Investigation of Dynamic Processes and Self-Organization by Spectroscopic Methods and Molecular Modelling

Ph.D. Thesis

Anasztázia Hetényi

Institute of Pharmaceutical Chemistry
University of Szeged

2005

Table of Contents

Abbreviations and Symbols

2D	two-dimensional
^3J	vicinal scalar coupling constant
δ	chemical shift
ΔE	conformational energy difference
AE	anomeric effect
B3LYP	Becke's gradient-corrected exchange-correlation density functional
CD	circular dichroism
cis-ACPC	*cis*-2-aminocyclopentanecarboxylic acid
COSY	correlated spectroscopy
DLS	dynamic light scattering
DOSY	diffusion-ordered spectroscopy
EN	electronegativity
EXSY	exchange spectroscopy
FT-IR	Fourier transform-infrared
HF	Hartree–Fock
HMBC	heteronuclear multiple bond correlation spectroscopy
HPLC	high-performance liquid chromatography
HSQC	heteronuclear single-quantum coherence
INEPT	insensitive nuclei enhanced by polarization transfer
ISPA	isolated spin pair approximation
LFER	linear free energy relationship
NBO	natural bond orbital
NMR	nuclear magnetic resonance
NOE	nuclear Overhauser effect
NOESY	nuclear Overhauser effect spectroscopy
PFGSE	pulsed field gradient spin echo
R	intensity ratio
RHF	spin-restricted Hartree–Fock theory
ROESY	rotating frame Overhauser effect spectroscopy
RSA	restrained simulated annealing
TEM	transmission electron microscopy
TFA	trifluoroacetic acid
TMS	tetramethylsilane
TOCSY	total correlated spectroscopy
trans-ACHC	*trans*-2-aminocyclohexanecarboxylic acid
VT-NMR	variable-temperature NMR

Publications

Related to the thesis

1. **Anasztázia Hetényi**, Zsolt Szakonyi, Karel D. Klika, Kalevi Pihlaja, Ferenc Fülöp:
   Formation and characterisation of a multicomponent equilibrium system derived from *cis*- and *trans*-1-aminomethylcyclohexane-1,2-diol.
   J. Org. Chem. **2003**, *68*, 2175-2182.

2. **Anasztázia Hetényi**, Tamás A. Martinek, László Lázár, Zita Zalán, Ferenc Fülöp:
   Substituent-dependent negative hyperconjugation in 2-aryl-1,3-*N,N*-heterocycles. Fine-tuned anomeric effect?
   J. Org. Chem. **2003**, *68*, 5705-5712.

3. **Anasztázia Hetényi**, István M. Mándity, Tamás A. Martinek, Gábor K. Tóth, Ferenc Fülöp:
   Chain-length-dependent helical motifs and self-association of *β*-peptides with constrained side chains.
   J. Am. Chem. Soc. **2005**, *127*, 547-553.

4. Tamás A. Martinek, **Anasztázia Hetényi**, Lívia Fülöp, István M. Mándity, Gábor K. Tóth, Imre Dékány, Ferenc Fülöp:
   Biomimicking tertiary structures of *β*-peptides form nano-sized fibrils and membranes.
   Angew. Chem. Int. Ed. submitted

Other papers

5. Zsolt Szakonyi, Tamás Martinek, **Anasztázia Hetényi**, Ferenc Fülöp:
   Synthesis and transformations of enantiomeric 1,2-disubstituted monoterpene derivatives.
   Tetrahedron: Asymmetry **2000**, *11*, 4571-4579.

6. Márta Palkó, **Anasztázia Hetényi**, Ferenc Fülöp:
   Synthesis and stereochemistry of indano[1,2-*d*][1,3]oxazines and thiazines, new ring systems.
   J. Heterocyclic Chem. **2004**, *41*, 69-75.

7. Ferenc Csende, **Anasztázia Hetényi**, Géza Stájer, Ferenc Fülöp:
   Synthesis and structure of cycloalkane- and norbornane-condensed 6-aryl-1,2,4,5-tetrahydropyridazinones.
   J. Heterocyclic Chem. **2004**, *41*, 259-261.

8. Ferenc Miklós, **Anasztázia Hetényi**, Pál Sohár, Géza Stájer:
   Preparation and structure of di-*exo*-condensed norbornane heterocycles.
   Monatsh. Chem. **2004**, *135*, 839-847.

9. István Szatmári, **Anasztázia Hetényi**, László Lázár, Ferenc Fülöp:
   Transformation reactions of the Betti base analog aminonaphthols.
   J. Heterocyclic Chem. **2004**, *41*, 367-373.

10. Zita Zalán, **Anasztázia Hetényi**, László Lázár, Ferenc Fülöp:
    Substituent effects in the ring–chain tautomerism of 4-aryl-1,3,4,6,7,11b-hexahydro-2H-pyrimido[6,1-*a*]isoquinolines.
    Tetrahedron **2005**, *61*, 5287-5295.

11. Zsolt Szakonyi, Szilvia Gyónfalvi, Enikő Forró, **Anasztázia Hetényi**, Norbert De Kimpe, Ferenc Fülöp:
    Synthesis of 3- and 4-hydroxy-2-aminocyclohexanecarboxylic acids by iodocyclization.
    Eur. J. Org. Chem. **2005**, in press.

Conference lectures related to the thesis

12. **Anasztázia Hetényi**, Karel D. Klika, Kalevi Pihlaja, Zsolt Szakonyi, Ferenc Fülöp:
    A study of the conformation and ring–chain tautomerism of aminodiol derivatives by NMR spectroscopy.
    2001 Nordic NMR Symposium, 23rd Finnish NMR Symposium
    Helsinki, Finland, 26-29 August, 2001, Abstr.: 33 old.

13. Zsolt Szakonyi, **Anasztázia Hetényi**, Ferenc Fülöp, Karel D. Klika, Kalevi Pihlaja:
    Stereoselective synthesis and ring enclosure of carbocyclic aminodiols.
    85th CSC Conference & Exhibition
    Vancouver, Canada, 1-5 June, 2002, Abstr.: 1136 OR PS.

14. Martinek A. Tamás, **Hetényi Anasztázia**, Zalán Zita, Lázár László, Fülöp Ferenc:
    Sztereoelektronikus stabilizáció detektálása nem periplanáris geometriák esetén. Finomhangolható anomer-effektus?
    Elméleti Szerveskémiai Munkabizottsági ülés
    Budapest, 2003. január 30.

15. **Hetényi Anasztázia**, Martinek Tamás, Lázár László, Zalán Zita, Fülöp Ferenc:
    Sztereoelektronikus hatások szerepe a 2-arilszubsztituált 1,3-X,N-heterociklusos epimerek szubsztituensfüggő stabilitáskülönbségében.
    Vegyészkonferencia 2003
    Hajdúszoboszló, 2003. június 26-28., Abst.: P-39.

16. **Hetényi Anasztázia**:
    Ciklusos β-aminosav oligomerek térszerkezetének változása a lánchosszal.
    A Szegedi Ifjú Szerves Kémikusok Támogatásáért Alapítvány és a SZAB Szerves és Gyógyszerkémiai Munkabizottság 4. tudományos előadásülése
    Szeged, 2004. január 14.

17. **Hetényi Anasztázia**, Mándity M. István, Martinek A. Tamás, Tóth K. Gábor, Fülöp Ferenc:
    Konformációsan gátolt β-peptidek lánchosszfüggő helikális szerkezete és önrendeződése.
    MTA-Peptidkémiai Munkabizottság ülése
    Balatonszemes, 2004. május 26-28.

18. **Anasztázia Hetényi**, István Mándity, Tamás A. Martinek, Gábor K. Tóth, Ferenc Fülöp:
    True folding of conformationally constrained β-peptides: chain length-dependent secondary structure.
    3rd International and 28th European Peptide Symposium
    Prague, Czech Republic, 5-10 September, 2004, Abstr.: P-590

19. Martinek Tamás, **Hetényi Anasztázia**, Mándity István, Fülöp Lívia, Tóth Gábor, Fülöp Ferenc:
    Béta-peptidek kiralitással szabályozott harmadlagos szerkezetei.
    MTA-Peptidkémiai Munkabizottság ülése
    Balatonszemes, 2005. május 30-június 1.

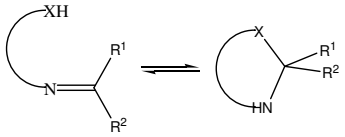
20. Martinek Tamás, Mándity M. István, **Hetényi Anasztázia**, Tóth K. Gábor, Forró Enikő, Fülöp Ferenc:
    További lépések a β-peptidek harmadlagos szerkezete felé.
    Vegyészkonferencia 2005
    Hajdúszoboszló, 2005. június 28-30., Abstr.: P-62.

1. Introduction and Goals

In recent decades, 1,3-*X,N*-heterocycles[1,2] and *β*-amino acid derivatives[3] have played an important role in drug therapy and drug research. These compounds serve as good models for a better comprehension of several important phenomena in chemistry, for example the dynamic processes of ring–chain tautomerism and ring–ring epimerization and the behaviour of self-organizing foldameric systems.

Ring–chain tautomerism involving the reversible addition of a heteroatom to a heteropolar double bond has been intensively studied for 1,3-*X,N*-heterocycles (Scheme 1). The equilibria between open-chain **A** and ring-closed **B** species can be investigated by NMR.

SCHEME 1



A **B**

X = O, S, NR

The tautomeric equilibria of the 2-aryl-1,3-*X,N*-heterocycles (K_X = **B/A**) can be described successfully by the Hammett–Brown linear free energy relation ship (LFER). The LFER cannot only describe two- or three-component equilibria,[1,2] but is also applicable to more complex systems.[4-7] The earlier results indicated that the proportions of the ring-closed forms strongly depend on the electronic character of the substituent on the aromatic ring. The scope and limitations of LFER can be studied from the aspect of the influence of the electronic effects of the substituent on the aromatic ring, and its applicability to complex multicomponent tautomeric mixtures. It is still an open question whether the substituent-dependent stability difference between the epimeric forms in the ring–ring epimerization can be explained by the stereoelectronic interactions related to the anomeric effect.

For a long time, only proteins, RNA and DNA were known to be able to form a well-defined hierarchy of self-assembling structural levels. The hallmark of folding is the ability of an amino acid sequence to reversibly acquire a well-defined, unique structure in a moderate amount of time, even though an exponentially large number are possible.[8] The unnatural foldamers with well-defined secondary structure preferences, especially *β*-peptides, are attracting increasing interest and have a wide range of potential applications because they are able to adopt specific, compact conformations,[9-15] and they could be used to create new types of tertiary structures. The folding propensity is influenced by local torsional, side-chain to backbone and long-range side-chain interactions. Although *β*-peptide foldamers are sensitive to the solvent, the systematic choice of the side-chain pattern and spatiality allows the design

of the desired specific secondary structure.[14,16-20] The conformationally constrained β-peptide oligomers containing cyclic side-chains are among the most thoroughly studied models in foldamer chemistry. The chain length-independence of the folding pattern of β-peptide oligomers raises the still open question of whether these oligomers exhibit a real folding process, or whether the conformational space of the monomers is too preorganized to allow partly folded or other stable secondary structures. Another major challenge in foldamer science is to prove that higher-order structural levels are available for β-peptides and that their formation can be tailored by the β-amino acid sequence. Certain designed β-peptide helices have been observed to exhibit self-association, but distinct tertiary structures with specific morphology on the nanometre scale have not been reported thus far.

In the field of the dynamic processes, the scope of my PhD thesis covers an analysis of the limitations of the Hammett–Brown equation in the ring–chain tautomerism of *cis*- and *trans*-1-aminomethylcyclohexane-1,2-diols and in the ring–ring epimerization of 2-aryl-1,3-*N,N*-heterocyclic derivatives. In the field of the self-organizing foldamers, the goal was to study the limits of the conformational flexibility and self-assembling properties of the β-peptides with constrained side-chains for the *trans*-2-aminocyclohexanecarboxylic acid and *cis*-2-aminocyclopentanecarboxylic acid foldamers.

In order to achieve the above goals, mainly NMR spectroscopy was utilized, together with some other complementary methods, including IR, CD, DLS and TEM, besides molecular modelling. Accordingly, a brief theoretical background of the techniques used and the experimental set-ups is presented in section 2. Finally, to provide the thesis with sufficient clarity, the two major topics (dynamic processes and self-organization phenomena) are discussed in separate sections.

2. Methods

2.1. NMR Spectroscopy

2.1.1. Homonuclear Coherent Magnetization Transfer

COSY (COrrelated SpectroscopY).[21,22] Cross-peaks arise through coherence transfer between scalar coupled spins. The 2D experiment is composed of a $90^o_x - t_1 - 90^o_x$ pulse sequence element. The first pulse creates transverse magnetization components for all allowed transitions. During the evolution period, the various magnetization components are labelled with their characteristic precession frequencencies. Finally, the mixing pulse transfers magnetization components among all those transitions that belong to *J*-coupled spins. For a two-spin system, initial I_{1z} magnetization evolves through the pulse sequence elements as:

$$I_{1z} \xrightarrow{\left(\frac{\pi}{2}\right)_x - t_1 - \left(\frac{\pi}{2}\right)_x} -I_{1z}\cos(\Omega_1 t_1)\cos(\pi J t_1) - 2I_{1x}I_{2y}\cos(\Omega_1 t_1)\sin(\pi J t_1) + I_{1x}\sin(\Omega_1 t_1)\cos(\pi J t_1) - 2I_{1z}I_{2y}\sin(\Omega_1 t_1)\sin(\pi J t_1) \quad (1)$$

where I_k for $k = 1, 2, \ldots, K$ for a *K*-spin system, Ω_k is a chemical shift of the k^{th} spin and J_{jk} is a scalar coupling constant between the j^{th} and k^{th} spins.

TOCSY (TOtal Correlation SpectroscopY).[23,24] Cross-peaks are generated between all members of a coupled spin network. The 2D experiment is composed of a $90^o_x - t_1 - 90^o_x - \tau_m - 90^o_x$ pulse sequence element. After the evolution period (t_1), the magnetization is spin-locked with a composite pulse scheme in which the magnetization behaves as a strongly coupled spin system and evolves under the influence of a "collective spin-mode". In that collective mode, coherence transfer is possible between all coupled nuclei in a spin system. Following the $90^o_x - t_1 - 90^o_x$ sequence, which is decribed by equation 1, the terms I_{1x} and $2I_{1z}I_{2y}$ are eliminated by the TOCSY phase cycle. During the isotropic mixing sequence, magnetization proportional to I_{1z} is transferred through the spin system via a strong coupling Hamiltonian as follows:

$$-I_{1_z}\cos(\Omega_1 t_1)\cos(\pi J t_1) \xrightarrow{\tau_m} -\sum_{k=1}^{K} I_{kz}a_{1k}(\tau_m)\cos(\Omega_1 t_1)\cos(\pi J t_1) \quad (2)$$

in which $a_{1k}(\tau_m)$ are mixing coefficients for the transfer of magnetization through the spin system from spin I_1 to spin I_k, and zero-quantum terms have been ignored.

$$a_{1k}(\tau_m):\ a_{11}(\tau_m) = a_{22}(\tau_m) = 1 - \sin^2\Phi\sin^2(q\tau_m)$$

$$a_{12}(\tau_m) = a_{21}(\tau_m) = \sin^2\Phi\sin^2(q\tau_m)$$

After the last 90^{o}_{x} pulse and Hahn echo, the density operator prior to t_2 is given by:

$$\xrightarrow{\left(\frac{\pi}{2}\right)_x} \sum_{k=1}^{K} I_{ky} a_{1k}(\tau_m) \cos(\Omega_1 t_1) \cos(\pi J t_1) \tag{3}$$

2.1.2. Heteronuclear Coherent Magnetization Transfer

HSQC (Heteronuclear Single-Quantum Coherence).[25-28] The experiment correlates protons with their directly attached *X*-nuclei. The experiment starts with proton magnetization. The first INEPT step is used to create proton antiphase magnetization, which is then transferred to the directly attached heteronuclei. This *X*-nucleus magnetization is left to evolve with its chemical shift. The effects of proton coupling and chemical shift are removed by the use of an 180° proton pulse applied at mid-evolution time.

For an *IS* heteronuclear spin system (I = ^{1}H, S = ^{13}C or ^{15}N), in which the *I* and *S* spins are covalently bonded, the proton *I* is scalar-coupled to a remote proton, *K*, and the delay $2\tau = 1/(2J_{IS})$, the pulse sequence elements are as follows:

$$I_z \xrightarrow{\frac{\pi}{2}(I_x+K_x)-\tau-\pi(I_x+K_x),\pi S_x-\tau-\frac{\pi}{2}(I_y+K_y),\frac{\pi}{2}S_x} -2I_z S_y \tag{4}$$

$$\xrightarrow{\frac{t_1}{2}-\pi(I_x+K_x)-\frac{t_1}{2}} 2I_z S_y \cos(\Omega_s t_1) - 2I_z S_x \sin(\Omega_s t_1) \tag{5}$$

$$\xrightarrow{\frac{\pi}{2}(I_x+K_x),\frac{\pi}{2}S_x-\tau-\pi(I_x+K_x),\pi S_x-\tau} -I_x \cos(\Omega_s t_1) - 2I_y S_x \sin(\Omega_s t_1) \tag{6}$$

HMBC (Heteronuclear Multiple Bond Correlation spectroscopy).[29,30] This is suitable for determining long-range ^{1}H-*X* connectivity. The pulse sequence utilizes zero- and double-quantum coherence between *J*-coupled protons and *X*-nuclei to label each proton with the frequency of a remote *X* in the F_1 dimension of a 2D experiment. The experiment converts transverse magnetization into zero- and double-quantum coherence, delay is matched to the inverse of the long-range coupling constant $^{n}J_{XH}$, and there is a filter to suppress cross-peaks arising from one-bond proton-carbon interactions. The one-bond filter has a delay that is matched to the inverse of $^{1}J_{XH}$:

$$I_z \xrightarrow{\frac{\pi}{2}(I_x+K_x)-2\tau-\frac{\pi}{2}S_x} -2I_x S_y \tag{7}$$

$$\xrightarrow{\frac{t_1}{2}-\pi(I_x+K_x)-\frac{t_1}{2}} -2I_x S_y \cos(\Omega_s t_1)\cos(\pi J_{IK} t_1) - 4I_y K_z S_y \cos(\Omega_s t_1)\sin(\pi J_{IK} t_1) \tag{8}$$

$$\xrightarrow{\frac{\pi}{2}S_x-2\tau} -I_y \cos(\Omega_s t_1)\cos(\pi J_{IK} t_1) + 2I_x K_z \cos(\Omega_s t_1)\sin(\pi J_{IK} t_1) \tag{9}$$

2.1.3. Homonuclear Incoherent Magnetization Transfer

NOESY (Nuclear Overhauser Effect SpectroscopY).[31,32] The experiment allows the correlation of nuclei through space (distance < 5 Å). The 2D experiment involves a $90^o_x - t_1 - 90^o_x - \tau_m - 90^o_x$ pulse sequence element. The second 90° pulse transmits some of the magnetization to the Z-axis. During the following mixing period, this non-equilibrium Z component will exchange magnetization through relaxation, which is known as NOE. After a shorter time than the relaxation time T_1, the transverse magnetization is restored and detected. If relaxation exchange (or chemical exchange) has taken place during the mixing time, cross-peaks will be observed in the spectrum. For a spin system containing two scalar-coupled spins, evolution through the $90^o_x - t_1 - 90^o_x$ sequence is described by equation 1. Evolution of the I_{1z} term in equation 1 during τ_m is governed by the Solomon equations, in which the initial condition is $-I_{1_z}\cos(\Omega_1 t_1)\cos(\pi J_{12} t_1)$ and the equilibrium magnetization, I^o_{1z}, is rejected by phase cycling for axial peak suppression. If $K - 1$ spins (I_k for $k = 2, \ldots, K$) are close in space to spin I_1, the resulting evolution during τ_m is

$$-I_{1z}\cos(\Omega_1 t_1)\cos(\pi J_{12} t_1) \xrightarrow{\tau_m} -\sum_{k=1}^{K} I_{kz} a_{1k}(\tau_m)\cos(\Omega_1 t_1)\cos(\pi J_{12} t_1) \tag{10}$$

in which $a_{1k}(\tau_m) = [\exp(-\mathbf{R}\tau_m)]_{1k}$ is the $(1, k)^{th}$ element of the matrix exponential and **R** is the matrix of the rate constants ρ_i and σ_{ij}. After the final 90° pulse, the density operator terms that result from the longitudinal magnetization are given by

$$\sum_{k=1}^{K} I_{ky} a_{1k}(\tau_m)\cos(\Omega_1 t_1)\cos(\pi J_{12} t_1) \tag{11}$$

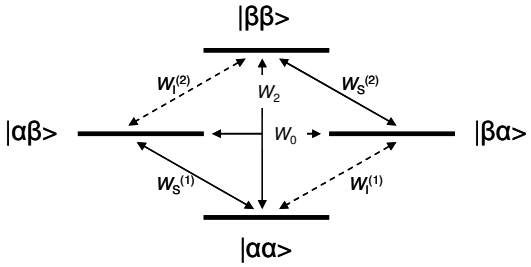


FIGURE 1. The relaxation-induced transitions, which are possible amongst the same set of levels.

This two-spin system can be described by the Solomon equations as follows:

$$\frac{d\Delta I_z(t)}{dt} = -\rho_I \Delta I_z(t) - \sigma_{IS}\Delta S_z(t);\ \frac{d\Delta S_z(t)}{dt} = -\rho_S \Delta S_z(t) - \sigma_{IS}\Delta I_z(t) \tag{12}$$

where ρ_I and ρ_S are the autorelaxation rate constants for the I and S spins, respectively, and $\sigma_{IS} = W_2 - W_0$ is the cross-relaxation rate, W_0 and W_2 are the transition probabilities for the zero-quantum and double-quantum transition, respectively, $\Delta I_z(t) = \langle I_z\rangle(t) - \langle I_z^0\rangle$, and $\langle I_z^0\rangle$ is the

equilibrium magnitude of the I_z operator. Corresponding relationships hold for S_z. Generally, for a multispin system in matrix nomenclature:

$$\frac{d\Delta\mathbf{M}_z(t)}{dt} = -\mathbf{R}\Delta\mathbf{M}_z(t) \tag{13}$$

in which **R** is an *N x N* matrix with elements $R_{kk} = \rho_k$ and $R_{kj} = \sigma_{kj}$, and $\Delta\mathbf{M}_z(t)$ is an N x 1 column vector with entries $\Delta M_z(t) = \Delta I_{kz}(t)$.

The cross-peak intensities contain information about the relative distances of the nuclei that contribute to the cross-peak. The build-up of cross-peak intensity in a multispin system is given by:

$$A_{ij}(\tau_m) = \{-R_{ij}\tau_m + \frac{1}{2}\sum R_{ij}R_{kj}{\tau_m}^2 + ...\}A_{ii}(0) \tag{14}$$

where $A(\tau_m)$ is the cross-peak intensity as a function of the mixing time τ_m, and R_{ij} is the relaxation rate between spins *i* and *j*.

For a short mixing time, the cross-peak intensity is directly proportional to the cross-relaxation rate and thus to the distance between the nuclei, because the quadratic term and those of higher order in τ_m can be neglected.

$$A_{NOE} \approx \sigma_{ij} \approx {r_{ij}}^{-6} \tag{15}$$

where

$$\sigma_{ij} = \frac{\gamma^4 h^2 \tau_c}{4\pi^2 10 {r_{ij}}^6}\left(\frac{6}{1+4\omega^2{\tau_c}^2} - 1\right) \tag{16}$$

in which σ is the cross-relaxation rate, γ is the gyromagnetic ratio, h is Planck's constant, τ_c is the correlation time, ω is the Larmor frequency, and r is the distance between nuclei.

On the basis of the isolated spin-pair approximation (ISPA), which is the assumption that the cross-relaxation can be treated as occurring between an isolated pair of spins *i* and *j*, since the influence of any other nearby nuclei is minimal, the distance between nuclei can be calculated as follows:

$$r = \left(\frac{I_{ref}}{I}\right)^{1/6} r_{ref} \tag{17}$$

where '*ref*' means the reference intensity and distance.

ROESY (Rotating frame Overhauser effect SpectroscopY).[33] The sequence is very similar to that of NOESY in that it provides information concerning the distance between nuclei. The 2D experiment involves a $90^o_x - t_1 - spinlock$ pulse sequence element. This technique is based on NOE in the rotating frame. For a scalar-coupled two-spin system, I_{1z} is given by:

$$I_{1z} \xrightarrow{(\pi/2)_x - t_1} -I_{1y}\cos(\Omega_1 t_1)\cos(\pi J_{12} t_1) + 2I_{1x}I_{2z}\cos(\Omega_1 t_1)\sin(\pi J_{12} t_1) \quad (18)$$
$$+ I_{1x}\sin(\Omega_1 t_1)\cos(\pi J_{12} t_1) + 2I_{1y}I_{2z}\sin(\Omega_1 t_1)\sin(\pi J_{12} t_1)$$

During the subsequent spin-locking period:

$$I_{1y} \Rightarrow I_{1z}^{'}\sin\theta_1 + I_{1y}^{'}\cos\theta_1$$
$$2I_{1y}I_{2z} \Rightarrow 2\left(I_{1z}^{'}\sin\theta_1 + I_{1y}^{'}\cos\theta_1\right)\left(I_{2z}^{'}\cos\theta_2 - I_{2y}^{'}\sin\theta_2\right) \quad (19)$$

in which θ_1 and θ_2 are the tilt angles of the effective field for spins I_1 and I_2. Those terms that commute with the spin-lock Hamiltonian are proportional to I_{1z}' and $2I_{1z}'I_{2z}$'. If K – 1 spins (I_k for k = 2, …, K) are close in space to spin I_1, the resulting evolution of the longitudinal magnetization is:

$$-I_{1z}^{'}\sin\theta_1\cos(\Omega_1 t_1)\cos(\pi J_{12} t_1) \xrightarrow{\tau_m} -\sum_{k=1}^{K} I_{kz}^{'} a_{1k}(\tau_m)\sin\theta_1\cos(\Omega_1 t_1)\cos(\pi J_{12} t_1) \quad (20)$$

in which $a_{1k}(\tau_m) = [\exp(-\mathbf{R}\tau_m)]_{1k}$ is the $(1, k)^{th}$ element of the matrix exponential and **R** is the matrix of the rotating frame relaxation rate constants $R_{kk}(\theta_i)$ and $\sigma_{jk}(\theta_i, \theta_j)$. Transforming back from the tilted frame to the rotating frame:

$$\sum_{k=1}^{K} I_{ky} a_{1k}(\tau_m)\sin\theta_1\sin\theta_k\cos(\Omega_1 t_1)\cos(\pi J_{12} t_1) \quad (21)$$

The cross-peaks are generated by cross-relaxation of transversal magnetization. In the rotating frame, σ is given by:

$$\sigma_{ij} = \frac{\gamma^4 h^2 \tau_c}{4\pi^2 10 r^6}\left(\frac{3}{1+4\omega^2\tau_c^{\,2}} + 2\right) \quad (22)$$

2.2. Other Techniques

2.2.1. Theoretical Methods

Force Field Calculations. A molecular mechanics/dynamics force field may reasonably be asked to reproduce all of a number of molecular properties, including the following: molecular geometries, conformational and stereoisomeric energies, torsional barriers and torsion-deformation energies, intermolecular-interaction energies and geometries, vibrational frequencies and heats of formation.

During our calculations, the MMFF94 force field was used. The potential energy of the MMFF94 can be written:[34]

$$E_{MMFF94} = \Sigma EB_{ij} + \Sigma EA_{ijk} + \Sigma EBA_{ijk} + \Sigma EOOP_{ijk;l} + \Sigma ET_{ijkl} + \Sigma EvdBW_{ij} + \Sigma EQ_{ij} \quad (23)$$

where ΣEB_{ij} relates to the bond-stretching, ΣEA_{ijk} to the angle-bending, ΣEBA_{ijk} to the stretch-bend interactions, $\Sigma EOOP_{ijk;l}$ to the out-of-plane bending at tricoordinate centres, ΣET_{ijkl} to the torsion interactions, $\Sigma EvdBW_{ij}$ to the van der Waals interactions, and ΣEQ_{ij} to the electrostatic interactions, and i, j, k, and l denote different atoms.

Quantum Mechanics. The starting point of quantum mechanics is the time-independent Schrödinger equation:

$$\left\{-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)+V\right\}\Psi(\mathbf{r})=E\Psi(\mathbf{r}) \tag{24}$$

where m is the mass of an electron, V is the external field, $\hbar$ is Planck's constant divided by 2π, E is the energy of the particle, Ψ is a wavefunction which characterizes the motion of the particle, and $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ is the position vector.

Hartree–Fock Methods. The Hartree–Fock (HF) equations are obtained when the first derivative of the energy, δE, is zero in the time-independent Schrödinger equation of a many-electron system in a coulombic potential of fixed nuclei. As the nuclei are modelled as fixed point charges, it is only applicable after the Born–Oppenheimer approximation has been made. The starting point for the HF method is a set of approximate orbitals. Typically, in HF calculations, the wavefunction is approximated as a product of one-electron wavefunctions, which are in turn approximated by a linear combination of atomic orbitals. HF calculations can be used as the starting point for more sophisticated methods, such as many-body perturbation theory.

Density Functional Methods. Density functional theory is an approach to the electronic structure of atoms and molecules. From the Hohenberg–Kohn theorem, the total electronic energy as a function of the electron density ρ is:

$$E(\rho)=E_{KE}(\rho)+E_C(\rho)+E_H(\rho)+E_{XC}(\rho) \tag{25}$$

where $E_{KE}(\rho)$ is the kinetic energy, $E_C(\rho)$ is the electron-nuclear interaction term, $E_H(\rho)$ is the electron-electron coulombic energy, and $E_{XC}(\rho)$ contains the exchange and correlation distribution.

During our calculations, the B3LYP method was used, which gives approximate solutions to both exchange and correlation energies, but is not based on a purely quantum mechanical solution. B3LYP is a hybrid method, which includes a mixture of HF exchange with DFT exchange-correlation. It can be described as follows[35]:

$$A * E_x^{\,Slater} + (1-A) * E_x^{\,HF} + B * \Delta E_x^{\,Becke} + E_C^{\,VWN} + C * \Delta E_C^{\,non-local} \tag{26}$$

Natural Bond Orbital (NBO) analysis. This process generates NBOs.[36-42] These orbitals are localized electron-pair orbitals for bonding pairs and lone pairs. The hybridization of the atoms and the weight of each atom in each localized electron-pair bond is calculated in the idealized Lewis structure. The exact analysis of the delocalization energy contributions to the substituent-dependent AE can be tackled through a second-order perturbative analysis of the Fock matrix in the NBO basis. Although this procedure is known to overestimate stabilizing interactions, they closely parallel the energies afforded by the more accurate Fock matrix deletion method. The second-order energy stabilization due to the electron donation from orbital *i* to orbital *j* is given by equation 27, where q_i is the donor orbital occupancy, ε_i and ε_j are diagonal elements and F_{ij} is the off-diagonal NBO Fock matrix element.

$$E_{ij} = q_i \frac{F_{ij}^2}{\varepsilon_j - \varepsilon_i} \tag{27}$$

2.2.2. Vibrational and Chiroptical Spectroscopies

FT-IR Specroscopy.[43] The IR spectra of polypeptides exhibit a number of amide bands, which represent different vibrational modes of the peptide bond. Of these, the amide I band is most widely used for secondary structure analyses. The amide I band results from the C=O stretching vibration of the amide group coupled to the bending of the N-H bond and the stretching of the C-N bond. These vibrational modes, present as IR bands in the range 1600–1700 cm^{-1}, are sensitive to H-bonding and coupling between the transition dipoles of adjacent peptide bonds, and hence are sensitive to the secondary structure.

Circular Dichroism Spectroscopy.[44] Circular dichroism (CD) spectroscopy is a form of light absorption spectroscopy that measures the difference in absorbance of right- and left-circularly polarized light by a substance. It has been shown that CD spectra between 260 and approximately 180 nm can be analysed for the different types of secondary structures: alpha helix, parallel and antiparallel β-sheet, turn, and other.

2.2.3. Particle Size Measurements

Pulse Field Gradient NMR Spectroscopy. Diffusion-Ordered SpectroscopY (DOSY)[45] seeks to separate the NMR signals of different species according to their diffusion coefficients. A series of gradient spin echo spectra are measured with different pulsed field gradient strengths, and the signal decays are analysed to extract a set of diffusion coefficients with which to synthesize the diffusion domain of a DOSY spectrum. If a spherical molecule is assumed, the diffusion coefficient *D* is described by the Stokes–Einstein equation:

$$D = \frac{kT}{6\pi\eta r_s} \tag{28}$$

where k is the Boltzman constant, T is the temperature, η is the viscosity of the liquid and r_s is the hydrodynamic radius of the molecule.

By the use of a field gradient, molecules can be spatially labelled. If they move after this encoding during the following diffusion time Δ, their new position can be decoded by a second gradient. The measured signal is the integral over the whole sample volume and the NMR signal intensity is attenuated, depending on Δ and the gradient parameters (g, and δ). This intensity change is described by:

$$I = I_0 e^{-D\gamma^2 g^2 \delta^2 (\Delta - \delta/3)} \tag{29}$$

where I is the observed intensity, I_0 is the reference intensity, γ is the gyromagnetic ratio of the observed nucleus, g is the gradient strength, and δ is the length of the gradient.

Dynamic Light Scattering (DLS).[46] Scattering experiments can provide a wealth of detailed information about the structural and dynamic properties of matter. Dynamic light scattering in particular has important applications in particle and macromolecule sizing. When a beam of light passes through a colloidal dispersion, the particles scatter some of the light in all directions. When the particles are very small compared with the wavelength of the light, the intensity of the scattered light is uniform in all directions; for particle diameters > approximately 250 nm, the intensity is angle-dependent. Analysis of the autocorrelation function in terms of the particle size distribution is done by fitting the data numerically with calculations based on assumed distributions.

Transmission Electron Microscopy (TEM).[47] Electron microscopes are scientific instruments that use a beam of highly energetic electrons to examine objects on a very fine scale. This examination can yield the following information: *a*. topography: the surface features of an "how it looks"; *b*. morphology: the shape and size of the particles making up the object; *c*. composition: the elements and compounds that the object is composed of, and the irrelative amounts; and *d*. crystallographic information: how the atoms are arranged in the object.

2.3. Experimental

The NMR spectra were recorded on JEOL L400 and Bruker *Avance* DRX 400 spectrometers, with the deuterium signal of the solvent as the lock. During ^{1}H and ^{13}C measurements, TMS was applied as internal standard. For ^{1}H spectra, the number of scans was usually 128 or 256. For kinetic measurements, 0.046, 0.19, 0.38, 0.51 and 0.675 M CD_3OD solutions were prepared with respect to both the aldehydes and the aminodiols. For study of ring–ring epimerization, 10 mg of heterocycle was dissolved in 500 μL of $CDCl_3$. In these cases, the solutions were left to stand at ambient temperature for 1 day for the equilibria to be established before the VT-NMR spectra were run at 223.1-263.1 K for **11a–d** and **12a–d**, or at 300.0–325.1 K for **11e–i** and **12e–i**. During the signal assignments of *β*-peptide homo-oligomers, 8 mM CD_3OD and DMSO solutions were used at 303.1 K.

The assignment of the ^{1}H and ^{13}C NMR signals could be systematically achieved via the application of the standard 2D methods, such as COSY, TOCSY, HSQC, HMBC, NOESY and ROESY. The COSY measurements were performed with a 45^{o}_{x} pulse for a different peak shape of vicinal and geminal protons. The number of scans was 40; 2k time domain points and 256 increments were applied. The TOCSY measurements were performed with the MLEV17 mixing sequence, with a mixing time of 80 ms; the number of scans was 32; 2k time domain points and 512 increments were applied. During the HSQC measurements, the number of scans was 40. For all the HSQC spectra, 2k time domain points and 256 increments were applied. During the HMBC measurements, the number of scans was 80; 2k time domain points and 512 increments were applied. The NOESY measurements were performed with a mixing time of 400 ms; the number of scans was 64; 2k time domain points and 512 increments were applied. The ROESY measurements were performed with a cw spinlock with gradients and the 3-9-19 pulse sequence for water suppression. For the ROESY spinlock, 225 ms and 400 ms mixing times were used; the number of scans was 64; 2k time domain points and 512 increments were applied. For processing, a cosine bell function was applied before transformation, except for the COSY spectra, for which a sine bell function was used.

The integrated peak volumes were determined by using the program NMRView.[48] The distances from the ROESY cross-peak intensities were calculated by application of the ISPA, using an approximate average of a geminal proton-proton distance of 1.78 Å as reference.

Force Field Calculations were carried out on an HP workstation xw6000 in the Chemical Computing Group's Molecular Operating Environment.[49] For the energy calculations, the MMFF94s force field was used, with a 15 Å cutoff for van der Waals and Coulomb interactions. The lowest-energy structures of the 10- and 14-helices for **22–24** were

determined by using relevant H-bond pillar atom distance restraints: NH_i–CO_{i+1} and NH_i–CO_{i+2} for the 10-helix and the 14-helix, respectively. Before the restrained simulated annealing (RSA), a random structure set of 100 molecules was generated by saving the conformations during a 100 ps dynamics simulation at 1000 K every 1000 steps. The RSA was performed for each structure with an exponential temperature profile in 75 steps, and a total duration of 25 ps, and the H-bond restraints were applied as a 10 kcal mol^{-1} $Å^{-2}$ penalty function. Minimization was applied after every RSA in a cascade manner, using the steepest-descent, conjugate gradient and truncated Newton algorithm.

The molecular structure, stereochemistry and geometry of the ACHC and ACPC oligomers were exclusively defined in terms of their *z*-matrix internal coordinate system. The optimizations were carried out with the Gaussian03 program by using the RHF/3-21G basis set with a default set-up. Single point energies were additionally calculated by using density-functional theory at the B3LYP/6-311G** level.

During the **NBO**, the *ab initio* calculations were carried out with Gaussian 94[50] at the HF/3–21G*//HF/6–31G* level. The NBO analysis[36-42] was performed by using the NBO 3.1 software implemented in Gaussian 94 with default parameters. The overlapping NBO orbitals were visualized with the molecular graphics package MOLEKEL.[51]

The **FT-IR** measurements were performed on a Bio-Rad Digilab Division FTS 65A/869 FT-IR spectrometer. The samples were prepared as 8 mM solutions in CD_3OH. The experiments were run with the spectral window 4000-400 cm^{-1} at an optical resolution of 4 cm^{-1} with 256 scans. The detector was DTGS (deuterated triglycyl sulfate); the sample was placed in a 0.1 mm liquid cell and KBr windows were applied.

CD spectra were measured on a Jobin–Yvon Mark VI dichrograph at 25 °C in a 0.02 cm cell. Four spectra were accumulated for each sample. The baseline spectum recorded with only the solvent was subtracted from the raw data. The concentration of the sample solutions was 4 mM in CD_3OH. Molar ellipticity, [Θ], is given in deg cm^2 $dmol^{-1}$. The data were normalized for the number of chromophores.

The **PFGSE NMR** measurements were performed by using the stimulated echo and longitudinal eddy current delay (LED) sequence. 1 ms was used for the dephasing/refocusing gradient pulse length (δ), and 250 ms for the diffusion delay (Δ). The gradient strength was

changed quadratically from 5% to 95% of the maximum value (B-AFPA 10 A gradient amplifier) and the number of steps was 16. Each measurement was repeated three times with 128 scans and 16k time domain points. For the processing, an exponential window function and single zero filling were applied. The fluctuation of the temperature was less than 0.1 K. Prior to the NMR scans, all the samples were equilibrated for 30 min at 302.98 K.

DLS were measured on a Zetasizer Nano ZS at 25 °C in a maintenance-free ζ-potential cell with a volume of 0.75 mL. The experiments were conducted in back-scattering mode to gain a picture of the particle size distribution in the solvent over a wide diameter range. The DLS autocorrelation data were analysed by utilizing the CONTIN[52] algorithm with maximum-entropy regularization implemented in the progam SEDFIT.[53] In the same experimental set-up, the ζ-potential was measured too.

TEM were measured on a Philips CM 10 transmission electron microscope at 100 kV, routinely at magnifications of 25,000 and 46,000. Following the method of Walsh et al.,[54] 10 µL samples were applied to 400 mesh carbon-coated copper grids, fixed with 0.5% (v/v) glutaraldehyde solution, washed three times with 10 µL droplets of dd water, and stained with 2% (w/v) uranyl acetate.

3. Substituent Effects on Tautomeric Equilibria

3.1. Preliminaries

Ring–chain tautomerism involving the reversible addition of a heteroatom to a heteropolar double bond has been studied for 1,3-*X,N*-heterocycles.[1] Qualitative studies on the equilibria could be established by NMR spectroscopy with the integration of well-separated indicator protons.[1] The tautomeric character of the 1,3-*X,N*-heterocycles can be used in the enantioselective synthesis of chiral amines or as aldehyde or ketone sources in the Hantzsch and Pictet–Spengler reactions.[55-59]

Aliphatic aminodiols play an important role in drug therapy and drug research.[60-65] Besides the pharmacological interest, aminodiols are also useful starting materials for the syntheses of oxazines or oxazolidines, depending upon which hydroxy group undergoes ring closure with the amino group. The products of the reaction between an aldehyde and an alicyclic aminodiol containing three or more functional groups can potentially present a novel example of ring−chain tautomerism.[66-70]

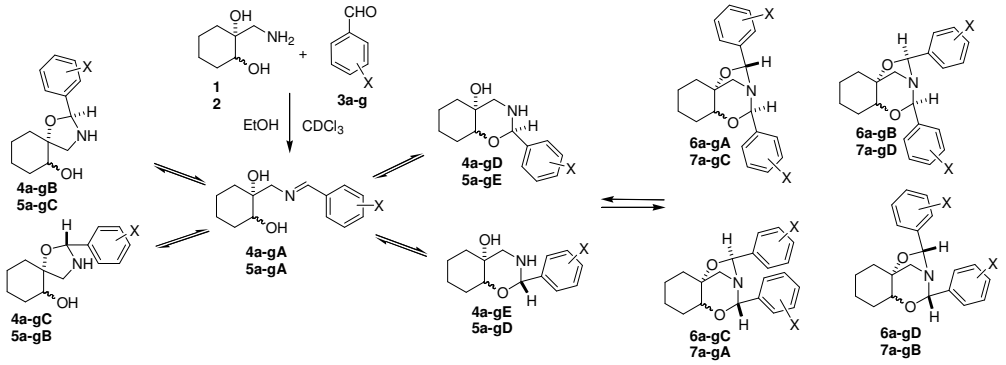
The 1,3-*X,N*-heterocycles are apt to participate in equilibrium processes involving opening of the heterocyclic ring at C2. Electron acceptors increase the relative contribution of the ring form, while electron donors decrease it in favour of the chain form. Electronic effects of the aryl substituent control the ring–chain ratio. These processes include ring–ring epimerization and ring–chain tautomerism.[1] The early works of Kleinpeter *et al*.,[71,72] our own previous studies on 2-aryl-1,3-*X,N*-heterocycles[73] and the recent work of Neuvonen *et al*.[74] turned our attention to the fact that not only the ring–chain equilibrium ratios, but also the ring–ring epimerization equilibrium can be influenced by the 2-aryl substituent.

Our present aim was to study the ring closures of stereoisomeric alicyclic aminodiols and the ring−chain tautomeric equilibria involved in the ring closures of *cis*- and *trans*-1-aminomethylcyclohexane-1,2-diols with various aromatic aldehydes. It was anticipated that both 1,3-oxazines and oxazolidines could be formed under the reaction conditions applied. We also sought an answer to the question of whether the substituent-dependent stability difference between the epimeric forms in the ring–ring epimerization of 2-aryl-1,3-*N,N*-heterocycles can be explained by the stereoelectronic interactions.

3.2. Results and Discussion

Ring–chain tautomerism. The reactions of the *cis*- and *trans*-aminodiols **1** and **2** with seven aromatic aldehydes **3a–g** in EtOH resulted in complex mixtures (Scheme 2). These involve five-component equilibria with the participation of the two epimeric spiro-oxazolidines (**4a–gB** and **4a–gC** for **1**, and **5a–gB** and **5a–gC** for **2**), the two epimeric condensed 1,3-oxazines (**4a–gD** and **4a–gE** for **1**, and **5a–gD** and **5a–gE** for **2**) and the corresponding Schiff bases (**4a–gA** for **1**, and **5a–gA** for **2**) and tricyclic compounds (**6** or **7**).

SCHEME 2



a, *p*-NO_2; **b**, *m*-Cl; **c**, *p*-Cl; **d**, H; **e**, *p*-Me; **f**, *p*-OMe; **g**, *p*-NMe_2

The relative configurations of the various structures were determined from the NOESY spectra via the observation of NOEs between the O-C*H*Ar-N hydrogen and the O-C*H* hydrogen of the ring forms (**4C**, **4E**, **5C** and **5E**) or one of the O-C*H*Ar-N hydrogens of the tricyclic forms (**6C**, **6D**, **7C** and **7D**). Interestingly, the conformation of the major component **4E** or **5E** was not forthcoming since the C-CH_2-NH methylene protons form a degenerate singlet and they therefore do not provide any information regarding possible coupling constants or NOE contacts and hence stereochemical relationships.

For the tautomeric study of these compounds, the samples were dissolved in $CDCl_3$, and, as illustrative examples, phenyl derivatives **4d** and **5d** are discussed in detail. Although the spectrum of the dissolved sample **4d** (or **5d**) was quite complex, structures could readily be assigned due to the well-separated singlets from the azomethine hydrogen resonating at 8.21 ppm (or 8.29 ppm) and the O-C*H*Ar-N hydrogens resonating in the region 5.25–6.18 ppm (or 5.17–5.75 ppm). By the application of various selective 1D and standard 2D homo- and heteronuclear experiments, the NMR analysis revealed that the signals in the region 5.25–6.18 ppm (or 5.17–5.75 ppm) are those of 1,3-oxazine **4dE** (or **5dE**) and the tricyclic compounds **6dB** and **6dC** (or **7dC**). Initially upon dissolution, the tricyclic compounds, the condensed 1,3-oxazine **4dE** (or **5dE**) and the corresponding Schiff base were the major components; the other condensed 1,3-oxazine **4dD** (or **5dD**) and the two epimeric spiro-

oxazolidines were not formed initially. A period of 2 months was found to be adequate for the establishment of equilibrium, resulting in an increase in concentration of the initially minor compounds at the expense of the tricyclic compounds, which were reduced in concentration by a factor of 10.

To test the applicability of the Hammett–Brown linear free energy equation, regression analysis was performed on plots of log *K* versus σ^+ for compounds **4B–D** and **5B–D** (Figures 2 and 3). The tautomeric ratios are based on the integration of the **4B–E**, **5B–E**, **6A–D** and **7A–D** ring form O-C*H*Ar-N protons and the **4A** and **5A** chain form N=CH proton singlets.

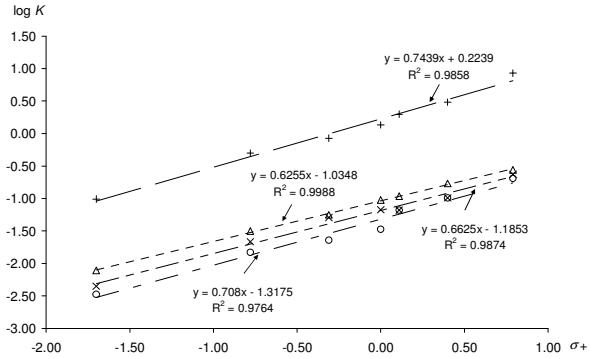


FIGURE 2. Plots of log K_x versus σ^+ for compounds **4a–g**: **B** (x), **C** (Δ), **D** (o), **E** (+).

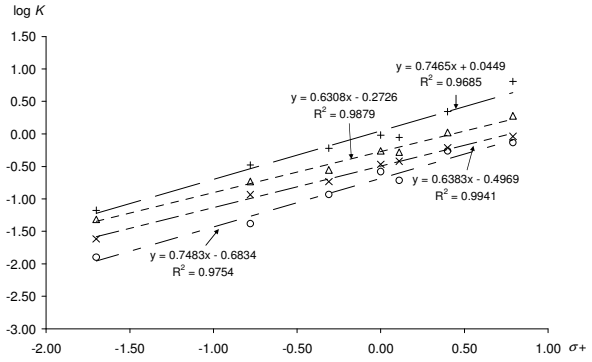


FIGURE 3. Plots of log K_x versus σ^+ for compounds **5a–g**: **B** (x), **C** (Δ), **D** (o), **E** (+).

Tables 1 and 2 list the chemical shifts of these indicator protons for compounds **4**–**7** and Tables 3 and 4 the relative contributions of the compounds to the multicomponent equilibria. Exact thermodynamic comparisons for the condensed 1,3-oxazines and the spiro-oxazolidines **B–E** were performed[7] separately for each of the ring-closed isomers. The slopes for the condensed 1,3-oxazines were determined to lie in the range 0.72–0.75, and those for the spiro-oxazolidines in the range 0.62–0.65, in close accord with those described in the literature.[1] Since the plots are linear and the regression coefficients are all higher than 0.96, the systems are deemed to be adequately described by the Hammett–Brown linear free energy

equation. An attempt was also made to fit the same relationship to the tricyclic compounds **6B**, **6C** and **7C**, but the resulting regression coefficients were all generally less than 0.95, using few experimental points. As regards the mole fractions within the multi-component equilibria (Tables 3 and 4), the highest amounts of tricyclic compounds were observed for **6dC** (i.e. when benzaldehyde was used) and **7bC** (i.e. when *m*-chlorobenzaldehyde was used), in their respective series.

TABLE 1. Chemical shifts (ppm) of the indicator protons[a] for compounds **4** and **6** in $CDCl_3$

	4A	4B	4C	4D	4E	6A	6B	6C	6D
a	8.33	5.74	5.55	5.67	5.32	–	–	–	–
b	8.14	5.47	5.42	5.57	5.19	–	5.79, 6.08	5.47, 5.57	–
c	8.18	5.59	5.45	5.61	5.21	–	–	5.50, 5.57	–
d	8.21	5.58	5.48	5.66	5.25	–	5.88, 6.18	5.56, 5.71	–
e	8.16	5.54	5.45	5.63	5.21	–	5.84, 6.15	5.51, 5.68	–
f	8.12	5.53	5.43	5.62	5.19	–	5.82, 6.13	5.50, 5.68	–
g	8.04	5.49	5.41	5.63	5.17	–	–	–	–

[a] For **4A**, N=CH; all others, O–C*H*Ar–N.

TABLE 2. Chemical shifts (ppm) of the indicator protons[a] for compounds **5** and **7** in $CDCl_3$

	5A	5B	5C	5D	5E	7A	7B	7C	7D
a	8.41	5.72	5.61	5.63	5.25	–	–	5.59, 5.68	–
b	8.25	5.62	5.45	5.47	5.14	–	–	5.58, 5.62	–
c	8.29	5.62	5.49	5.51	5.17	–	5.58, 5.61	5.65, 5.76	–
d	8.29	5.62	5.49	5.51	5.17	–	–	5.65, 5.75	–
e	8.17	5.57	5.47	5.51	5.22	–	–	5.52, 5.68	–
f	8.29	5.57	5.44	5.47	5.13	–	–	5.59, 5.72	–
g	8.10	5.53	5.42	5.17	5.10	–	–	–	–

[a] For **5A**, N=CH; all others, O–C*H*Ar–N.

TABLE 3. Tautomeric compositions (%) at equilibrium for compounds **4** and **6** in $CDCl_3$ at 300 K

	σ^+	4A	4B	4C	4D	4E	6A	6B	6C	6D
a	0.79	9.8	2.3	2.7	2.0	83.3	0.0	0.0	0.0	0.0
b	0.40	8.8	0.9	1.5	0.9	26.8	0.0	14.3	46.8	0.0
c	0.11	25.7	1.7	2.8	1.7	50.9	0.0	0.0	17.1	0.0
d	0.00	17.9	1.2	1.7	0.6	24.2	0.0	3.2	51.2	0.0
e	-0.31	39.3	2.0	2.2	0.9	33.3	0.0	2.3	20.1	0.0
f	-0.78	60.4	1.3	1.9	0.9	30.2	0.0	0.5	5.0	0.0
g	-1.70	89.8	0.4	0.7	0.3	8.8	0.0	0.0	0.0	0.0

TABLE 4. Tautomeric compositions (%) at equilibrium for compounds **5** and **7** in $CDCl_3$ at 300 K

	σ^+	5A	5B	5C	5D	5E	7A	7B	7C	7D
a	0.79	8.3	7.6	15.7	6.1	53.4	0.0	0.0	8.9	0.0
b	0.40	10.0	6.1	10.5	5.5	22.0	0.0	0.0	46.0	0.0
c	0.11	14.5	5.5	7.6	2.8	12.8	0.0	14.7	42.1	0.0
d	0.00	19.4	6.6	10.6	5.1	18.6	0.0	0.0	39.7	0.0
e	-0.31	40.2	7.4	11.2	4.7	24.2	0.0	0.0	12.3	0.0
f	-0.78	57.8	6.7	10.7	2.4	19.1	0.0	0.0	3.4	0.0
g	-1.70	86.8	2.1	4.2	1.1	5.8	0.0	0.0	0.0	0.0

Reaction Pathway for the Tricyclic Compounds. As concerns the possible reaction pathways for the formation of the tricyclic compounds **6** (or **7**), three hypotheses were postulated. On the basis of the well-known literature finding that oxazolidines and tetrahydro-1,3-oxazines can be used as aldehyde sources,[56-59] the tricyclic forms **6** (or **7**) could be formed

by the reaction between two condensed 1,3-oxazines **4E** (or **5E**), or between the condensed 1,3-oxazine **4E** (or **5E**) and the Schiff base **4A** (or **5A**), or between the condensed 1,3-oxazine **4E** (or **5E**) and an aldehyde **3a**–**g** in the five-component tautomeric system. To select the most appropriate of these different hypotheses, the following experiments were performed.

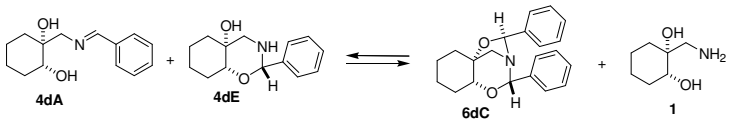
First, 2D EXSY measurements were carried out in CD_3OD solutions of the phenyl-substituted systems **4dA–E** and **6dA–D** (or **5dA–E** and **7dA–D**), in order to detect direct magnetization transfers between the indicator protons of tricyclic compounds **6dC** (or **7dC**) and the indicator protons in question of the compounds in the different reaction pathways. First-order chemical exchange cross-peaks were observed between the well-separated N=C*H*Ar proton singlet of the corresponding Schiff base **4dA** (or **5dA**) and one of the O-C*H*Ar-N protons of the tricyclic compound **6dC** (or **7dC**), and also between the other O-C*H*Ar-N proton of **6dC** (or **7dC**) and the indicator proton of the condensed 1,3-oxazine **4dE** (or **5dE**).

However, a reaction pathway involving a slower addition of the condensed 1,3-oxazine to aldehyde cannot be ruled out with the above method, if the reaction rate is too low as compared to the NMR time scale. If the tricyclic forms **6** (or **7**) are formed by the reaction between the condensed 1,3-oxazine and an aldehyde, the pure tricyclic molecules after dissolution decompose under the reaction conditions applied to form the multicomponent equilibrium. On the other hand, the reaction between two condensed 1,3-oxazines **4E** (or **5E**) or between the condensed 1,3-oxazine **4E** (or **5E**) and the Schiff base **4A** (or **5A**) involve an aminodiol elimination. Assuming the microscopic reversibility of the reaction, the aminodiol is necessary in the first step of the decomposition of the pure tricyclic compounds. In order to test the stability of the pure forms and the effect of the aminodiol in solution, the pure tricyclic products were prepared. Preparation of the major isomeric tricyclic compounds **6aC** (or **7aC**) was accomplished by reacting the appropriate aminodiol **1** (or **2**) with 3 equivalents of *p*-nitrobenzaldehyde in EtOH under reflux. After evaporation of the solvent and recrystallization of the resulting crude product, pure samples of **6aC** (or **7aC**) were isolated in each case. After dissolution in $CDCl_3$, the pure *p*-nitrophenyl-substituted **6aC** (or **7aC**) did not decompose during an observation period of 3 months. When the aminodiol was added to the solution, a multicomponent equilibrium formed immediately.

These results indicate that the reaction pathway for the formation of the tricyclic compounds is via the reaction between the condensed 1,3-oxazine and the Schiff base in the five-component ring–chain tautomeric system. For the examined system, this means that the tricyclic compound **6dC** (or **7dC**) was formed from the Schiff base **4dA** (or **5dA**) and the

condensed 1,3-oxazine **4dE** (or **5dE**) by aldehyde transfer and aminodiol elimination (Scheme 3).

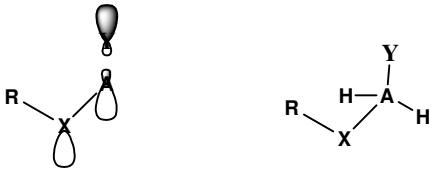
SCHEME 3



To answer the question of whether the substituent-dependent stability difference can be explained by the stereoelectronic interactions, the epimerization equilibria of 2-aryl-1,3-*N,N*-heterocycles were used.

The generalized anomeric effect (AE) is recognized in an Lp-X-A-Y moiety as a preference for an antiperiplanar arrangement of the lone pair (Lp) and the A-Y bond, where X is a heteroatom, A is an element with intermediate electronegativity (EN) and Y is a group with higher EN.[75,76] During the anomeric effect, interaction is observed between a lone pair n and a neighboring antibonding σ^* orbital (Scheme 4).

SCHEME 4

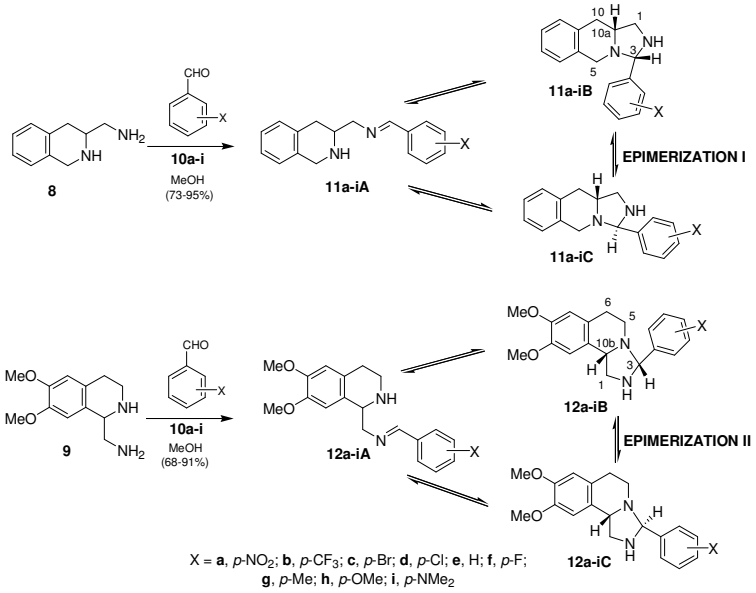


X=N, O, S, or Hal

It is generally accepted that the stability difference observed between two specific geometries of the same molecular moiety stems from two-electron/two-orbital hyperconjugative interactions resulting in an excess stabilization energy.[77,78] There are four major factors that influence the excess of the electron delocalization energy: (i) the EN difference along the A-Y bond, (ii) the acceptor orbital energy level, (iii) the nature of the atom X, and (iv) the Lp-X-A-Y dihedral angles observed in the compared structures. Handling the stability difference as a function of the acceptor ability of the A-Y bond and that of the Lp-X-A-Y dihedral angle allows a further generalization of the concept. This approach has resulted in a number of studies involving the use of various functional groups as Y and elements with intermediate EN as A.[79-81] Despite the difficulties in the detection of the low energy differences in systems containing carbon as Y (*e.g.* -COR, -COOR or -Ar),[82-84] it has been revealed that the stability difference between the C2 epimers in the epimerization equilibrium of 2-aryl-*trans*-1,3-*X,S*-decalins is affected by the electronic properties of the aromatic substituent.[71,72] As regards the local geometry, observations of AE have been reported in conformations other than antiperiplanar.[85]

Ring–ring epimerization. The condensations of diamines **3** and **8** with equivalent amounts of nine aromatic aldehydes in $CDCl_3$ at 300 K resulted in three-component tautomeric mixtures **11A-C** and **12A-C** (two C-3 epimeric imidazoisoquinolines and the corresponding Schiff base) (Scheme 5), containing angularly and linearly condensed imidazoisoquinolines in epimerization equilibria.

SCHEME 5



After the attainment of equilibrium, the spectra of compounds **11a-i** or **12a-i** contained well-separated singlets resonating from the azomethine group at 8.31-8.46 ppm or 8.23-8.44 ppm and the two N-C*H*Ar-NH hydrogens in the region 4.11-4.63 ppm or 4.22-4.92 ppm. With the aim of an unambiguous signal assignment, standard 2D homo- and heteronuclear spectra were recorded. The proportions K_X of the chain (**11A, 12A**) and diastereomeric ring forms (**11B**, **11C** and **12B**, **12C**) of the tautomeric equilibria **11a-i** and **12a-i** were determined by integration of the well-separated N=C*H*Ar (chain) and N-C*H*Ar-NH (ring) proton singlets in the ^{1}H NMR spectra in $CDCl_3$ at 300 K (Table 5).

TABLE 5. Proportions (%) of tautomeric forms in tautomeric equilibria for compounds **11**and **12**.

Compd.	σ^+	**11A**	**11B**	**11C**	**12A**	**12B**	**12C**
a	0.79	3.4	88.7	7.9	4.1	25.8	70.1
b	0.612	4.8	88.2	7.0	4.7	32.0	63.3
c	0.15	7.6	86.1	6.3	14.7	28.9	56.4
d	0.114	6.9	88.0	5.1	8.1	38.1	53.8
e	0	9.0	86.3	4.7	12.8	28.1	59.1
f	-0.073	9.6	85.4	5.0	9.9	38.6	51.5
g	-0.311	11.6	85.0	3.4	14.3	28.9	56.8
h	-0.778	14.3	82.5	3.2	24.3	29.5	46.2
i	-1.7	34.6	63.7	1.7	57.0	14.6	28.4

LFER analysis of the reaction free energies (Scheme 2, calculated at 300 K) demonstrates an acceptable linear correlation with the Hammett–Brown substituent constants (σ^+) for epimerization I (K_{11} = **11C/11B**), while the linear fit falls off for epimerization II (K_{12} = **12C/12B**) (Figure 4).

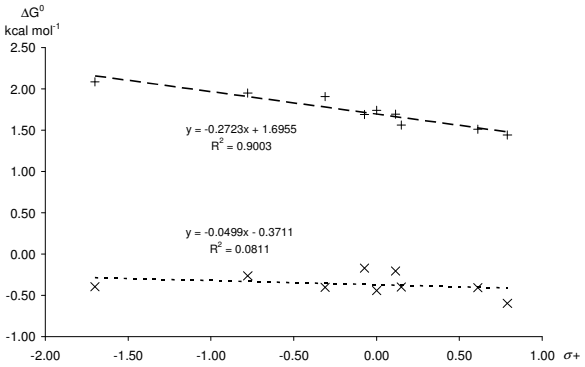


FIGURE 4. Plots of ΔG^0 *vs.* Hammett–Brown σ^+ values for epimerizations I ('+') and II ('x').

The conformationally restrained condensed ring systems **11B** and **11C** display significant substituent effects on the epimerization equilibrium, but the angularly condensed imidazoisoquinolines **12B** and **12C** do not exhibit net substituent effects on the relative stabilities. Our conclusion at this point is that the selected models are suitable for detailed analysis of the stereoelectronic interactions probably responsible for the substituent-dependent AE.

Conformational Analysis. Since the stereoelectronic interactions are highly dependent on the geometry of the studied molecules, thorough conformational analysis was performed. Our goal was to determine the predominant geometry for all the models and to identify the possible minor conformers influencing the correlation between the calculated and the measured data. The resulting structures proved to be rigid, since no minor conformation was found within the 6 kcal/mol energy window. The final conformations for **11eB**, **11eC**, **12eB** and **12eC** are shown in Figure 5.

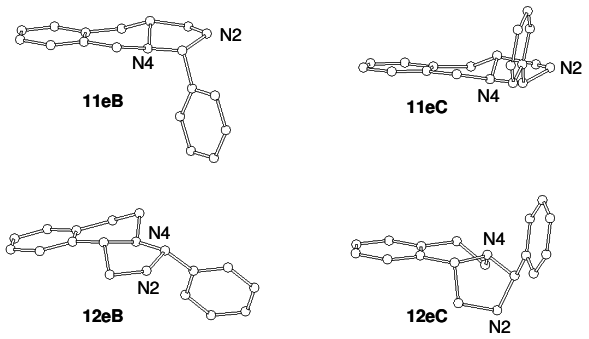


FIGURE 5. Final predominant minimum energy molecular structures for **11eB**, **11eC**, **12eB** and **12eC**, obtained by using *ab initio* HF/3-21G* calculations.

The conformations obtained are supported by the NOE patterns determined via the NOESY spectra. It is clear from the structures that the ideal, antiperiplanar arrangement cannot be found for the nitrogen lone pairs and the possible endocyclic and exocyclic substituents on C3.

NBO Analysis. In the present model compounds, each nitrogen lone pair can overlap with three different vicinal antibonding orbitals associated with C3: $\sigma^*_{C3\text{-}N}$, $\sigma^*_{C3\text{-}Ar}$ and $\sigma^*_{C3\text{-}H}$. (From here, the exact numbering of the model ring system followed.) Since we were interested in the 3-aryl substituent-dependent stereoelectronic interactions, six orbital overlaps around C3 were taken into consideration in our analysis. The results of the NBO calculations revealed insignificant amounts of stabilization energy for n_{N4}-$\sigma^*_{C3\text{-}N}$ and n_{N2}-$\sigma^*_{C3\text{-}H}$ in **11B**, for n_{N4}-$\sigma^*_{C3\text{-}N}$, n_{N4}-$\sigma^*_{C3\text{-}H}$, n_{N2}-$\sigma^*_{C3\text{-}N}$ and n_{N2}-$\sigma^*_{C3\text{-}H}$ in **11C**, for n_{N4}-$\sigma^*_{C3\text{-}N}$ and n_{N2}-$\sigma^*_{C3\text{-}H}$ in **12B**, and for n_{N4}-$\sigma^*_{C3\text{-}H}$ and n_{N2}-$\sigma^*_{C3\text{-}H}$ in **12C**, which was accompanied by negligible variations. The signed sum of the E_{ij} values gives the theoretical hyperconjugative stabilization energy $\Delta\Sigma E_{ij}$ in equation 30, where *r* and *l* designate the right-hand side and the left-hand side, respectively, of the epimerization equilibria.

$$\Delta\Sigma E_{ij} = \Sigma E^{r}_{ij} - \Sigma E^{l}_{ij} \tag{30}$$

To correlate the experimental stability differences with the theoretical results, linear regression analysis was carried out with $\Delta\Sigma E_{ij}$ against ΔG^0 (Figures 6 and 7).

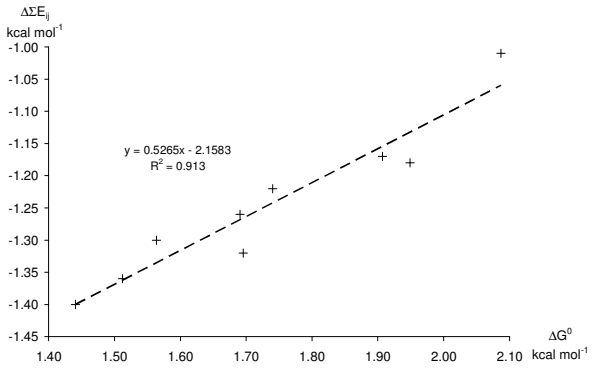


FIGURE 6. Plots of theoretical stereoelectronic stabilization energy $\Delta\Sigma E_{ij}$ *vs.* ΔG^0 for epimerization I.

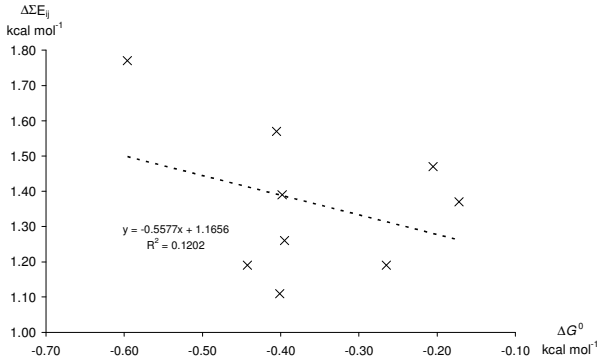


FIGURE 7. Plots of theoretical stereoelectronic stabilization energy $\Delta\Sigma E_{ij}$ *vs.* ΔG^0 for epimerization II.

For epimerization I we obtained a good linear relationship, while in the case of epimerization II there was no significant correlation. The variances of ΔG^0 and $\Delta\Sigma E_{ij}$ were of the same magnitude and the linear correlations were acceptable for epimerization I. These findings suggest that the observed 3-aryl substituent-dependent stability difference is stereoelectronic in origin. The lack of any linear relationship for epimerization II indicates that the variations in ΔG^0 cannot be explained by an electron delocalization energy excess. The correlation between $\Delta\Sigma E_{ij}$ and σ^+ exhibits a similar pattern (Figure 8). The regression analysis for epimerization I gave a linear fit, while $\Delta\Sigma E_{ij}$ for epimerization II proved independent of σ^+, which is in good agreement with the experimental observations.

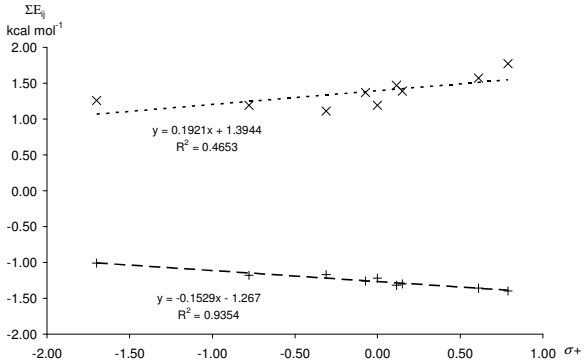


FIGURE 8. Plots of theoretical stereoelectronic stabilization energy $\Delta\Sigma E_{ij}$ *vs.* Hammett–Brown σ^+ values for epimerizations I ('+') and II ('x').

It is clear that the experimentally detected substituent-dependent stability difference for epimerization I is connected with the AE comprising all the possible n–σ^* interactions around C3. For epimerization II, the electronic properties of the 3-aryl substituent do not seem to exert a detectable effect either on the experimental reaction free energy or on the calculated stereoelectronic stabilization energy. These results raise two questions: (i) Are the electron delocalization energies of the individual orbital overlaps also substituent-dependent? (ii) What specific NBO donor-acceptor interactions are responsible for the substituent-dependent stabilization?

The Hammett–Brown substituent dependence of the E_{ij} values was tested by univariate linear regression for both epimerization equilibria. The calculation was carried out for each NBO donor-acceptor interaction. All the linear fits exceeding a significance limit of 99% are given.

The results demonstrate that, regardless of the nature of the antibonding orbitals associated with C3, the 3-aryl substituent can have a significant influence on the two orbital electron delocalization energies, even in compounds **12B** and **12C**. It is also seen that the σ^+ values are in an acceptably linear correlation with the E_{ij} values, despite the fact that the

interacting orbitals are not in a periplanar conformation. The evaluation of the slope of the linear models (ρ) demonstrates that the 3-aryl substituents have a strong positive influence on the n_N–$\sigma^*_{C3\text{-}Ar}$ interaction energies, i.e. the higher the value of σ^+, the higher the absolute value of the stabilization energy. It is interesting that the interactions n_N–$\sigma^*_{C3\text{-}N}$ and n_N–$\sigma^*_{C3\text{-}H}$ change in an opposite sense: electron-withdrawing substituents decrease the stabilizating delocalization energy. The n_N–$\sigma^*_{C3\text{-}H}$ overlaps are subjected to a weak modulation by σ^+ as compared with those obtained for n_N–$\sigma^*_{C3\text{-}Ar}$.

Our observations reveal that all the possible vicinal n–σ^* interactions around C3 that have a conformational arrangement other than orthogonal may contribute to the substituent-dependent AE (Figure 9).

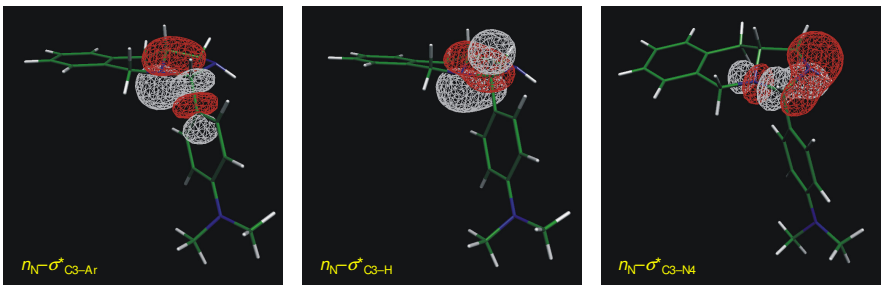
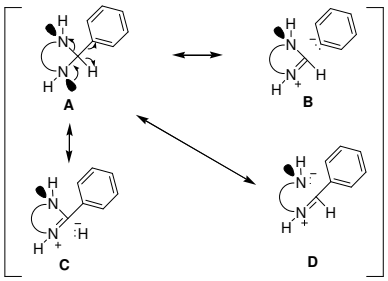


FIGURE 9. Graphical representations of the selected NBO interactions contributing to the substituent-dependent stereoelectronic stabilization. The orbital overlaps are displayed for **11iB**.

Translating the results of the NBO analysis to the language of valence bond theory gives an intuitive picture of three types of double-bond–no-bond resonance structures responsible for the phenomenological behaviour of the studied models (Scheme 6).

SCHEME 6



The growing electron demand of the 3-aryl substituent increases the stabilizing contribution of the resonance structure **B**, while the weights of the negative hyperconjugations **C** and **D** taper off. Depending on the exact geometry of the species in the epimerization equilibrium, the influence of the substituent on the donor-acceptor interactions can result overall in an experimentally observable substituent-dependence of the equilibrium ratios (epimerization I), which may be interpreted as a fine-tuned AE. It is also possible that the

individual hyperconjugative contributions cancel out, leading to an equilibrium ratio that is independent of σ^+ (epimerization II).

Equation 27 contains two important parameters determining the E_{ij} values: the reciprocal donor-acceptor orbital energy difference ($\Delta\varepsilon = \varepsilon_j - \varepsilon_i$) and the square of the NBO Fock matrix off-diagonal element (F_{ij}). $\Delta\varepsilon^{-1}$ changes in parallel with σ^+ for all types of acceptor orbitals, and the relative increase in $\Delta\varepsilon^{-1}$ on going from *p*-NMe_2 to *p*-NO_2 is only a few per cent (Figure 10).

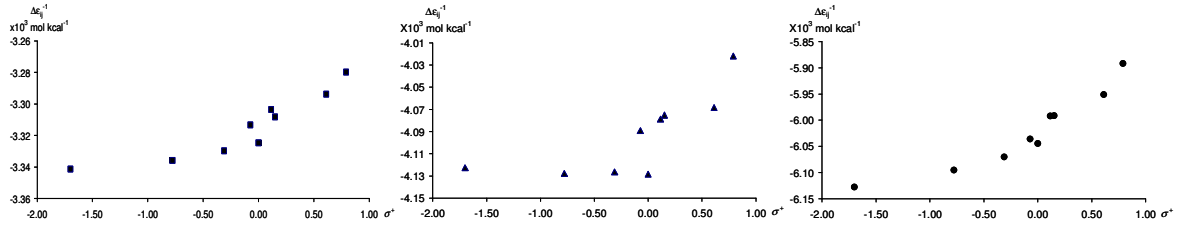


FIGURE 10. Plots of selected reciprocal donor-acceptor orbital energy difference ($\Delta\varepsilon^{-1}$) in **11B**, obtained from NBO calculation. The symbols '■', '▲' and '●' designate the orbital overlaps n_{N2}–$\sigma^*_{C3–N4}$, n_{N2}–$\sigma^*_{C3–Ar}$ and n_{N2}–$\sigma^*_{C3–H}$, respectively.

This is in accordance with the picture of decreasing acceptor orbital energy due to the increasing electron-withdrawing effect of the 3-aryl substituent. Inspection of the substituent dependence of F_{ij}^2 (Figure 11) allows the conclusion that the changes in E_{ij} are mainly explained by the variations in F_{ij}, since the changes are higher as compared with those in $\Delta\varepsilon^{-1}$, and the opposite trend of the correlation with σ^+ can also be detected. We note that the Fock matrix off-diagonal element for the n_N–$\sigma^*_{C3–H}$ overlap does not exhibit a good linear fit against σ^+.

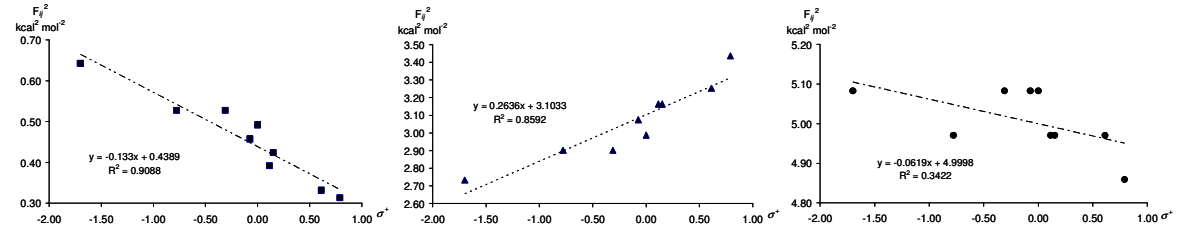


FIGURE 11. Plots of selected squared Fock matrix off-diagonal elements (F_{ij}^2) in **11B**, obtained from NBO calculations. The symbols '■', '▲' and '●'designate the orbital overlaps n_{N2}–$\sigma^*_{C3–N4}$, n_{N2}–$\sigma^*_{C3–Ar}$ and n_{N2}–$\sigma^*_{C3–H}$, respectively.

F_{ij} is strongly correlated with the orbital overlap matrix element S_{ij}, which finally boils down to a dihedral angle and acceptor orbital polarization dependence.[35] The substituent-dependent conformational changes for these model compounds are around 1°, i.e. negligible. However, the polarization coefficients of the C3 sp^3 atomic orbitals increase in $\sigma^*_{C3–Ar}$ and

decrease in the acceptor bonds $\sigma^*_{C3–N}$ and $\sigma^*_{C3–H}$ when the electron demand of the 3-aryl substituent is augmented (Figure 12).

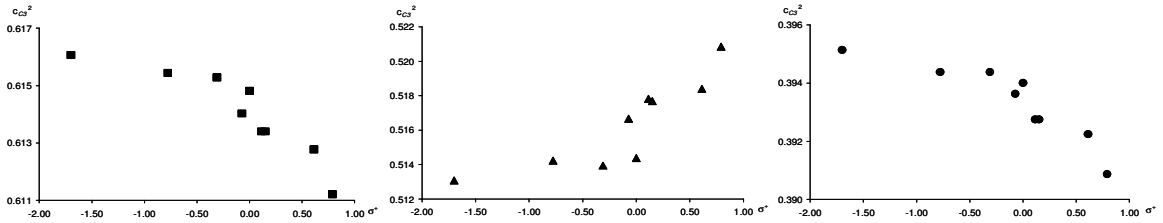


FIGURE 12. Substituent-dependence of the polarization coefficients in the C3-associated antibonding orbitals in **11B**. The symbols '■', '▲' and '●'designate the orbital overlaps $\sigma^*_{C3–N4}$, $\sigma^*_{C3–Ar}$ and $\sigma^*_{C3–H}$, respectively.

To summarize, the 3-aryl substituent effect contracts or expands the antibonding orbital in the direction of the donor lone pairs (Figure 13), which accounts exactly for the opposite trends of the variations in F_{ij}, and thus for the observed substituent dependence of the individual orbital overlap stabilization energies.

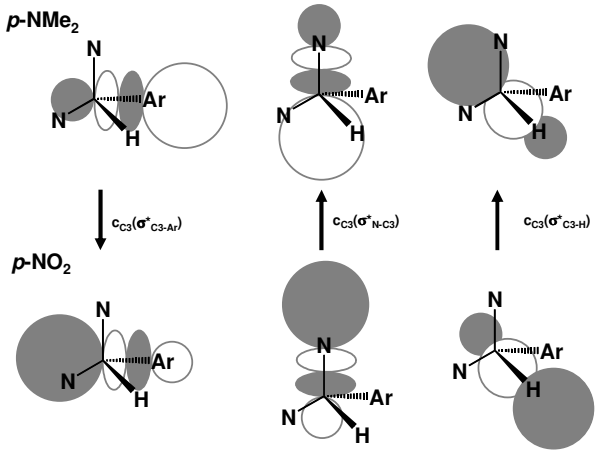


FIGURE 13. Graphical interpretation of the substituent dependence of the polarization coefficients on C3. On going from *p*-NMe_2 to *p*-NO_2, the C3–Ar antibonding orbital expands, while the C3–N and C3–H antibonding orbitals contract at C3.

3.3. Conclusions

Condensation of aminodiol **1** or **2** with one equivalent of an aromatic aldehyde, **3a–g**, resulted in a complex, multicomponent equilibrium mixture consisting of a five-component, ring–chain tautomeric system, involving the Schiff base, two epimeric spiro-oxazolidines and two epimeric condensed 1,3-oxazines, and some tricyclic compounds. The reaction pathway for the formation of the tricyclic compounds is via the reaction between one 1,3-oxazine and one Schiff base in the five-component, ring–chain tautomeric system. The five-component, ring–chain tautomeric system was found to be adequately described by the Hammett–Brown linear free energy equation.

An extended view of the stereoelectronic interactions affords a detailed description of the substituent-dependent stability changes exhibited by the 2-aryl-1,3-*X,N*-heterocycles. The electron-withdrawing properties of the 2-aryl substituent alter the polarization along all the single bonds associated with C2; it therefore changes the extents of the orbital overlaps between the nitrogen lone pairs and the antibonding orbitals. The donor-acceptor hyperconjugative stabilization energies display an acceptable linear correlation with the Hammett–Brown substituent constant, and the theoretical stereoelectronic stabilization energies calculated as the sums of the individual orbital overlap electron delocalization energies explain the experimentally observed substituent-dependent AE. Application of the results presented in this work may help resolve the question of the substituent-dependence of the epimerization equilibrium constants in the field of ring–chain tautomerism. In accord with the currently increasing interest in dynamic covalent bonds allowing the construction of supramolecular structures and combinatorial libraries under equilibrium conditions,[86] the condensed 2-aryl-1,3-*X,N*-heterocycle derivatives were recently recognized as useful building blocks for the synthesis of novel supramolecular host compounds and dynamic combinatorial libraries. It therefore appears important to provide explanations as to the exact electronic mechanisms of the effects exerted by the substituents and the stereochemistry on the equilibrium processes of these compounds.

4. Self-organization of Biomimicking β-Peptides

4.1. Preliminaries

The artificial secondary structures of β-peptides are of major importance in the field of self-organizing systems by virtue of the wide range of their potential applications, due to their propensity to adopt side-chain-controllable compact ordered conformations.[3,16-20] Homo-oligomers constructed by using *trans*-2-aminocyclohexanecarboxylic acid (*trans*-ACHC) monomers, which have protecting groups at both ends, **16**-**19** (Scheme 5), are known to form a highly stable 14-helix in various solvents (methanol or pyridine).[10,14,87-93]

This structural motif is so stable that it is always predominant when the chain is longer than three monomers. The torsional angles φ, θ and ψ of the β-residues extracted from the literature data[16,18,94-96] for the 10- and the 14-helices with the same helicity occupy adjacent positions in the Ramachandran plot, which allows a smooth rearrangement between the two conformational states without unfolding of the helical pattern if the 10-helix is energetically available under the corresponding conditions. Accordingly, by utilizing CD spectra together with a molecular dynamics simulation, Seebach *et al.* proposed an equilibrium between the 10- and 14-helices in β-hexapeptide derivatives with simple proteinogenic side-chains,[97,98] while Fleet *et al.* demonstrated experimentally that an oxetane-based β-residue leads directly to 10-helix formation.[99] Such a conformational polymorphism is an important feature of any folded system that is designed to have a complex dynamic function in general; it is also observed for the natural α-peptides: the interplay between the α-helix and the 3_{10}-helix motif can be a crucial factor during the folding process.[100] In order to reveal similar intrinsic properties of the β-peptides and experimentally capture the conformational polymorphism in high-resolution solution structures, short oligomers with conformationally constrained side-chains without terminal protecting groups were studied in the present work.

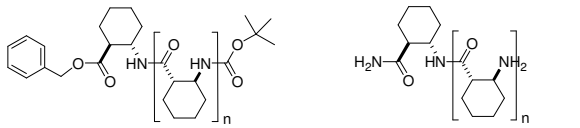
A major challenge in foldamer science is to prove that higher-order structural levels are available for β-peptides and that their formation can be tailored by the β-amino acid sequence. Certain designed β-peptide helices have been observed to exhibit self-association in solution,[93] but distinct tertiary structures with specific morphology on the nanometre scale have not been reported thus far. The construction of tertiary structures via the building of protein-sized β-peptide sequences is a formidable task because of synthetic difficulties. One feasible approach is to design discrete secondary structure units, which tend to self-assemble into tertiary structure motifs exhibiting specific morphology. The building blocks of choice are the homo-oligomers of cyclic β-amino acid residues: *cis*-(1*R*,2*S*)-ACPC$_7$ and *trans*-

(1*S*,2*S*)-ACHC$_6$. They are known to form a self-stabilizing helix and a non-polar strand, resembling the α-helix and the β-strand of the α-peptides, respectively. As a consequence of their unlike secondary structures, essentially different tertiary structures can be anticipated if self-association is intrinsically possible for these β-peptides in solution.

4.2. Results and Discussion

In our approach, the potential energy hypersurfaces of *trans*-ACHC homo-oligomers without protecting groups, **21**-**24** (Scheme 7), were probed first by a standard restrained simulated annealing conformational search protocol, carried out by using molecular mechanics and distance restraints between the relevant H-bond pillar atoms.

SCHEME 7



16, n = 0; **17**, n = 1; **18**, n = 2; **19**, n = 3; **20**, n = 5; **21**, n = 2; **22**, n = 3; **23**, n = 4; **24**, n = 5

The resulting most stable structures for **22**-**24** were optimized without restraints for both the 10- and 14-helices. Final minimizations were carried out by using an *ab initio* quantum mechanical method[50] at the RHF/3-21G level in vacuum, as this has been reported to provide a good approximation for the geometry of the β-peptides.[101,102] The *ab initio* structures converged to the corresponding local minimum of the potential energy surface. In order to take into account the effects of more diffuse basis sets and the electron correlation, the energies were calculated at the B3LYP/6-311G** level too (Figure 14). The conformational energy differences between the 10- and 14-helices ($\Delta E = E_{10}$-E_{14}) proved to be rather low in terms of the non-bonding interaction energy scale, indicating a possible energetic availability of the 10-helix. The ΔE values clearly reveal a trend to an increasing relative stability for the 10-helix motif as the chain length decreases. The energy difference obtained with the different approximation levels for **22** predicts a conformational equilibrium between the 10- and the 14-helix, with a slight preference for the 10-helix in the case of the RHF/3-21G level. These findings presage the tendency for the unprotected tetramer to adopt the 10-helical motif, which is surprising in light of the earlier literature results on similar constrained β-peptides with protecting groups and offers a possibility for experimental testing.

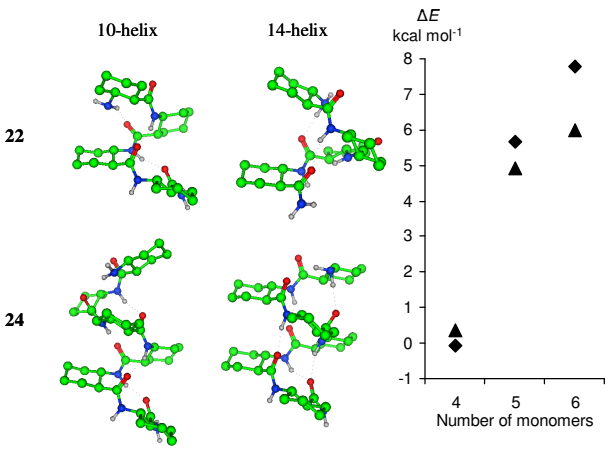


FIGURE 14. Some of the resulting *ab initio* structures, and the chain length-dependence of the conformational energy difference between the 10- and 14-helices. ♦ RHF/3-21G, ▲ B3LYP/6-311G**

For assessment of the conformational stability of the peptides in CD_3OD, ^1H-^2D amide proton exchange was utilized. For **21**, an unfolded conformation was found, as deciphered from the immediate amide proton exchange. Following the dissolution of **22**, ^1H NMR indicated that the amide protons exchanged completely only in 2 weeks. The amide protons in **23** and **24** still showed strong resonances after the same period of time. The shielded amide protons of the residues in the corresponding oligomers were assigned (Figure 15).

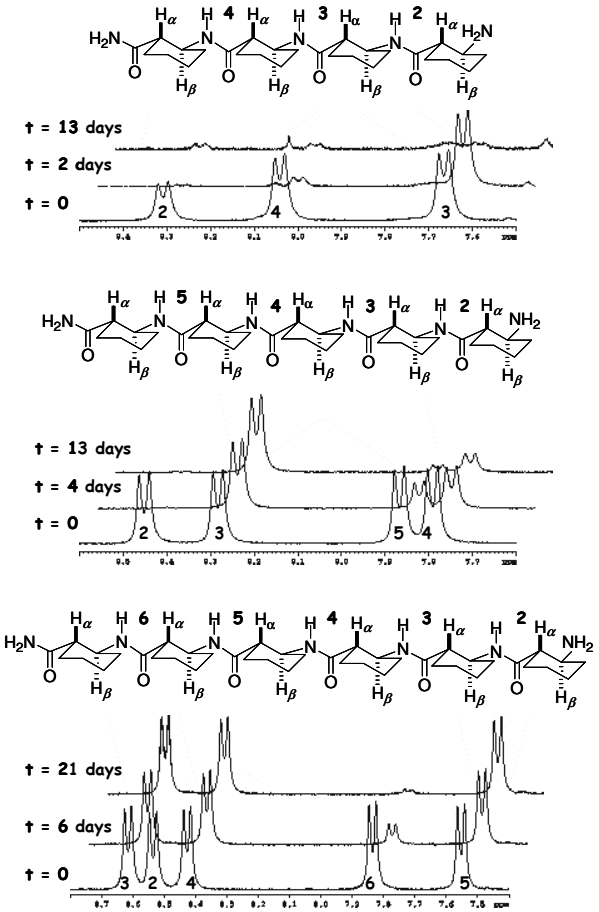


FIGURE 15. NH/HD exchange results for **22**-**24** in CD_3OD.

The observations confirmed that the studied oligomers adopt very stable intramolecular H-bonded folding patterns in CD_3OD solution. As concerns the previously reported *trans*-ACHC oligomers with protecting groups, the amide proton signals were lost within 65 h,[87] which may suggest considerably more shielded NH protons in our models, but a direct comparison must be made with caution because of the incidentally different concentrations of residual acid (TFA) stemming from the chromatographic separation and sample preparation protocol. The ^1H-^2D amide proton exchange experiments were therefore repeated for **24** with an elevated TFA concentration of 1.0 v/v%. The residual signals of the most shielded protons exhibit a relative intensity of 0.2 after 65 h for **24** (Figure 16). The usual TFA concentration during HPLC separations is 0.1 v/v%; hence, the applied TFA level in these ^1H-^2D exchange experiments is drastically higher than one would expect in a normal set-up. Accordingly, these results still support the view of highly shielded and stable H-bonds in the studied compounds.

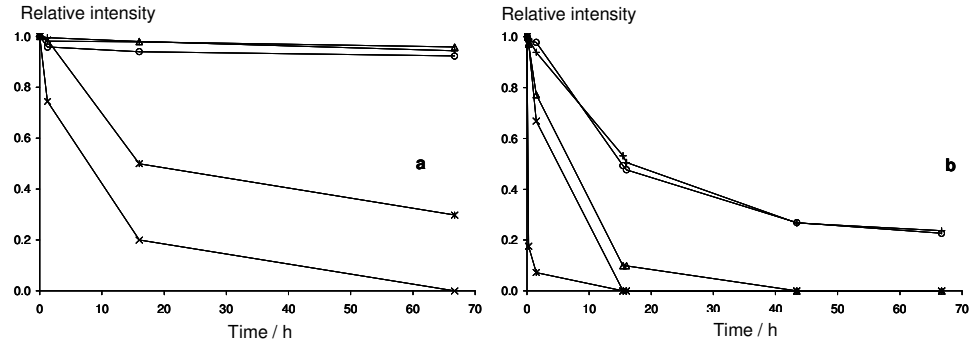


FIGURE 16. Comparison of the NH/ND exchange rate: a) without additional TFA; b) with 1% TFA. The displayed intensities are assigned as follows: x NH_2, o NH_3, + NH_4, Δ NH_5, * NH_6

The high-resolution 3D structure assignment was carried out by running ROESY spectra with 225 ms (Figure 17) and 400 ms spin-lock in CD_3OD. The intensity ratio (R) of the cross-peaks NH_i-$C^{\alpha}H_{i-1}$ and NH_i-$C^{\beta}H_i$ calculated according to equation 24 is a sensitive indicator of the overall fold of the oligomers because it is dependent on the 6th power of the relative distances:[14]

$$R = I_{NOE}(NH_i\text{-}C^{\alpha}H_{i-1})/I_{NOE}(NH_i\text{-}C^{\beta}H_i) \qquad (31)$$

where I_{NOE} designates the integrated and offset-compensated ROESY cross-peak intensities. On the use of secondary structure motifs of 1*S*,2*S*-ACHC obtained by molecular modelling, representative R values are estimated as 3.51, 5.83 and 5.61 for the right-handed 12-helix, left-handed 10-helix and left-handed 14-helix, respectively. The random coil gives an R value of 2.1. It is clear that the 10- and 14-helices can not be distinguished in this way, considering their similar R values and the usual experimental error; nevertheless, the overall fold can be determined.

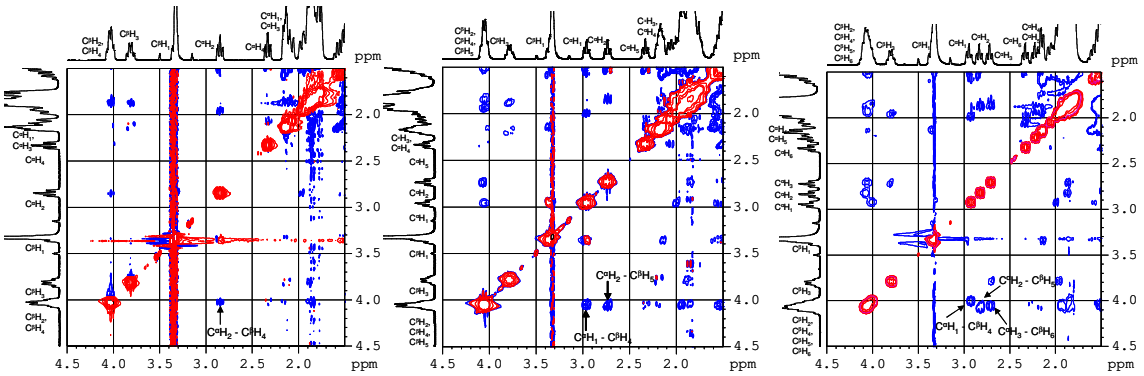


FIGURE 17. Diagnostic cross-peaks in ROESY spectra and signal assignments for a) **22**, b) **23** and c) **24** in CD_3OD, red: positive intensity levels, blue: negative intensity levels.

TABLE 6. *R* values from *ab initio* structures and NMR measurements

	24			**23**			**22**		
i	Calcd. 10-helix	Calcd. 14-helix	Exp.	Calcd. 10-helix	Calcd. 14-helix	Exp.	Calcd. 10-helix	Calcd. 14-helix	Exp.
2	6.1	4.9	3.8	5.3	6.1	3.7	5.8	6.3	0.0
3	7.1	4.9	5.9	5.2	6.3	4.6	5.5	4.4	4.2
4	6.6	4.7	5.7	6.0	6.1	4.8	6.0	6.5	4.2
5	5.7	5.7	6.0	4.8	6.3	6.0			
6	5.9	5.1	7.5						

Comparison of the experimental intensity ratios and the theoretical values reveals good agreement with the left-handed 10- or 14-helical fold for each compound, with the exception of the N-terminal residues (Table 6). For **22**, the NH_2-$C^{\alpha}H_1$ NOE interaction is missing (i.e. $R = 0$, Figure 18), which is a major deviation from the regular helix structure, pointing to a fraying or unusually oriented N-terminal residue. This observation is supported by the higher exchange rate of NH_2 and the considerable upfield shift of $C^{\alpha}H_1$ (2.14 ppm) in comparison with those for **23** (2.96 ppm) and **24** (2.95 ppm). For **23** and **24**, the *R* values involving the first residue are slightly decreased, but there is not a dramatic difference. The antiperiplanar arrangement of the NH_i-$C^{\beta}H_i$ protons in both the 10- and 14-helices was confirmed by the uniform $^3J(NH_i\text{-}C^{\beta}H_i)$ values of ca. 9 Hz.

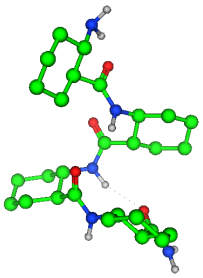


FIGURE 18. Prevailing structure of **22** determined by NMR.

As regards the theoretical models, another major difference between the two helical structures is that the 10-helix should give $C^{\alpha}H_i$-$C^{\beta}H_{i+2}$, and the 14-helix $C^{\alpha}H_i$-$C^{\beta}H_{i+3}$ NOE interactions. For **22**, a marked signal can be observed between the protons at 2.85 ppm, unambiguously assigned to $C^{\alpha}H_2$, and at 4.02 ppm. The resonances of $C^{\beta}H_2$ and $C^{\beta}H_4$ overlap near 4.02 ppm, but analysis of the vicinal coupling constant pattern clearly demonstrates the prevailing *trans*-diaxial orientation of $C^{\beta}H_2$ and $C^{\alpha}H_2$, which rules out their proximity of <2.5 Å, estimated from the cross-peak intensity. The non-overlapping intraresidue $C^{\beta}H_i$-$C^{\alpha}H_i$ NOE signals for residues 3 and 4 are of much lower intensities (25%). The cross-peak in question is therefore assigned to the long-range $C^{\alpha}H_2$-$C^{\beta}H_4$ interaction. The $C^{\alpha}H_1$-$C^{\beta}H_3$ interaction in the ROESY spectrum at the given signal-to-noise ratio is indistinct, in parallel with the irregular N-terminal residue orientation deciphered from the missing NH_2-$C^{\alpha}H_1$ NOE cross-peak. The proximity of the $C^{\alpha}H_2$-$C^{\beta}H_4$ can only be maintained by the cooperation of two H-bonds of CO_4-NH_3 and CO_3-NH_2 forming a stable 10-helix motif involving residues 2-4, while the absence of the $C^{\alpha}H_1$-$C^{\beta}H_3$ and NH_2-$C^{\alpha}H_1$ NOE interactions can be explained by an irregularly oriented N-terminal residue. The rotation of the first residue is not hindered sterically or by hydrophobic stabilization in the 10-helix due to the lack of *i*-(*i*+3) side-chain – side-chain proximity (Figure 18). The preference for the distorted 10-helix motif in **22** can be explained by the fact that only three amide bonds are available for H-bonding stabilization. In the 10-helix conformation, two H-bonds can be formed for such a short chain length, while the 14-helix facilitates only one H-bond. The extra electrostatic stabilization counteracts the conformational strain encountered at the monomer level.[3] In contrast with the ACHC tetramer with an unprotected N-terminal, the ACHC tetramer studied by Gellman *et al.*[88,90] possesses four amide groups. The additional amide NH on the first residue can also participate in a H-bond, furnishing two stabilizing H-bonds in the 14-helix, and here the contribution from the third H-bond can not counteract the conformational strain occurring when the N-terminal residue is also regularly positioned in the 10-helix. These findings suggest that the formation of a stable 14-helix requires at least four amide bonds in the β-peptide chain and the helix nucleation with three residues can start with the 10-helix motif.

For **23** and **24**, all the $C^{\alpha}H_i$-$C^{\beta}H_{i+3}$ interactions could be readily identified (Figure 19). These results strongly suggest that **23** and **24** adopt a left-handed 14-helix pattern, which is in accord with the earlier literature results and proves that removal of the protecting groups does not affect the stability of the 14-helix at this chain length. The lower R values found for the first residues supports a fraying N-terminal, but the presence of the $C^{\alpha}H_1$-$C^{\beta}H_4$ NOE interaction is promoted by the sterically hindered rotation caused by the well-known *i*-(*i*+3) side-chain – side-chain interaction in the 14-helix.

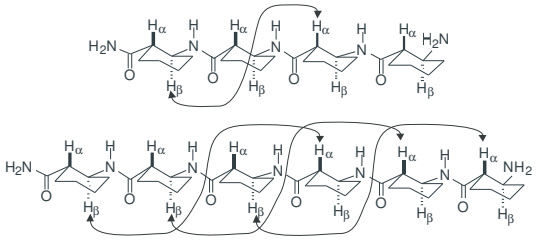


FIGURE 19. The NOE cross-peaks defining long-range $C^{\alpha}H_2$-$C^{\beta}H_4$ interaction for **22** and $C^{\alpha}H_i$-$C^{\beta}H_{i+3}$ interactions for **24**.

The presence of the H-bonds stabilizing the helices was tested by using IR spectroscopy in the intramolecular H-bond-promoting solvent CD_3OH at a concentration of 8 mM (Figure 20). The carbonyl stretching vibrations of the peptide backbone are sensitive markers of the peptide secondary structure, as the vibrational frequency of each C=O bond depends on the H-bonding and the interactions between the amide units, both of which are influenced by the secondary structure. For all the molecules, two characteristic bands were observed, at 1679 cm^{-1} and 1648 cm^{-1}. The high-frequency band was assigned to amide CO groups not involved as acceptors in H-bonding, and the low-frequency band to intramolecular H-bonds. The relative intensity of the 1648 cm^{-1} band increases on going from **22** to **24**, which is in line with the assigned secondary structures. In the amide I region for **22,** a weak shoulder was observed at 1656 cm^{-1}, which may be indicative of a weaker H-bond. This observation can be correlated with the fraying N-terminal of the tetramer and may provide additional evidence of the structural dissimilarity of **22**.

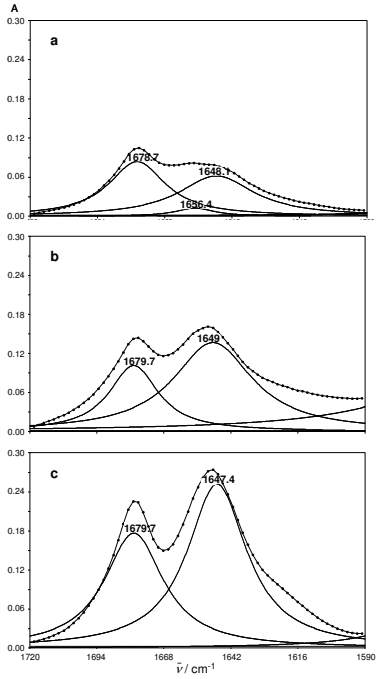


FIGURE 20. Comparison of the IR spectra of *trans*-ACHC oligomers recorded in CD_3OH: a) **22**, b) **23** and c) **24**. The thick curve denotes the measured data, and the thin curves the component bands obtained from non-linear fitting.

A series of publications have revealed that the interpretation of the CD spectra of β-peptides requires special care, as they are extremely sensitive to minor variations in the structure that do not affect the overall fold of the molecule.[88,97,98] The best assignment between the CD features and the 10- and 14-helical patterns appears to be as follows: (i) the 10-helix motif gives rise to a single strong Cotton effect near 205 nm, and (ii) the 14-helix motif contributes with Cotton effects of opposite sign near 200 nm and near 215 nm.[97] We compared the CD spectra for the model compounds (Figure 21).

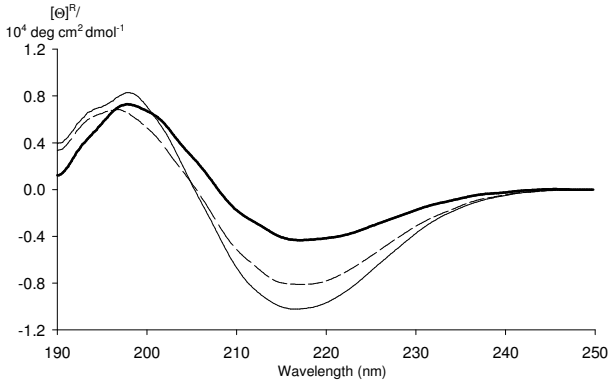


FIGURE 21. Comparison of the CD spectra of *trans*-ACHC oligomers recorded in CD_3OH: thick curve: **22**, dashed curve: **23**, thin curve: **24**. The mean residue ellipticity is depicted.

The data obtained for **22-24** display similar features: a positive Cotton effect at 197 nm and a negative Cotton effect at 217 nm, but there are marked differences in the relative intensities of the Cotton effects for **22** as compared with the longer oligomers. The intensity of the negative Cotton effect at around 217 nm is significantly less for **22**, while the positive Cotton effect, at slightly lower frequency, gives a signal comparable in strength with those for **23** and **24**. The assignment rule of Seebach *et al.*[97,98] allows the considerably lower signal near 217 nm and the strong positive Cotton effect for **22** to be assigned to a conformational ensemble containing the 10-helix motif in significant amount or possibly as the predominant conformation in light of the NMR results. The clear unbiased Cotton effects with opposite sign for **23** and **24** can be assigned to the 14-helix.

The intriguing question of the much higher shielding from the solvent for the NH protons may be explained by the self-association phenomenon on the basis of the early works of Gellman *et al.*[88,93] For characterization of the self-association, diffusion NMR measurements with a stimulated gradient echo and longitudinal eddy current delay were used.[45] Through exploitation of the relationship between the apparent hydrodynamic radius and the diffusion constant, the aggregation numbers (N) were determined by following the literature methods,[45,103] via the following equations:

$$N = \frac{(m_{TMS} / m)^3 \times V_{TMS}}{V_i} \quad (32)$$

$$m = -\frac{\ln(I / I_0)}{k} \quad (33)$$

$$k = (\gamma G\delta)^2(\Delta\text{-}\delta/3) \quad (34)$$

where I is the corresponding NMR signal intensity, G is the gradient strength, V is the molecular volume, and γ is the gyromagnetic ratio, while Δ and δ are the diffusion delay and gradient pulse length parameters for the stimulated gradient echo pulse sequence. The aggregation numbers were determined from three independent measurements with an average error of 0.12. As internal reference, TMS was used. This measurement samples the ensemble average of the molecules; hence, the resulting aggregation number is an average over the aggregated species, giving rise to non-integral values. The diffusion model utilized assumes spherical symmetry for the species, this condition being met for the helix monomers, where both the helix diameter and the helix length are approximately 10 Å. The aggregated species diverge from spherical symmetry, but the trends in the aggregation propensities can be followed reasonably well.

The measured aggregation numbers, 5.91, 6.60 and 7.49, revealed an increasing trend for **22**, **23** and **24**, respectively, in CD_3OD at 303.1 K. The driving force for the aggregation may be a specific head-to-tail H-bond pattern leading to a rod-like superstructure, and/or the solvent-driven side-by-side hydrophobic interactions between the helical units, resulting in helix bundles. Head-to-tail NOE interactions were not detected and a rod-like assembly would not explain the extra NH shielding observed in the exchange experiments; this hypothesis can therefore be eliminated. The temperature-dependent aggregation number measurements furnished a maximum curve (Figure 22), which is characteristic of ionic tensides.[104,105] For this type of molecules, a lower temperature is advantageous for the formation of a stable solvation shell preventing aggregation, while a higher temperature affords a higher disaggregating kinetic energy. This result supports solvent-driven helix bundle formation, which can effectively increase the shielding of the NH protons from the solvent. The self-association does not affect the secondary structure stability, as constant helical patters were observed at the disaggregating temperature of 310 K.

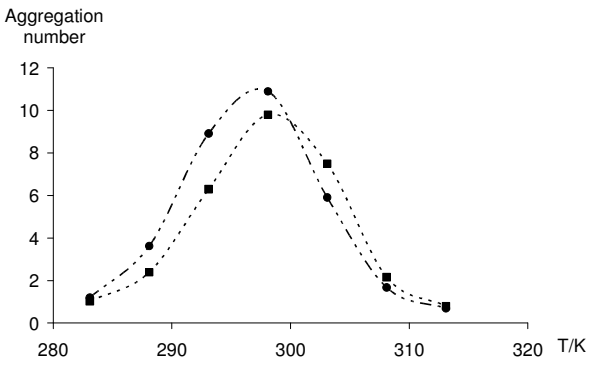


FIGURE 22. Temperature dependence of the aggregation numbers: • for **22**, and ■ for **24**.

From earlier results, it is known that the change in the relative configuration of the cyclic *β*-petides residues (*cis* or *trans*) provides an efficient control over the secondary structures of the helix and strand.[14] On the basis of the results of the DOSY measurements for *trans*-ACHC homo-oligomers, we made an attempt to test the self-assembling behaviour of the *cis*-ACPC homo-oligomers, which adopt a self-stabilizing six-strand secondary structure in DMSO. The preliminary NMR measurements in CD_3OD revealed that the peptide-bond NH/ND exchange rate strongly decreases as the *cis*-$ACPC_7$ strand concentration increases (Figure 23). These observations cannot be explained purely within the framework of extremely stable secondary structures, and they therefore oriented our attention to self-association phenomena.

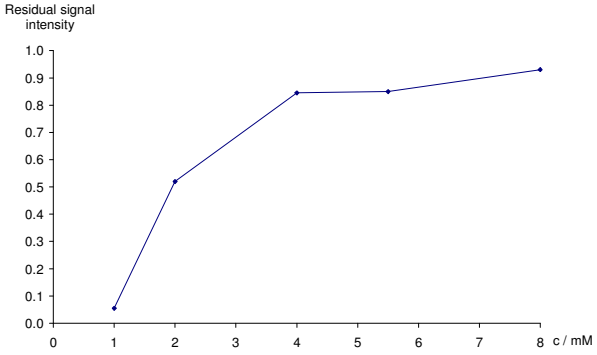


FIGURE 23. Concentration-dependent residual NH NMR signal of *cis*-$ACPC_7$ in CD_3OD after dissolution, based on integration of signals of NH at 7.48 ppm and $C^{\beta}H_1$ at 3.69 ppm.

First and most importantly, the morphology of the self-assembling structures was established by using TEM with negative staining (uranyl acetate). The time-dependent nature of the tertiary structure formation was followed by means of particle size measurements. DLS experiments were conducted in back-scattering mode to gain a picture of the particle size distribution in the solvent over a wide diameter range. In the same experimental setup, the ζ-potential too was measured. A diffusion NMR method was applied to monitor the self-

assembly at the oligomer level (diameter < 10 nm), where the DLS method may be unreliable because of the concomitant larger particles.

During the PFGSE NMR measurements on *cis*-ACPC$_7$, we found that the average diameter of the forming oligomers in methanol after 1 week at a concentration of 1 mM was 5.4 nm. Immediately after dissolution, only monomers were observed. The oligomerization measured without delay after sample preparation displayed a concentration-dependence (Figure 24), which points to a concentration-dependent self-association rate. Interestingly, the oligomerization in water was less pronounced, possibly because of the poorer solubility. The NMR gradient echo intensity did not exhibit a multiexponential delay, suggesting relatively fast *cis*-ACPC$_7$ strand-unit exchange between the dissolved monomers and the oligomers.

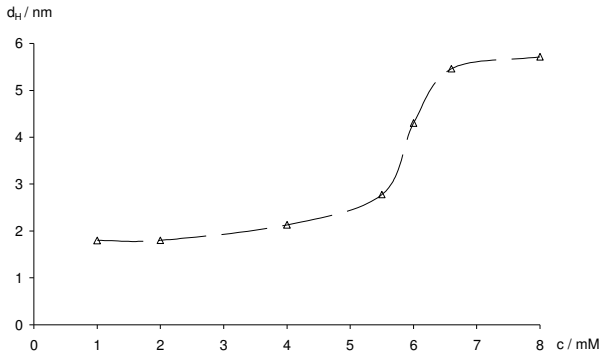


FIGURE 24. Concentration-dependence of average hydrodynamic diameter of *cis*-ACPC$_7$ oligomers, determined by diffusion NMR in CD_3OD after dissolution.

For the *cis*-ACPC$_7$ strand units, the TEM images clearly show ribbon-like fibrils in methanol after incubation of a 1 mM sample for 1 week at room temperature (Figure 25b).

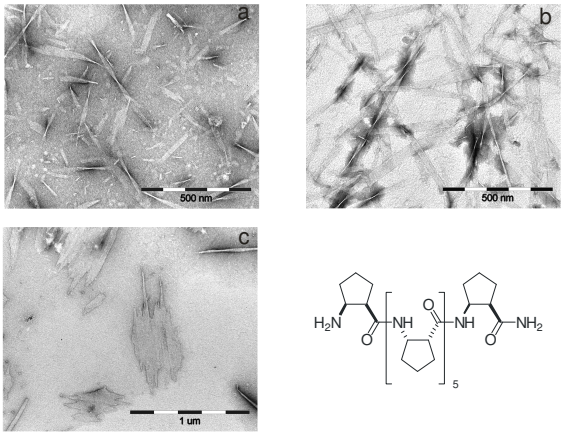


FIGURE 25. TEM images and chemical structure of *cis*-ACPC$_7$ fibrils. **a**. Fibrils detected after 1 day in 1 mM MeOH. **b**. Web of µm long fibrils observed after 1 week in 1 mM MeOH. **c**. Fibrils and sheets measured in water.

The ribbons are 30-50 nm wide and 2-3 µm long. In water, ribbons can likewise be recognized, but these are considerably wider, sheet formation being observed (Figure 25c). A

very interesting feature of the ribbon-like fibrils is the occurrence of helical turns, with a periodicity of 150-300 nm. The helical turns allow estimation of the average ribbon thickness, which is approximately 2.5 nm. When the self-association process is sampled after 1 day, considerably shorter (50-400 nm) fibril fragments are seen (Figure 25a). The solvent- and time-dependent size and morphology strongly support tertiary structure formation by a self-assembling process. The DLS data recorded in methanol (Figure 26) unequivocally reveal an increasing trend for the particle diameter, corroborating the above findings. After 1 week, particles in the μm range can be detected, the scattering intensity for these tapering off due to the sedimentation of the macroscopic species. The samples exhibit scattering intensity in the monomer/oligomer size range. The ζ-potential of the nanoribbons was 34 mV, which proves that the free N-termini of the *cis*-$ACPC_7$ strands are situated on the surface and are partially protonated. The significant ζ-potential endows the fibrils with stability, preventing coagulation.

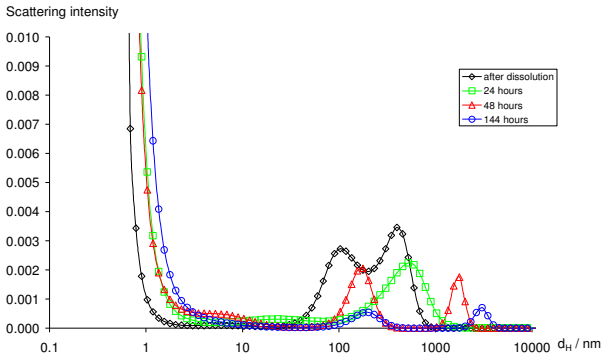


FIGURE 26. Size distribution of light scattering intensity for *cis*-$ACPC_7$ in 1 mM MeOH.

For the *trans*-$ACHC_6$ helix units, the TEM images demonstrated spherical objects after 1 day at a concentration of 1 mM both in methanol (Figure 27a) and in water.

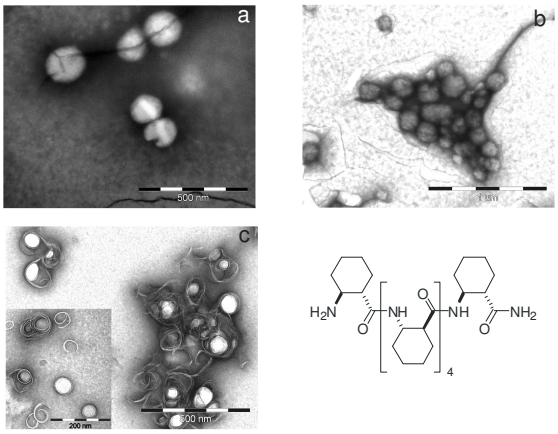


FIGURE 27. TEM images and chemical structure of *trans*-$ACHC_6$ vesicles. **a**. Vesicles detected after 1 day in 1 mM MeOH. **b**. Coagulated vesicles observed after 1 week in 1 mM MeOH. **c**. Disintegrated multilamellar vesicles and membranes measured after sonication of sample illustrated in diagram **a**.

These species are phenomenologically very similar to phospholipid vesicles in both shape and size.[106] The diameters of the spheres are in the range 60-180 nm. The shape and the size of the spheres did not vary in time, apart from a significant coagulation of the spheres (Figure 27b). In contrast with the structurally very informative dimensions of the *cis*-ACPC$_7$ nanoribbons, the *trans*-ACHC$_6$ nanospheres did not provide any hints as to the high-resolution structure. Following the hypothesis that the spheres behave like mono- or multilamellar phospholipid vesicles, samples were sonicated for 15 min (35 kHz, 34 W), which should disclose any membrane-like fine structure. After the treatment, the TEM images undoubtedly exhibited partially disintegrated vesicles built up from predominantly multilamellar walls (Figure 27c). The unilamellar wall thickness measured at the narrowest edges was 2.2 nm on average. The results of the DLS measurements support the TEM observations; the size distribution curve indicates particles in the range 10-200 nm, with a trend towards larger diameters in accordance with the slow coagulation (Figure 28). The scattering intensity disappears in a few days as a result of the sedimentation. The population in the monomer/oligomer size range observed by DLS is supported by the diffusion NMR finding of an average diameter of 3.4 nm at the oligomer level. The oligomerization was apparently independent of the concentration; the average diameter dropped only at the concentration of 0.05 mM. ζ-potential measurements furnished a value of 13 mV. It may be concluded that, despite the identical sample preparation protocol, the surface concentration of the basic N-terminal nitrogens was much lower than that of the *cis*-ACPC$_7$ ribbons. The low ζ-potential explains the increased coagulation propensity too.

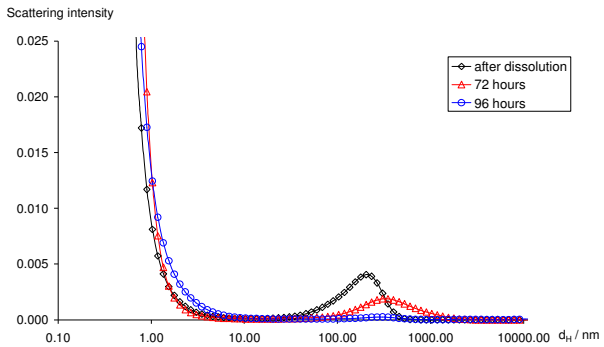


FIGURE 28. Size distribution of light scattering intensity for *trans*-ACHC$_6$ in 1 mM MeOH.

To interpret the structural features of the tertiary structures, molecular modelling was performed. On the basis of the measured nanoribbon dimensions and the known high-resolution structure of the similarly-behaving Alzheimer disease-related amyloid fibrils,[107,108] the construction of the ribbon model is rather straightforward. The ribbon height of 2.5 nm corresponds to the distance between the N- and C-termini of a single *cis*-ACPC$_7$ strand. The

peptide bonds are parallel to the long fibril axis, and the continuous pleated sheets associate in a sandwich tertiary structure, which is stabilized by hydrophobic interactions. *Ab initio* calculations carried out on the H-bond-stabilized and fully protonated parallel and antiparallel dimers of the *cis*-$ACPC_7$ strands at an HF/3-21G level prove that the pleated sheet is stable (Figure 29a); there is no steric repulsion and the stabilization energy *in vacuo* is estimated to be 40-50 kcal mol^{-1} (Figure 30).

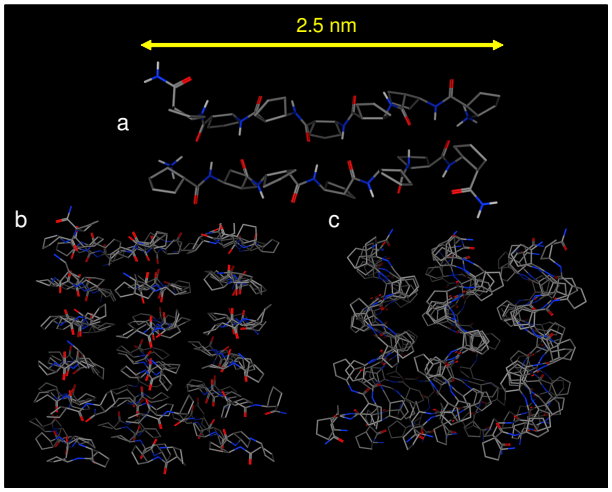


FIGURE 29. Molecular model of *cis*-$ACPC_7$ fibrils. a. *Ab initio* (HF/3-21G) model of antiparallel β-peptide sheet. b. Side-view of fibril fragment model after 3 ns MD simulation (explicit solvent not displayed). c. Top-view of fibril fragment model after 3 ns MD simulation.

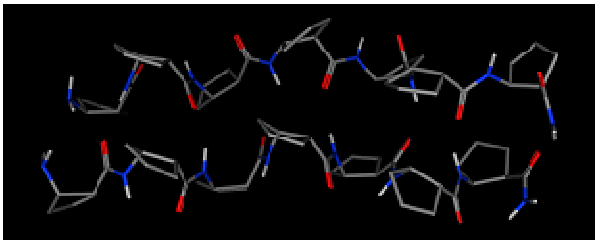


FIGURE 30. *Ab initio* model of parallel *cis*-$ACPC_7$ dimer.

The antiparallel pleated sheet is approximately 20 kcal mol^{-1} more stable than the parallel arrangement. The above findings indicated that the molecular model of the fragment of the *cis*-$ACPC_7$ nanoribbon was built up with anti-parallel strand orientations. Minimization and subsequent long-term molecular dynamics (MD) simulation (3 ns) were run with explicit water, periodic boundary conditions and the MMFF94[34] force field implemented in the Chemical Computing Group's Molecular Operating Environment. The ribbon model converged to a stable energy level; only side-chain vibrations were observed. Interestingly, the model structure displays signs of the helical twist along the axis parallel to the peptide bonds (Figure 30).

As regards the observed 2.2 nm wall thickness for the *trans*-$ACHC_6$ membranes and the low ζ-potential, we hypothesize head-to-head dimerization of the helices (Figure 31a). *Ab*

initio calculations resulted in a stable helix dimer stabilized by two H-bonds; the pillar atoms are the terminal NH_2 and the C=O on the second residue on both sides. The helix axes are tilted relative to each other by ca. 40°. The length of the dimer corresponds to the experimental membrane thickness, and the fact of the buried N-termini is in good accord with the low ζ-potential. Finally, it is safe to assume that the helix dimers form a hydrophobic surface sufficiently large to facilitate the solvent-driven association into a continuous helix bundle tertiary structure. Similar packing was observed in the solid phase by using X-ray crystallography. The cog-wheel-like mutual fit of the helix side-chains further increases the stabilizing hydrophobic surface (Figure 31). The *trans*-$ACHC_6$ membrane segment was constructed by following the above lines and a 3 ns MD was run with the same set-up as used for the nanoribbon model. The simulation revealed a stable membrane bilayer with side-chain vibrations.

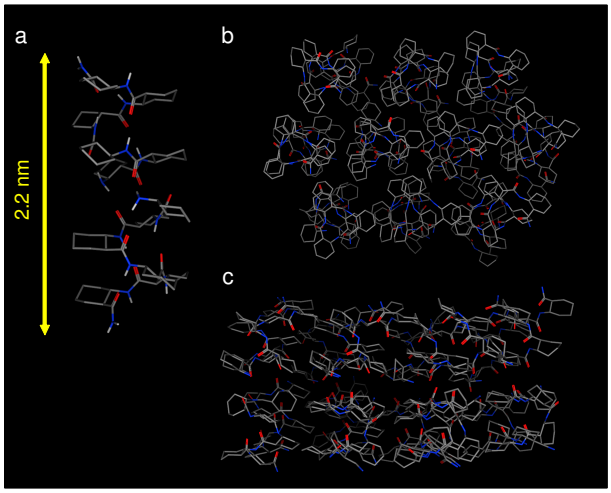


FIGURE 31. Molecular model of *trans*-$ACHC_6$ membranes. **a**. *Ab initio* (HF/3-21G) model of head-to-head helix dimer. **b**. Top-view of helix bundle membrane fragment model after 3 ns MD simulation (explicit solvent not displayed). **c**. Side-view of helix bundle membrane fragment model after 3 ns MD simulation.

4.3. Conclusions

It may be concluded that the conformationally constrained backbone of the studied models is flexible enough to afford both 10- and 14-helical motifs; in turn, this observation provides evidence of the true folding process and the question arises of whether the 10-helix is a nucleating intermediate in the folding of the 14-helix. In summary, it may be stated that a sharper picture of the folding rules of conformationally constrained β-peptides was obtained. The evidence of the chain length-dependent conformational preference indicates the flexibility of the studied models despite the conformationally constrained side-chains, which could be exploited to construct folded structures with dynamic functionality. The studied models undergo hydrophobically stabilized helical bundle formation in CD_3OD, where the propensity for self-association increases with the chain length. This finding draws the attention of the β-peptide community to the fact that oligopeptide aggregation might always be present, even if the usual indicators such as NMR signal broadening can not be observed. This points to the intrinsic self-associating propensity of helical units with hydrophobic side-chains in polar solvents.

We conclude that the designed β-peptide secondary structure units are capable of forming self-assembling intrinsic tertiary structures: fibrils of pleated sheet sandwiches and helix bundle membranes. Tuning of the relative configurations of the β-peptide backbone allows not only the secondary structure, but also the morphology of the nanostructured tertiary structures to be controlled. These observations and the stereochemical control for the tertiary structure design are unprecedented in biomimicking foldamer science. We believe that such newly-generated nanostructured materials may gain applications in the fields of nano- and biotechnology[109-111] and may contribute to a deeper understanding of the molecular mechanism of protein misfolding diseases.

5. Summary

1. The reactions of *cis*- or *trans*-1-aminomethylcyclohexane-1,2-diols (**1** or **2**) with one equivalent of aromatic aldehydes **3a–g**, at 300 K in $CDCl_3$ resulted in a complex multicomponent equilibrium mixture of **4a–g** and **6a–g** (or **5a–g** and **7a–g**), in each case consisting of a five-component ring–chain tautomeric system **4A–E** (or **5A–E**), involving the Schiff base, two epimeric spiro-oxazolidines, two epimeric condensed 1,3-oxazines, and some of the four tricyclic compounds **6A–D** (or **7A–D**). In the complex multicomponent equilibrium mixture, the five-component, ring–chain tautomeric system **4A–E** (or **5A–E**) was found to be adequately described by the Hammett–Brown linear free energy equation.

2. During a search for the possible reaction pathways of the formation of the tricyclic compounds **6A–D** (or **7A–D**) we found that they proceed via the reactions between the condensed 1,3-oxazines **4E** (or **5E**) and the Schiff base **4A** (or **5A**) by aldehyde transfer and aminodiol elimination.

3. An extended view of the stereoelectronic interactions affords a detailed description of the substituent-dependent stability changes exhibited by the 2aryl-1,3-*N*,*N*-heterocycles **11** (or **12**) in the corresponding ring–ring epimerization. The measured reaction free energies of the epimerization reactions of conformationally inflexible 2-aryl-1,3-*N*,*N*-heterocycles **11B**, **11C** (or **12B**, **12C**) were found to correlate well with the sum of the hyperconjugative stabilization energies of all the vicinal donor-acceptor orbital overlaps around C2, obtained from *ab initio* NBO analysis, and both quantities correlated linearly with the Hammett–Brown substituent constant.

4. The individual stereoelectronic interactions (n_N–$\sigma^*_{C2–N}$, n_N–$\sigma^*_{C2–Ar}$ and n_N–$\sigma^*_{C2–H}$) in **11B**⇌**11C** (or **12B**⇌**12C**) were also observed to exhibit a substituent-dependence, despite their distance from the 2-aryl substituent and the non-periplanar arrangement of the localized molecular orbitals. The higher the electron-withdrawing effect of the 2-aryl substituent, the larger was the stabilization for n_N–$\sigma^*_{C2–Ar}$, while the overlaps n_N–$\sigma^*_{C2–N}$ and n_N–$\sigma^*_{C2–H}$ changed in the opposite sense. The different polarization of the acceptor σ^* orbitals caused by the 2-aryl substituent accounted for the observed propagation of the substituent effect. These results promote a detailed explanation of the useful tautomeric behaviour of the 2-aryl-1,3-*X*,*N*-heterocycles, and reveal the nature of the

connection between the anomeric effect and the Hammett-type linear free energy relationship.

5. The conformationally constrained backbone of the homo-oligomers of *trans*-2-aminocyclohexanecarboxylic acid is flexible enough to afford both 10- and 14-helical motifs; in turn, this observation provides evidence of the true folding process of the *β*-petides. Homo-oligomers without protecting groups constructed by using *trans*-2-aminocyclohexanecarboxylic acid monomers **21**-**24** showed that the tetramer **22** tends to adopt a 10-helical motif, while both *ab initio* theory and NMR measurements point to the pentamer **23** and hexamer **24** forming the known 14-helix.

6. Such a conformational polymorphism is an important feature of any folded system that is designed to have a complex dynamic function in general; it is also observed for the natural *α*-peptides: the interplay between the *α*-helix and the 3_{10}-helix motif can be a crucial factor during the folding process. The revealed similar instrinsic properties of *β*-peptides suggest that the 10-helix may be a potential conformational intermediate in the folding process towards the thermodynamically stable 14-helix.

7. For the first time, direct evidence is presented on the tertiary structure of *β*-peptide foldamers, which proves that natural biopolymers are not unique in their highly structured conformational behaviour.

8. Stereochemically controlled secondary structure units of *β*-peptide strands and helices intrinsically self-assemble into sandwiches of pleated sheets and helix bundles, respectively. By residue control, the self-organization can be directed toward nano-sized fibrils or multilamellar vesicles. The results can have implications for new nanostructured materials and for a deeper understanding of protein folding and misfolding processes.

6. References

1. Valters, R. E.; Fülöp, F.; Korbonits, D. *Adv. Heterocycl. Chem.* **1996**, *66*, 1–71.
2. Lázár, L.; Fülöp, F. *Eur. J. Org. Chem*. **2003**, 3025–3042.
3. Martinek, T. A; Fülöp, F. *Eur. J. Biochem*. **2003**, *270*, 3657–3666.
4. Tice, C. M.; Ganem, B. *J. Org. Chem.* **1983**, *48*, 5043−5048.
5. Alcaide, B.; Plumet, J.; Rodriguez-Campos, I. M.; García-Blanco, S.; Martínez-Carrera, S. *J. Org. Chem.* **1992**, *57*, 2446–2454.
6. Lázár, L.; Lakatos, A. G.; Fülöp, F.; Bernáth, G.; Riddell, F. G. *Tetrahedron* **1997**, *53*, 1081−1088.
7. Martinek, T.; Lázár, L.; Fülöp, F.; Riddell, F. G. *Tetrahedron* **1998**, *54*, 12887−12896.
8. Saven, J. G. *Chem. Rev*. **2001**, *101*, 3113–3130.
9. Iverson, B. L. *Nature* **1997**, *385*, 113–115.
10. Appella, D. H.; Christianson, L. A.; Klein, D. A.; Powell, D. R.; Huang, X. L.; Barchi, J. J.; Gellman, S. H. *Nature* **1997**, *387*, 381–384.
11. Seebach, D.; Abele, S.; Gademann, K.; Jaun, B. *Angew. Chem. Int. Edit.* **1999**, *38*, 1595–1597.
12. Seebach, D.; Matthews, J. L. *Chem. Commun.* **1997**, 2015–2022.
13. Porter, E. A.; Wang, X.; Lee, H. S.; Weisblum, B; Gellman S. H. *Nature* **2000**, *404*, 565–565.
14. Martinek, T. A.; Tóth, G. K.; Vass, E.; Hollósi, M.; Fülöp, F. *Angew. Chem. Int. Ed.* **2002**, *41*, 1718–1721.
15. Cheng, R. P.; Gellman, S. H.; DeGrado, W. F. *Chem. Rev.* **2001**, *101*, 3219–3232.
16. Hill, D. J.; Mio, M. J.; Prince, R. B.; Hughes, T. S.; Moore, J. S. *Chem. Rev.* **2001**, *101*, 3893–4011.
17. Venkatraman, J.; Shankaramma, S. C.; Balaram, P. *Chem. Rev*. **2001**, *101*, 3131–3152.
18. Seebach, D.; Abele, S.; Gademann, K.; Guichard, G.; Hintermann, T.; Jaun, B.; Matthews, J. L.; Schreiber, J. V. *Helv. Chim. Acta* **1998**, *81*, 932–982.
19. Epand, R. F.; Raguse, T. L.; Gellman, S. H.; Epand, R. M. *Biochemistry* **2004**, *43*, 9527–9535.
20. Schmitt, M. A.; Weisblum, B.; Gellman, S. H. *J. Am. Chem. Soc*. **2004**, *126*, 6848–6849.
21. Jeener, J. Ampere Summer School, Basko Polje, Yugoslavia, 1971.
22. Aue, W. P.; Bartholdi, E.; Ernst, R. R. *J. Chem. Phys*. **1976**, *64*, 2229–2246.
23. Braunschweiler, L.; Ernst, R. R. *J. Magn. Reson*. **1983**, *53*, 521–528.
24. Bax, A.; Davis, D. G. *J. Magn. Reson*. **1985**, *65*, 355–360.
25. Müller. L. *J. Am. Chem. Soc*. **1979**, *101*, 4481–4484.
26. Bax, A.; Ikura, M.; Kay, L. E.; Torchia, D. A.; Tschudin, R. *J. Magn. Reson*. **1990**, *86*, 304–318.
27. Bodenhausen, G.; Ruben, D. J. *Chem. Phys. Lett*. **1980**, *69*, 185–189.
28. Norwood, T. J.; Boyd, J.; Heritage, J. E.; Soffe, N.; Campbell, I. D. *J. Magn. Reson*. **1990**, *87*, 488–501.
29. Bax, A.; Griffey, R. H.; Hawkins, B. L. *J. Mag. Res*. **1983**, *55*, 301–315.

30. Bax, A.; Subramanian, S. *J. Magn. Reson*. **1986**, *67*, 565–569.
31. Jeener, J.; Meier, B. H.; Bachmann, P.; Ernst, R. R. *J. Chem. Phys*. **1979**, *69*, 4546–4553.
32. Wagner, G.; Wüthrich, K. *J. Mol. Biomol*. **1982**, *155*, 347–366.
33. Bax, A.; Davis, D. G. *J. Magn. Reson*. **1985**, *63*, 207–213.
34. Halgren, T. A. *J. Comput. Chem*. **1996**, *17*, 490–519.
35. Becke, A. D. *Phys. Rev*. **1988**, *38*, 3098–3100.
36. Arnaud, R. *J. Comput. Chem*. **1994**, *15*, 1341–1356.
37. Alabugin, I. V.; Zeidan, T. A. *J. Am. Chem. Soc*. **2002**, *124*, 3175–3185.
38. Salzner, U.; von Ragué Schleyer, P. *J. Org. Chem*. **1994**, *59*, 2138–2155.
39. Cortés, F.; Tenorio, J.; Collera, O.; Cuevas, G. *J. Org. Chem.* **2001**, *66*, 2918–2924.
40. Reed, A. E.; Weinstock, R. B.; Weinhold, F. *J. Chem. Phys.* **1985**, *83*, 735–746.
41. Reed, A. E.; Curtiss, L. A.; Weinhold, F. *Chem. Rev.* **1988**, *88*, 899–926.
42. Foster, J. P.; Weinhold, F. *J. Am. Chem. Soc*. **1980**, *102*, 7211–7218.
43. Haris, P. I.; Chapman, D. *Biopolymers*, **1995**, *37*, 251–263.
44. Johnson, W. C. *Proteins* **1990**, *7*, 205–214.
45. Antalek, B. *Concept Magnetic Res*. **2002**, *14*, 225–258.
46. Egelhaaf, S. U.; Lobaskin, V.; Bauer, H. H.; Merkle, H. P.; Schurtenberger, P. *Eur. Phys. J. E* **2004**, *13*, 153–164.
47. Valluzzi, R.; Kaplan, D. L. *Biopolymers* **2000**, *53*, 350–362.
48. Johnson, B. A.; Blevins, R. A. *J. Biomol. NMR* **1994**, *4*, 603–614.
49. http://www.chemcomp.com
50. http://www.gaussian.com
51. Portmann, S; Lüthi, H. P. *Chimia* **2000**, *54*, 766–770.
52. Provencher, S. W. *Comp. Phys. Comm.* **1982**, *27*, 213–227.
53. Schuck, P. *Biophysical Journal* **2000**, *78*, 1606–1619.
54. Walsh, D. M.; Lomakin, A.; Benedek, G. B.; Condron, M. M.; Teplow, D. B. *J. Biol. Chem*. **1997**, *272*, 22364–22372.
55. Singh, H.; Sarin, R.; Singh, K. *Heterocycles* **1986**, *24*, 3039–3042.
56. Singh, H.; Singh, K. *Tetrahedron* **1989**, *45*, 3967–3974.
57. Singh, K.; Deb, P. K. *Tetrahedron Lett.* **2000**, *41*, 4977–4980.
58. Singh, K.; Deb, P. K.; Venugopalan, P. *Tetrahedron* **2001**, *57*, 7939–7949.
59. Singh, K.; Deb, P. K.; Behal, S. *Heterocycles* **2001**, *55*, 1937–1942.
60. Kempf, D. J.; Sowin, T. J.; Doherty, E. M.; Hannick, S. M.; Codavoci, L.; Henry, R. F.; Green, B. E.; Spanton, S. G.; Norbeck, D. W. *J. Org. Chem*. **1992**, *57*, 5692–5700.
61. Wang, G. T.; Li, S.; Wideburg, N.; Krafft, G. A.; Kempf, D. J. *J. Med. Chem.* **1995**, *38*, 2995–3002.
62. Chen, P.; Cheng, P. T. W.; Alam, M.; Beyer, B. D.; Bisacchi, G. S.; Dejneka, T.; Evans, A. J.; Greytok, J. A.; Hermsmeier, M. A.; Humphreys, W. G.; Jacobs, G. A.; Kocy, O.; Lin, P. F.; Lis,

K. A.; Marella, M. A.; Ryono, D. E.; Sheaffer, A. K.; Spergel, S. H.; Sun, C. Q; Tino, J. A.; Vite, G.; Colonno, R. J.; Zahler, R.; Barrish, J. C. *J. Med. Chem.* **1996**, *39*, 1991–2007.

63. Alcon, M.; Poch, M.; Moyano, A.; Pericas, M. A.; Riera, A. *Tetrahedron: Asymmetry* **1997**, *8*, 2967–2974.
64. Alexander, C. W.; Liotta, D. C. *Tetrahedron Lett.* **1996**, *37*, 1961–1964.
65. Beaulieu, P. L.; Gillard, J.; Bailey, M.; Beaulieu, H.; Duceppe, J. S.; Lavallée, P.; Wernic, D. *J. Org. Chem.* **1999**, *64*, 6622–6634.
66. Maiereanu, C.; Darabantu, M.; Plé, G.; Berghian, C.; Condamine, E.; Ramondenc, Y.; Silaghi-Dumitrescu, I.; Mager, S. *Tetrahedron* **2002**, *58*, 2681–2693.
67. Monge, S.; Sélambarom, J.; Carré, F.; Verducci, J.; Rogue, J. P.; Pavia, A. A. *Carbohydr. Res.* **2000**, *328*, 127–133.
68. Star, A.; Goldberg, I; Fuchs B. *J. Organomet. Chem.* **2001**, *630*, 67–77.
69. Star, A.; Goldberg, I; Fuchs B. *Angew. Chem. Int. Ed.* **2000**, *39*, 2685–2689.
70. Star, A.; Fuchs B. *J. Org. Chem.* **1999**, *64*, 1166–1172.
71. Tschierske, C.; Köhler, H.; Zaschke, H.; Kleinpeter, E. *Tetrahedron* **1989**, *45*, 6987–6998.
72. Köhler, H.; Tschierske, C.; Zaschke, H.; Kleinpeter, E. *Tetrahedron* **1990**, *46*, 4241–4246.
73. Lázár, L.; Göblyös, A.; Martinek, T. A.; Fülöp, F. *J. Org. Chem*. **2002**, *67*, 4734–4741.
74. Neuvonen, K.; Fülöp, F.; Neuvonen, H.; Koch, A.; Kleinpeter, E.; Pihlaja K. *J. Org. Chem.* **2001**, *66*, 4132–4140.
75. Kirby, A. J. The anomeric effect and related stereoelectronic effects at oxygen; Springer Verlag: Berlin, **1983**.
76. Juaristi, E.; Cuevas, G. *Tetrahedron* **1992**, *48*, 5019–5087.
77. Alabugin, I. V. *J. Org. Chem*. **2000,** *65*, 3910–3919.
78. Schleyer, P. V.; Kos, A. J. *Tetrahedron* **1983**, *39*, 1141–1150.
79. Carballeira, L.; Pérez-Juste, I. *J. Phys. Chem*. **2000**, *104*, 9362–9369.
80. Mikolajczyk, M.; Garaczyk, P.; Wieczorek, M. W.; Bujacz, G. *Angew. Chem. Int. Ed*. **1991**, *30*, 578–580.
81. Reed, A. E.; Schleyer, P. V. *J. Am. Chem. Soc*. **1987**, *109*, 7362–7373.
82. Anderson, C. B.; Sepp, D. T. *J. Org. Chem*. **1968**, *33*, 3272–3276.
83. Juaristi, E.; Tapia, J.; Méndez, R. *Tetrahedron* **1986**, *42*, 1253–1264.
84. Juaristi, E.; González, E. A.; Pinto, B. M.; Johnston, B. D.; Nagelkerke, R. *J. Am. Chem. Soc*. **1989**, *111*, 6745–6749.
85. King, J. F.; Rathore, R.; Guo, Z. R.; Li, M. Q.; Payne, N. C. *J. Am. Chem. Soc*. **2000**, *122*, 10308–10324.
86. Rowan, S. J.; Cantrill, S. J.; Cousins, G. R. L.; Sanders, J. K. M.; Stoddart, J. F. *Angew. Chem. Int. Ed*. **2002**, *41*, 898–952.
87. Appella, D. H.; Christianson, L. A.; Karle, I. L.; Powell, D. R.; Gellman, S. H. *J. Am. Chem. Soc*. **1996**, *118*, 13071–13072.

88. Appella, D. H.; Christianson, L. A.; Karle, I. L.; Powell, D. R.; Gellman, S. H. *J. Am. Chem. Soc*. **1999**, *121*, 6206–6212.
89. Appella, D. H.; Christianson, L. A.; Klein, D. A.; Richards, M. R.; Powell, D. R.; Gellman, S. H. *J. Am. Chem. Soc*. **1999**, *121*, 7574–7581.
90. Barchi, J. J., Huang, X. L.; Appella, D. H.; Christianson, L. A.; Durell, S. R.; Gellman, S. H. *J. Am. Chem. Soc*. **2000**, *122*, 2711–2718.
91. Gellman, S. H. *Acc. Chem. Res*. **1998**, *31*, 173–180.
92. Appella, D. H.; Barchi, J. J.; Durell, S. R.; Gellman, S. H. *J. Am. Chem. Soc*. **1999**, *121*, 2309–2310.
93. Raguse, T. L.; Lai, J. R.; LePlae, P. R.; Gellman, S. H. *Org. Lett*. **2001**, *3*, 3963–3966.
94. Banerjee, A.; Balaram, P.; *Curr. Sci.* **1997**, *73*, 1067–1077.
95. Seebach, D.; Gademann, K.; Schreiber, J. V.; Matthews, J. L.; Hintermann, T.; Jaun, B.; Oberer, L.; Hommel, U.; Widmer, H. *Helv. Chim. Acta* **1997**, *80*, 2033–2038.
96. Günther, R.; Hofmann, H. J. *Helv. Chim. Acta* **2002**, *85*, 2149–2168.
97. Seebach, D.; Schreiber, J. V.; Abele, S.; Daura, X.; van Gunsteren, W. F. *Helv. Chim. Acta* **2000**, *83*, 34–57.
98. Daura, X.; Bakowies, D.; Seebach, D.; Fleischhauer, J.; van Gunsteren, W. F.; Krüger, P.; *Eur. Biophys. J*. **2003**, *32*, 661–670.
99. Claridge, T. D. W.; Goodman, J. M.; Moreno, A.; Angus, D.; Barker, S. F.; Taillefumier, C.; Watterson, M. P.; Fleet, G. W. J. *Tetrahedron Lett.* **2001**, *42*, 4251–4255.
100. Bolin, K. A.; Millhauser, G. L. *Acc. Chem. Res.* **1999**, *32*, 1027–1033.
101. Beke, T.; Csizmadia, I. G.; Perczel, A. *J. Comp. Chem.* **2004**, *25*, 285–307.
102. Mohle, K.; Günther, R.; Thormann, M.; Sewald, N.; Hofmann, H. J. *Biopolymers* **1999**, *50*, 167–184.
103. Valentini, M.; Vaccaro, A.; Rehor, A.; Napoli, A.; Hubbell, J. A.; Tirelli, N. *J. Am. Chem. Soc*. **2004**, *126*, 2142–2147.
104. González-Pérez, A.; Del Castillo, J. L.; Czapkiewicz, J.; Rodríguez, J. R. *Colloids Surfaces A* **2004**, *232*, 183–189.
105. Jolicoeur, C.; Philip, P. R. *Can. J. Chem*. **1974**, *52*, 1834–1839.
106. Luisi, P. L.; Stano, P.; Rasi, S.; Mavelli, F. *Artifical Life* **2004**, *10*, 297–308.
107. Soto, C. *Nature Rev. Neurosci.* **2003**, *4*, 49–60.
108. Branden, C.; Tooze, J. Introduction to protein structure; Garland Publishing: New York, **1998**.
109. Ball, P. *Nanotechnology* **2005**, *16*, R1–R8.
110. Hartgerink, J. D.; Beniash, E.; Stupp, S. I. *Science* **2001**, *294*, 1684–1688.
111. Silva, G. A.; Czeisler, C.; Niece, K. L.; Beniash, E.; Harrington, D. A.; Kessler, J. A.; Stupp, S. I. *Science* **2005**, *303*, 1352–1355.

7. Acknowledgements

This work was carried out in the Institute of Pharmaceutical Chemistry, University of Szeged, during the years 2001-2005.

I would like to express my warmest thanks to Professor Ferenc Fülöp, Director of the Institute, for his guidance of my work, his inspiring ideas, his useful advice and his constructive criticism.

My grateful thanks are also due to my co-supervisor, Dr. Tamás Martinek, for his continuous support and interest in my activities. His advice and help have been invaluable during all stages of my work.

I am greatly indebted to Professor Kalevi Pihlaja, Department of Chemistry, University of Turku, for providing me with the opportunity to work for 6 months in his research group.

Many thanks are due to Dr. Frank G. Riddell, School of Chemistry, University of St. Andrews, for providing me with the opportunity to work for 6 months in his research group.

I would like to thank Dr. Livia Fülöp for the TEM measurements.

I express many thanks to all my co-authors for the synthetic work and discussions.

I am grateful to Dr. David Durham for revising the English language of the Thesis.

I owe very much to my colleagues and my friends for creating all the circumstances enabling me to carry out this work.

Appendix