*, 2009, **81**, 3560; (c) R. Cirilli, R. Costi, R. Di Santo, F. Gasparrini, F. La Torre, M. Pierini and G. Siani, *Chirality*, 2009, **21**, 24.
- 38 M. Jung, *QCPE Bull.*, 1992, **12**, 52.
- 39 O. Trapp and V. Schurig, *Comput. Chem.*, 2001, **25**, 187.
- 40 P. Wang and A. Anderko, *Fluid Phase Equilib.*, 2001, **186**, 103.

TABLE OF CONTENTS

Thermodynamic and Kinetic Investigation of Monoketo-Aldehyde-Peroxyhemiacetal-(MKA), a Stereolabile Degradation Product of Dihydroartemisinin.

Dorina Kotoni, Monica Piras, Walter Cabri, Fabrizio Giorgi, Andrea Mazzanti, Marco Pierini, Marco Quaglia, Claudio Villani, Francesco Gasparrini.

The $\beta \rightleftharpoons \alpha$ epimerization process of Monoketo-Aldehyde-Peroxyhemiacetal-(MKA), a stereolabile degradation product of the powerful antimalarial drug Dihydroartemisinin, has been in depth studied under a thermodynamic and kinetic point of view in several solvents. A rationalization of the factors responsible for the position assumed by the equilibrium has been achieved by resorting to LSER analyses and molecular modeling calculations.

