Eur. Phys. J. B **17**, 499–513 (2000)

THE EUROPEAN PHYSICAL JOURNAL B

EDP Sciences

Self-assembly and structure transformations in living polymers forming fibrils

I.A. Nyrkova[1,a], A.N. Semenov[1], A. Aggeli[2], M. Bell[2], N. Boden[2], and T.C.B. McLeish[2]

[1] Department of Applied Mathematics, University of Leeds, Leeds LS2 9JT, UK
[2] Centre for Self-Organising Molecular Systems, University of Leeds, Leeds LS2 9JT, UK

Received 28 June 1999

Abstract. The classical isodesmic one-dimensional model for equilibrium polymerization is generalized in order to describe self-assembly in systems forming fibrils. The model was applied to peptide solutions forming β-sheet tapes which can further aggregate into stacks of various thickness: double tapes and fibrils (several double tapes stacked together). We found that in some cases the model yields several step-like transitions as the concentration increases: first from monomers to single or double tapes, and then to fibrils. The abruptness of the first transition is controlled by the free energy penalty for transformation of a peptide from random coil to a straight β-strand conformation (the latter is characteristic for tapes). If both single and double tapes are allowed, the length of the aggregates after the first transition can be very large with high scission energies. For very low energies of attraction between double tapes, the transition from double tapes to fibrils happens separately (above the first transition), and it is even more abrupt and produces extremely long fibrils. The theoretical findings are used to extract the characteristic molecular parameters for the self-assembly of the *de novo* peptide DN1 forming polymeric β-sheets in water.

PACS. 36.20.Ey Conformation (statistics and dynamics) – 61.46.+w Clusters, nanoparticles, and nanocrystalline materials – 87.15.-v Biomolecules: structure and physical properties

List of symbols

f_β: fraction of peptides in all kinds of β-sheet tapes, equation (17);

$\langle m\rangle^{(p)}$: number average length of p-fibril (hence, $\left(p\,\langle m\rangle^{(p)}\right)$ is the average aggregation number), equation (11);

v_β: effective volume of chemical bonds between peptides in the same single β-sheet, equation (B.4);

v_p: effective attraction volume of the tapes in a p-folded tape;

v_2: effective attraction volume for single tapes inside a double tape, equation (B.2), v_{fib} is such volume for double tapes forming a $2p$-fibril, equation (B.3), $v_{2p}^{2p-1} \equiv v_2^p v_{\rm fib}^{p-1}$;

$E_{\rm tr} \equiv \exp(-\varepsilon_{\rm tr})$, $\varepsilon_{\rm tr}$: transformation energy (free energy difference between coiled peptide and a peptide in a rod-like conformation), equation (7);

$E_\beta \equiv (v_0/v_\beta)\exp(-\varepsilon_\beta)$, ε_β: β-scission energy (total bond energy between two rods inside a single β-sheet), equation (7);

$E_p \equiv \exp(-\varepsilon_p)$, ε_p: fibrillization energy for a fibril made of p single tapes (mean free energy gain per peptide in a p-folded stack as compared to peptide in a single tape), equation (9);

$E_2 \equiv \exp(-\varepsilon_2) \equiv \exp(-\varepsilon_{\rm dbl}/2)$, $\varepsilon_{\rm dbl}$: cross-tape (side bond) free energy gained due to aggregation into double tape, ε_2 is the corresponding free energy gain per peptide, equation (33);

$E_{2p} \equiv \exp(-\varepsilon_{2p}) \equiv E_2 E_{\rm net}$, $\varepsilon_{2p} \equiv \varepsilon_2 + \frac{p-1}{2p}\varepsilon_{\rm fib}$, $E_{\rm net} \equiv \exp\left(\frac{p-1}{2p}\varepsilon_{\rm fib}\right)$, $\varepsilon_{\rm fib}$: net energy (per an extra inter-peptide contact) gained when double tapes condense to form fibrils (made of p double tapes), equation (51);

p_*: the stacking number of the most favourable fibril, $\Lambda_* \equiv \min_{p=2}^{p_0}\{E_p\} \equiv E_{p_*}$, equation (13), Section 3.3;

$c_* \equiv K_1 \equiv (E_\beta \Lambda_*)/(E_{\rm tr} v_0) \equiv (v_\beta)^{-1}\exp(\varepsilon_{\rm tr} - \varepsilon_\beta - \varepsilon_{p_*})$,

[a] e-mail: irina@amsta.leeds.ac.uk

equation (18); *e.g.* in Section 4.2: $\tilde{\tilde{K}}_1 \equiv (E_\beta E_2)/(E_{\rm tr} v_0)$, equation (34), and in Section 4.3: $\tilde{K}_1 \equiv (E_\beta E_{2p})/(E_{\rm tr} v_0) \equiv \tilde{\tilde{K}}_1 E_{\rm net}$, equation (53);

$K_2 \equiv E_{\rm tr}^2 v_0/E_\beta = v_\beta \exp(\varepsilon_\beta - 2\varepsilon_{\rm tr})$, equation (24);

$K_3 \equiv (E_{\rm tr} E_\beta/\Lambda_*)\,(1-\Lambda_*)^{-4}\,(v_2/v_0)$, equation (36); *e.g.* in Section 4.2: $\tilde{\tilde{K}}_3 \equiv (E_{\rm tr} E_\beta/E_2)\,(1-E_2)^{-4}\,(v_2/v_0)$, and in Section 4.3: $\tilde{K}_3 \equiv (E_{\rm tr} E_\beta/E_{2p})\,(1-E_{2p})^{-4}\,(v_2/v_0)$, equation (53).

$K_4 \equiv E_\beta^2\,(1-E_2)^{-4}\left(v_2 v_{\rm fib}/v_0^2\right)$, equation (60).

1 Introduction

Self-assembling systems based upon protein-like molecules are particularly interesting. They are biocompatible and biodegradable water-based systems, potentially highly responsive to moderate changes in the media properties (like pH, temperature, ionic composition, etc.). The structures of native peptides are very diverse, being determined by regular arrangements of hydrogen bonds as well as side-chain interactions [1]. Depending on the solution conditions, the same peptide (primary structure) can form various higher-order secondary structures, hence lability of the peptide systems.

Researchers have already tried to exploit the self-assembling properties of natural peptides by designing new oligomeric peptide chains which form either solid crystals of well-defined architecture [2], or peptide nanotubules [3], or macroscopic membranes [4]. However, a self-assembling of proteins is also the molecular mechanism for the development of pathological situations such as Alzheimer's, prions and other amyloidosis diseases [5], and for the aggregation of lens proteins which leads to cataract [6].

The two main regular secondary structures found in peptides are alpha-helices and beta-sheets. In the former the hydrogen bonds link aminoacid residues which are relatively close to each other (3-4 aminoacid residues apart) inside the same peptide chain. On the contrary, the β-sheet structures involve links between distant fragments of the same peptide chain or even between fragments of different peptide molecules, leading to self-assembly of these molecules. In native proteins, short β-sheets or barrels are widely displayed, and there are also extended β-sheets in silk [1]. Where identical short peptide molecules are involved, these self-assembling sheets may be very long, *i.e.* tape-like. The formation of such self-assembling β-sheet tapes was reported recently for several types of *de novo* designed oligopeptides [7–9]. When placed in their respective solvents, these oligopeptides aggregate into long semi-flexible tapes already well below millimole (mM) concentrations. This self-assembly into tapes is often accompanied by gel formation at higher concentrations for some oligopeptides; liquid crystalline ordering and formation of higher ordered self-assembled structures have been observed as well [7–9].

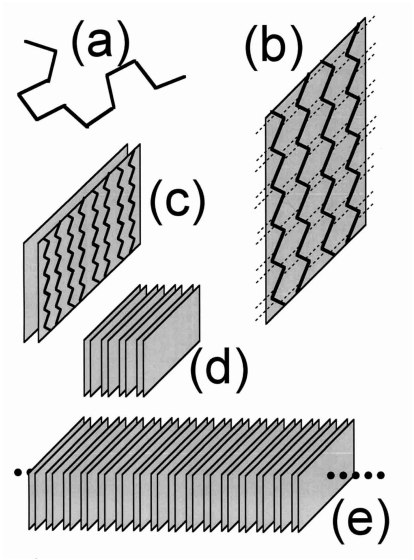

Fig. 1. Self-assembling structures in solutions of DN1 peptide molecules: coiled lone molecule (a), single β-sheet tape (hydrogen bonds are shown with dotted lines) (b), double tape (primary tape to form fibrils) (c), stack of several double tapes (fibril) (d). A natural expectation for the stack structure in the absence of twist: wide ("infinite") sheet (e).

One of the novel peptides, the 11-residue peptide DN1 (CH_3CO-Gln-Gln-Arg-Phe-Gln-Trp-Gln-Phe-Glu-Gln-Gln-NH_2), was observed to form a very rich variety of structures in water [9]. Moreover, some of the structures resemble those found in native peptides, hence it can be considered as a suitable model peptide for biological applications.

At very low concentrations in water the DN1 peptide molecules are mostly in the form of lone *coils* (just one peptide molecule in a random coil or a helix conformation depending on the solvent properties), Figure 1a. However, at higher concentrations DN1 forms β-sheet *tapes* (effectively one-dimensional structures consisting of rod-like (β-strand) peptides connected with each other *via* hydrogen bonds and stabilized additionally with inter-strand side chain interactions), see Figure 1b. For the particular case of DN1, one side of such tape is very hydrophobic, thus favouring the formation of *double tapes*, see Figure 1c. The outer sides of such double tapes can still be slightly mutually attractive in water, thus causing a formation of *stacks* formed by many double tapes, Figure 1d.

The most natural expectation is that at high enough concentrations these tapes would form very wide stacks (two-dimensional sheets), Figure 1e, and that the width of the stack would increase with concentration in analogy with other self-assembling systems (*e.g.* living polymers). However, the actual result is a formation of *fibrils* which are thread-like structures with well-defined diameter, see Figure 2a, references [9,10].

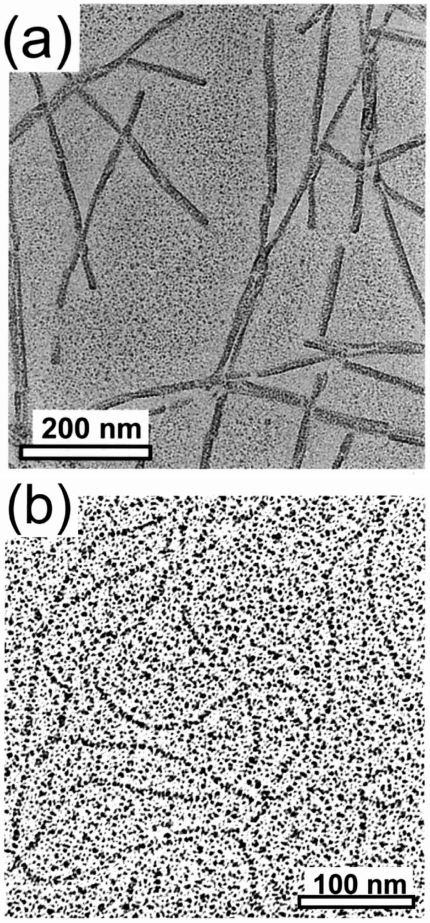

Fig. 2. Structures formed in solutions of DN1 peptide in pure water as revealed by electron microscopy (platinum shadowing technique). High concentration structures: thick rod-like fibrils ($[c] \simeq 6$ mM) (a). Lower concentration structures: thinner flexible objects, presumably double tapes ($[c] \simeq 200$ μM) (b).

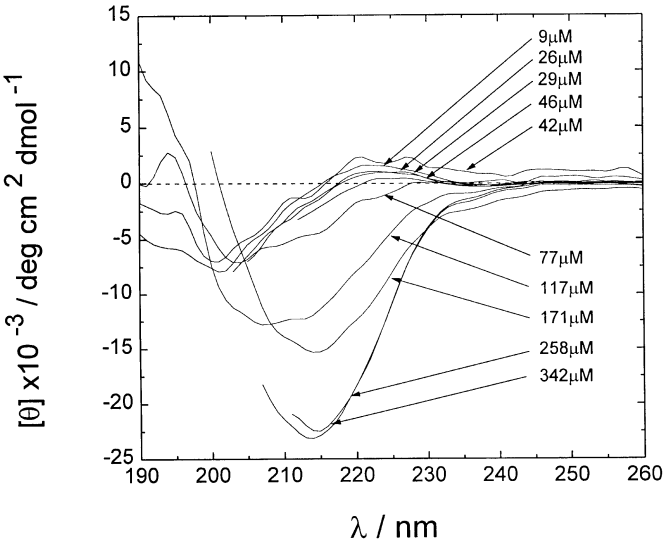

Fig. 3. Far-UV CD equilibrium spectra of DN1 water solutions, the concentrations are indicated along the lines.

We have shown in a separate publication [10] that this stabilisation of the stack size has its origin in the twisted nature of the primary tapes. Indeed, DN1 peptide molecules contain L-aminoacid residues which are chiral, hence the equilibrium structure of the double tape (being a building block for stacks) is twisted. When forming a stack, as in Figure 1d, the primary (double) tapes should either untwist, or bend, or experience both types of deformation. As a result, a stack formation leads not only to gaining a negative energy of face-to-face attraction between tapes, but also to increase of their elastic energy. Depending on the molecular parameters of the primary tapes, the final result may be either a formation of twisted fibrils (*e.g.* stacks of intertwisted and bent tapes with a finite width which is determined by a balance of attraction and elastic energies) or a formation of an "infinite" sheet (a sheet in which the tapes are completely untwisted and lay parallel to each other, the width of such sheet is concentration dependent and tends to infinity in the limit of high concentrations) [10]. Formation of infinite sheets is only possible if the attraction energy is high enough to compensate a complete untwisting of the tapes. Twisted fibrils are formed with smaller attraction energies, the formation of such fibrils being especially easy if the bending energy is small [10]. It is finite size fibrils with practically monodisperse diameter which are formed in solutions of DN1 peptide at high enough concentrations.

Thus the following structures are formed in the DN1 peptide aqueous solutions: a) lone (monomeric) peptides, b) single tapes, c) double tapes, d) fibrils formed by several double tapes glued together [9].

For understanding of the self-assembling processes in such peptide tape systems it is important to define the characteristic energy parameters which govern the structural transformations. Many peculiarities of the behaviour of peptide-based β-sheet tapes are not common for other linear self-assembling systems (*e.g.* living polymers, cylindrical surfactant vesicles, etc. accounted for by isodesmic one-dimensional self-assembling model [11]):

1) The formation of self-assembling tape structures happens abruptly, see Figure 3. At lower than c_* concentrations, the tape fraction keeps low and grows slowly, and the tape length is small. However, at the threshold peptide concentration c_* the tapes start to be observable both with UV spectroscopy and in electron microscopy images. The length of the observable tapes turns out to be very long, Figure 2b. This is contrast to usual living polymers which exhibit a smooth length growth [11].

2) Similar abruptness accompanies a transformation from tapes to fibrils as concentration is increased.

3) All transformations are very slow. The equilibrium distributions can be achieved only within many days (in some cases within weeks or even months/years).

All these facts suggest that peptide tapes represent a self-assembling system with very special (energetic) characteristics. Our aim is to explore the corresponding possibilities, and to analyse the most general scenarios of self-assembly in these systems. We thus introduce a general

model, analyze it and compare with the related experimental data available for the DN1 peptide system. We will explore mainly the only precise enough data on DN1 solutions as to how the equilibrium composition of DN1 solution changes with concentration.

The model allow us to do estimations for the parameters controlling self-assembling behaviour in peptide tape systems and to explain the peculiarities mentioned above. The results for the energetic parameters obtained from the composition data are important on their own, because there is no other way to get these parameters which must control kinetics and rheology of tape systems.

This paper deals only with equilibrium properties of peptide solutions. The kinetic aspects will be considered in a separate publication. Next section is devoted to experimental details. In the third section we formulate and analyse the basic model which allows to explore concentration dependence of composition of solution of self-assembling entities in the general case. In Section 4 we analyse some particular cases of this general model and consider them in view of the data on DN1 system, the minimum set of the molecular parameters needed to characterize DN1 system is established and the estimations for these parameters are given. Section 5 is devoted to a concluding discussion.

2 Experimental

The details of the synthesis, purification, structure analysis, samples preparation of DN1 peptide, as well as the details of transmission electron microscopy (TEM) and far-UV CD spectroscopy can be found elsewhere, see references [8,10]. DN1 was rationally designed to self-assemble into polymeric hydrogen-bonded β-sheet tapes in water (Fig. 1b). Moreover, DN1 β-sheets combine into double tapes (Fig. 1c) because of the presence of the aromatic rings in the 4th, 6th and 8th side-chains, which gives rise to a hydrophobic "adhesive" stripe running along one side of each single tape, causing them to associate in pairs in water. In a rod-like β-strand conformation (inside a β-sheet) DN1 molecule is characterized by the length $b_1 \simeq 37$ Å (equivalent to the width of the tape), the adjacent peptide-peptide periodicity inside the β-structure $b_2 \simeq 4.7$ Å and the thickness across a single β-sheet $a_1 \simeq 10 \div 12.5$ Å (hence the thickness of the double tape is $a = 2a_1 \simeq 20 \div 25$ Å).

The presence of two morphologically different kinds of self-assembling structures (tapes and fibrils) in DN1 water solutions is apparent from the TEM images, Figure 2; it is also revealed in far-UV CD spectroscopy, see references [8–10]. At higher concentrations (above *ca.* 0.6 mM) the DN1 peptide forms fibrils which look like rigid rod-like polymers, $D \simeq 8 \div 10$ nm wide, with apparent persistence length of the order of several dozen of micrometers, see Figure 2a. These fibrils are monodisperse in width, and we believe that they are stacks of *ca.* 4 double tapes stuck together by their longer sides due to face-to-face attraction, Figure 1d. At lower concentrations (*ca.* $80 \div 600$ μM) DN1 peptide forms flexible chains less than 5 nm wide, with apparent persistence length of the order of a micrometer, Figure 2b. These chains are apparently double tapes. If the fibril solution is diluted below 300 μM and left at room temperature, over a course of several months, the fibrils are seen to unwind into several tapes [9]. Note that the contour length observed in TEMs Figure 2 may be limited by multiple ruptures of the fibrils and tapes during preparation of the samples for TEM imaging.

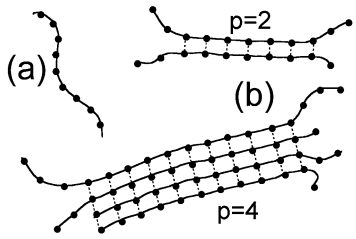

Fig. 4. Typical self-assembling structures: a single tape (a), many-folded sheaf-like tapes (fibrils): double tape and a fibril with $p = 4$ (b). Peptide molecules are drawn by thick points, junctions between neighboring peptides in β-sheets are shown by thin solid lines, and face-to-face attraction — by dotted lines.

The equilibrium far-UV CD spectra of DN1 peptide in pure water at various concentrations are shown in Figure 3. For the most dilute solutions (below 50 μM) the spectra have a negative ellipticity at *ca.* 200 nm and a positive ellipticity at *ca.* 222 nm, characteristic of random coil monomeric peptide. At 150 $\mu\text{M} < [c] < 600$ μM they have a negative CD band at *ca.* 214 nm and a positive band below *ca.* 195 nm, typical of a β-sheet conformation, which we attribute to the double DN1 tapes. The growth of the negative β-sheet band intensity $[\theta]_{214}$ reflects the transition from monomeric state to tapes and can be used for fitting of our theoretical models. In such cases we take the band intensity $[\theta]_{214}$ to be a linear function of the fraction of peptides in β-sheets, f_β:

$$[\theta]_{214} = k_1 f_\beta + k_2. \tag{1}$$

3 Equilibrium statistics of self-assembling systems forming tape-like structures

3.1 Model and the basic general equations for the composition

Let us consider a solution of N elementary objects (peptide molecules) which can form tapes. Among the tapes we consider single tapes, double ones, triple ones, and so on up to p_0-fold tapes. Each single tape has a structure of a living polymer: the peptides are connected into one-dimensional sequence, Figure 4a. In many-folded tapes, single ones are glued to each other by their longer sides. For the sake of simplicity we suppose that the tapes have a sheaf-like structure: the main "body" of a p-folded tape is indeed a stack of p single β-sheets (each made of m peptides) combined together; at both ends the tape splits into p "legs" which are short fragments of single tapes,

Figure 4b. As we will see below, usually these legs should be rather short (made of not more than a few peptide molecules, see Eq. (12) below), hence their influence on the global solution structure should be quite weak, hence we can ignore more complicated end structures of many-fold tapes.

The peptides inside a tape adopt a β-strand (rod-like) conformation. Lone (free) peptides are characterized by a different conformational state (random coil or alpha-helix, depending on solvent conditions) which generally has lower internal free energy than the rod-like conformation. However, peptides joint into a tape gain a linking (scission) energy which is enough to counterbalance the energy difference between the rod-like conformation and the lone peptide state (cp. condition (5) below).

The composition of such peptide solution can be characterized by the numbers of each type of tapes, so that the total number of peptide in volume V is fixed:

$$N_1 + \sum_{m=2}^{\infty} mN_m^{(1)} + \sum_{p=2}^{p_0} \sum_{m=2}^{\infty} \sum_{n_1,n_2,\ldots n_{2p}=0}^{\infty} \times \left(mp + \sum_{i=1}^{2p} n_i \right) N^{(p)}_{m,n_1,\ldots,n_{2p}} = N. \quad (2)$$

Here N_1 is the number of monomers (lone coils), $N_m^{(1)}$ is the number of single tapes of length m ($m > 1$ as the rod-like conformation is stable only inside β-sheet), $N^{(p)}_{m,n_1,\ldots,n_{2p}}$ is the number of p-fold tapes with the body of length m and with $2p$ legs of lengths $\{n_i\}$ ($1 \le i \le 2p$). The free energy of a solution with such composition can be written as

$$\begin{aligned}
\frac{\mathcal{F}}{k_BT} &= N_1 \ln\left(\frac{N_1 v_0}{eV}\right) + \sum_m N_m^{(1)} \ln\left(\frac{N_m^{(1)} v_0}{eV}\right) \\
&+ \sum_p \sum_m \sum_{n_1,n_2,\ldots n_{2p}} N^{(p)}_{m,n_1,\ldots,n_{2p}} \ln\left(\frac{N^{(p)}_{m,n_1,\ldots,n_{2p}} v_0}{eV}\right) \\
&- N_1\varepsilon_{\rm tr} + \sum_m \left(-\varepsilon_\beta - \ln\left(\frac{v_\beta}{v_0}\right)\right)(m-1)N_m^{(1)} \\
&+ \sum_p \sum_m \sum_{n_1,n_2,\ldots n_{2p}} N^{(p)}_{m,n_1,\ldots,n_{2p}} \\
&\times \left\{ -\varepsilon_\beta \left[(m-1)p + \sum_i n_i \right] - \varepsilon_p mp \right. \\
&\left. - \left[(m-1)p + \sum_i n_i \right] \ln\left(\frac{v_\beta}{v_0}\right) - (p-1)\ln\left(\frac{v_p}{v_0}\right) \right\}, \quad (3)
\end{aligned}$$

where all summations are the same as in equation (2). Here the first three terms represent translational free energies of the species, and other terms represent the internal free energy inputs. $e = \exp(1)$, v_0 is the microscopic elementary volume (its value is irrelevant for what follows), v_β is the effective volume of chemical bonds between peptides in the same single tape (it characterizes the degree of freedom left for a rod connected with the others in a β-sheet), v_p is the effective attraction volume of single tapes in a p-folded tape (it characterizes how freely each tape in a many-folded stack can move around while preserving the general structure of the p-folded aggregate, Fig. 4b), *cf.* with estimations of the volumes in Appendix B. The following energy parameters enter equation (3): $\varepsilon_{\rm tr}$ is the free energy difference between a lone free peptide and a peptide in a rod-like conformation as inside a β-sheet (*transformation energy*), ε_β is the total bond energy between two rods inside a single β-sheet (β *-scission energy*), and ε_p is the mean free energy gain per peptide in a p-folded stack (p-fibril) as compared to a peptide in a single tape (*fibrillization energy*):

$$\varepsilon_p = -\mathcal{E}_{p-\rm fibril} + \mathcal{E}_{\rm single-tape} \quad (4)$$

where $\mathcal{E}_{p-\rm fibril}$ is the internal free energy of a long p-fibril (per peptide) and $\mathcal{E}_{\rm single-tape}$ is the same quantity for a long single tape. Thus ε_p incorporates the free energy of surface face-to-face attraction between the tapes and the increment of elastic energy of the tapes associated with fibril formation. In a separate publication devoted to fibril formation [10] we show that usually ε_p has a well-defined maximum at some particular p_* which indicates the most favourable fibrils. All energies are measured in $k_{\rm B}T$ units. Note also that all three energy parameters ε_β, $\varepsilon_{\rm tr}$ and ε_p are assumed to be positive.

The natural condition that the β-sheet structure is stable at some concentration implies that the scission energy exceeds the transformation energy:

$$\varepsilon_\beta > \varepsilon_{\rm tr} > 0. \quad (5)$$

Also the fact, that at very low concentrations the fibrils split into still long tapes, implies that scission is accompanied by larger energy gain than fibrillization:

$$\varepsilon_\beta + \ln(v_\beta/v_p) > \varepsilon_p > 0. \quad (6)$$

In order to find the equilibrium composition of the solution, one should minimize the free energy (3) with respect to the numbers N_1, $N_m^{(1)}$, ... satisfying the condition (2). This procedure yields the following formulae for the fractional composition:

$$\frac{N_1 v_0}{V} = \Lambda \frac{E_\beta}{E_{\rm tr}}, \quad E_\beta \equiv \frac{v_0}{v_\beta}\exp(-\varepsilon_\beta),$$
$$E_{\rm tr} \equiv \exp(-\varepsilon_{\rm tr}) < 1; \quad (7)$$

$$\frac{N_m^{(1)} v_0}{V} = \Lambda^m E_\beta, \quad (8)$$

$$\frac{N^{(p)}_{m,n_1,\ldots n_{2p}} v_0}{V} = \left(\frac{E_\beta}{E_p^m}\right)^p \left(\frac{v_p}{v_0}\right)^{p-1} \Lambda^{pm} \prod_{i=1}^{2p} \Lambda^{n_i},$$
$$\frac{v_p}{v_0} E_\beta < E_p \equiv \exp(-\varepsilon_p) < 1, \quad (9)$$

(in Eqs. (8, 9) $m \ge 2$, $n_i \ge 0$) with $\ln(\Lambda E_\beta)$ being the Lagrange multiplier ($0 < \Lambda < 1$). The latter can be found

using the condition (2): it is determined indirectly by the total peptide concentration $c = N/V$:

$$c \equiv \frac{N}{V} = \frac{E_\beta}{E_{\rm tr} v_0}\Lambda + \frac{E_\beta}{v_0}\Lambda^2 \frac{2-\Lambda}{(1-\Lambda)^2} + \sum_{p=2}^{p_0} v_p^{p-1} \left(\frac{E_\beta}{v_0}\right)^p \frac{p\,(\Lambda/E_p)^{2p}}{(1-\Lambda)^{2p}\,(1-(\Lambda/E_p)^p)} \times \left[\frac{2-(\Lambda/E_p)^p}{(1-(\Lambda/E_p)^p)} + \frac{2\Lambda}{1-\Lambda}\right]. \quad (10)$$

Here the first term is concentration of monomers $c^{(0)} \equiv N_1/V$, the second term is concentration of peptides in the single tapes $c^{(1)} \equiv \sum \left(m N_m^{(1)}/V\right)$ and the last term is concentration of peptides in the p-fold tapes $c^{(p)} \equiv \sum \left([mp + \sum n_i]\, N^{(p)}_{m,n1,\ldots n_{2p}}/V\right)$.

An increase of the chemical activity of the monomers (*e.g.* increase of parameter Λ) results in an increase of all terms in the r.h.s. of equation (10), thus reflecting an increase in concentrations of all types of tapes and in the total concentration N/V. One can easily perform numerical analysis of equation (10) and get the composition of the system $\left\{c^{(j)}\right\}$ $(0 < j < p_0)$ as a function of the total concentration for any values of molecular parameters $\{E_\beta/v_0, v_p, E_{\rm tr}, E_2, \ldots E_{p_0}\}$. Below we start with general analysis of equation (10), and then consider some particular combinations of molecular parameters.

3.2 The tape lengths

First we find the typical length of the tapes (fibrils). Using equations (7–9) we get the number average length of the "bodies" of p-folded tapes:

$$\langle m\rangle^{(p)} = \frac{\sum_m \left[\sum_{n_1,\ldots n_{2p}}\right] m N^{(p)}_{m,n_1,\ldots,n_p}}{\sum_m \left[\sum_{n_1,\ldots n_{2p}}\right] N^{(p)}_{m,n_1,\ldots,n_p}} = 1 + \frac{1}{1-(\Lambda/E_p)^p}, \quad (11)$$

and the length of the "legs":

$$\langle n_{i_0}\rangle^{(p)} = \frac{\sum_m \left[\sum_{n_1,\ldots n_{2p}}\right] n_{i_0} N^{(p)}_{m,n_1,\ldots,n_p}}{\sum_m \left[\sum_{n_1,\ldots n_{2p}}\right] N^{(p)}_{m,n_1,\ldots,n_p}} = \frac{\Lambda}{1-\Lambda}\,. \quad (12)$$

Now we note that the value of Λ in equation (10) should be additionally restricted from the above:

$$0 < \Lambda < \Lambda_* \equiv \min_{p=2}^{p_0} \{E_p\} < 1, \quad (13)$$

as $E_p < 1$, equation (9). Hence, the leg length is always small ($\lesssim 1$), if $\Lambda < \min_p E_p \lesssim 1/2$ (cp. with Eq. (44) below), and thus the leg structure is indeed insignificant (we have used this statement at the beginning of this section where we introduced classification of the tapes under consideration). It is important to compare $\langle n_i\rangle^{(p)}$ with the rigidity segment of the primary tape (or "leg"). If $\langle n_i\rangle^{(p)}$ is larger than the corresponding persistence segment, the fibril structure would be very defective. On the other hand, if $\langle n_i\rangle^{(p)}$ is small compared to the tape persistence segment, the structure of many-folded tapes must be sheaf-like. It is the latter case that is considered below since persistence length of many-fold tapes was observed to be not less than $P \sim 0.5$ μm (see Eqs. (A.1, A.2)) corresponding to more than $P/c \sim 10^3$ peptides per persistence length.

3.3 High concentration limit

As Λ approaches the boundary Λ_*, see equation (13), the contribution of fibrils with $p = p_*$ to the total concentration defined in equation (10) becomes dominant (here p_* corresponds to the lowest E_p: $E_{p_*} < E_p$ if $p \neq p_*$): the p_*-contribution is of the order $\mathcal{O}\left(\delta\Lambda^{-2}\right)$ where

$$\delta\Lambda \equiv (1-(\Lambda/\Lambda_*)^p) \to 0 \quad \text{as } \Lambda \to \Lambda_*. \quad (14)$$

All other terms remain finite when $\delta\Lambda$ tends to zero. Hence, the corresponding concentrations approach some limiting values $c_*^{(p)}$:

$$c \simeq c_*^{(0)} + \sum_{j\neq p_*} c_*^{(j)} + \frac{\text{const.}}{\delta\Lambda^2}\,, \quad \text{as } \delta\Lambda \to 0; \quad (15)$$

$$c_*^{(0)} \equiv \frac{E_\beta}{E_{\rm tr} v_0}\Lambda_*; \quad c_*^{(1)} \equiv \frac{E_\beta}{v_0}\Lambda_*^2\,\frac{2-\Lambda_*}{(1-\Lambda_*)^2}, \quad \text{etc.} \quad (16)$$

As a result, the fraction f_β of β-sheet peptide follows a *universal curve* for high concentrations:

$$f_\beta \equiv \frac{c-c^{(0)}}{c} \simeq 1 - \frac{c_*}{c}, \quad \text{if } \frac{c}{c_*} \gg 1, \quad c_* \equiv K_1, \quad (17)$$

where

$$K_1 \equiv \frac{E_\beta}{E_{\rm tr} v_0}\Lambda_* \equiv \frac{E_\beta}{E_{\rm tr} v_0}\min\{E_p\} \equiv \frac{1}{v_\beta}\exp\left(\varepsilon_{\rm tr} - \varepsilon_\beta - \varepsilon_{p_*}\right), \quad (18)$$

i.e. K_1 is determined by the energy difference between the free coil and the peptide inside the most favourable p_*-fold tape. Note that $\Lambda/\Lambda_* - 1 \ll 1$ in the region $c/c_* \gg 1$ since

$$E_{\rm tr} \ll 1, \quad E_\beta \ll v_0/v_p \quad (19)$$

(*cf.* Eqs. (5, 6)).

The equilibrium lengths of the tapes with $p \neq p_*$ saturate in the high concentration region: using equations (11, 14) we get

$$\langle m\rangle^{(p)} - 1 = \frac{1}{\delta\Lambda + \Lambda^p\left[\Lambda_*^{-p} - E_p^{-p}\right]} \simeq \frac{1}{1-(\Lambda_*/E_p)^p}, \quad c/c_* \gg 1, \quad p \neq p_*. \quad (20)$$

The limiting length $\langle m\rangle^{(p)}$ can be large if the corresponding E_p is close to Λ_* defined in equation (13). The length of the most favourable fibrils (with $p = p_*$) is concentration dependent. Using equations (10, 11, 14, 15) we get:

$$\langle m\rangle^{(p_*)} - 1 = \frac{1}{\delta\Lambda} \simeq [c - c_*]^{1/2}\,(1-\Lambda_*)^p \times \left(p_* v_{p_*}^{p_*-1}\left(\frac{E_\beta}{v_0}\right)^{p_*}\right)^{-1/2}, \quad \frac{c}{c_*} \gg 1. \tag{21}$$

3.4 Low concentration region

For low concentrations we can approximate equation (10) as

$$c \simeq \frac{E_\beta}{E_{\rm tr}v_0}\Lambda + 2\frac{E_\beta}{v_0}\Lambda^2, \quad \Lambda \lesssim \frac{1}{2}\Lambda_* < \frac{1}{2} \tag{22}$$

(see uneqs. (19)). The first term here corresponds to the monomers (lone peptides) and the second term is due to single tapes (with $m = 2$). Note that the first term in r.h.s. of equation (22) dominates, therefore the β-sheet peptide fraction in this regime is

$$f_\beta \simeq \frac{2\Lambda^2 E_\beta}{\Lambda E_\beta/E_{\rm tr} + 2\Lambda^2 E_\beta} \simeq 2\Lambda E_{\rm tr} \simeq 2K_2 c, \quad \frac{c}{c_*} \lesssim \frac{1}{2}, \tag{23}$$

where

$$K_2 \equiv \frac{E_{\rm tr}^2}{E_\beta}v_0 = v_\beta \exp(\varepsilon_\beta - 2\varepsilon_{\rm tr}), \tag{24}$$

i.e. K_2 is determined by the energy difference between two lone free peptides and a single β-sheet "tape" of two peptides.

4 Particular scenarios of tape/fibril formation

We now proceed to an analysis of the simplest particular situations. Our aim is to extract as many molecular parameters of DN1 peptides as possible using their far-UV CD spectra, Figure 3, and the corresponding electron micrographs, Figure 2.

4.1 Only single tapes are possible

We start with the simplest situation when single tapes ($p = 1$) only are allowed. In this case, the basic equation (10) includes only two terms:

$$c \equiv \frac{N}{V} = \frac{E_\beta}{E_{\rm tr}v_0}\Lambda + \frac{E_\beta}{v_0}\Lambda^2\frac{2-\Lambda}{(1-\Lambda)^2}\,. \tag{25}$$

The critical value of the chemical activity is $\Lambda_* = 1$ (cp. Eq. (13)), and the critical concentration is $c_* = E_\beta/E_{\rm tr}v_0$

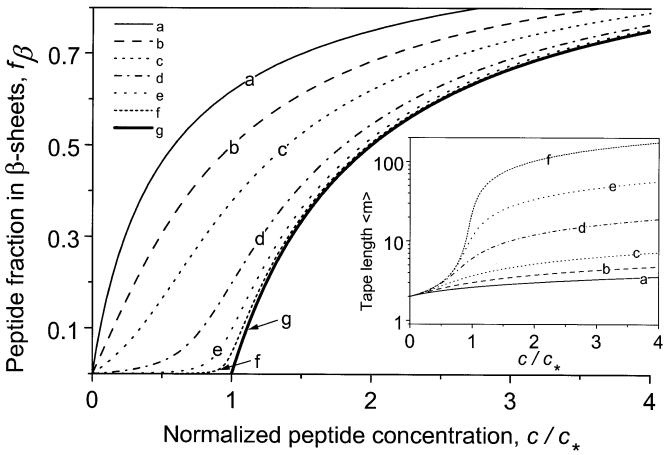

Fig. 5. Concentration dependencies of the tape fraction f_β for the case when only single tapes are allowed (drawn in accordance with Eq. (25)). The values of the transformation energy exponent: $E_{\rm tr} = 1$ (a), 0.33 (b), 0.1 (c), 0.01 (d), 10^{-3} (e), 10^{-5} (f). The last curve (g) corresponds to the asymptotics (27) ($E_{\rm tr} \to 0$). The value E_β/v_0 does not affect the curves $f_\beta(c/c_*)$. The transition concentration $c_* \equiv E_\beta/E_{\rm tr}v_0$, equation (26).

(see Eqs. (17, 18)). At low concentrations the tape fraction is approximately defined in equation (23), and at high concentrations it is defined in equation (17):

$$f_\beta \simeq \left(2\frac{E_{\rm tr}^2}{E_\beta}v_0\right)c \equiv 2E_{\rm tr}\frac{c}{c_*}, \quad \text{if } c \ll c_* \equiv \frac{E_\beta}{E_{\rm tr}v_0}; \tag{26}$$

$$f_\beta \simeq 1 - \frac{E_\beta}{E_{\rm tr}v_0 c} \equiv 1 - \frac{c_*}{c}, \quad \text{if } \frac{c}{c_*} - 1 \gg E_{\rm tr}^{1/3}. \tag{27}$$

The typical tape length can be estimated as (see Eqs. (11, 25)):

$$\langle m\rangle \equiv \frac{\sum_m m N_m^{(1)}}{\sum_m N_m^{(1)}} = \frac{2-\Lambda}{1-\Lambda} \simeq \left(\frac{c-c_*}{E_\beta}v_0\right)^{1/2} \equiv \left(\left(\frac{c}{c_*}-1\right)\frac{1}{E_{\rm tr}}\right)^{1/2}, \quad \text{if } \frac{c}{c_*} - 1 \gg E_{\rm tr}^{1/3}. \tag{28}$$

Note that equation (28) is different to equation (21): the latter is valid for $p_* \geq 2$.

The concentration dependence of f_β is determined by two parameters, $E_{\rm tr}$ and E_β/v_0 (see Eqs. (25–27)). Note that the value of E_β/v_0 does not affect the *shape* of the composition curve: indeed, Λ (and f_β) as defined by equation (25) is invariant with respect to the transformation

$$E_\beta/v_0 \to \alpha E_\beta/v_0; \quad c \to \alpha c \tag{29}$$

with any positive α. Hence the shape of the dependence of f_β *vs.* c/c_* is determined by the value of the transformation energy only, see Figure 5.

For $E_{\rm tr} = 1$ (the transformation energy $\varepsilon_{\rm tr} = 0$, hence there is no initial energetic barrier for the tape formation), we recover the classical formula for living polymers [11]. In this case the transition from monomers to living chains

(tapes) is very broad, the tape length in the transition region is moderate, and the curve $f_\beta(c)$ is convex for any c (see the curve (a) in Fig. 5).

For lower $E_{\rm tr}$ this transition is more pronounced. Indeed, a decrease of $E_{\rm tr}$ leads to a decrease of both the initial slope $\mathrm{d}f/\mathrm{d}(c/c_*)$ (see Eq. (26)), and the correction to the asymptotics given by equation (27). The typical tape length just above the transition point scales as $E_{\rm tr}^{-1/2}$, *i.e.* it is increasing as $E_{\rm tr} \to 0$ (see Eq. (28)). Figure 5 illustrates these transformations. If $E_{\rm tr} \gtrsim 0.25$ (*i.e.* $\varepsilon_{\rm tr} \lesssim 1.5$), then the transition to tapes is still very broad and the curve $f_\beta(c)$ is convex as for the classical situation (cp. the curves (a) and (b) in Fig. 5). At $E_{\rm tr} \sim 0.1$ ($\varepsilon_{\rm tr} \simeq 2.3$) the initial part of the curve becomes concave, however deviations from the asymptotic behavior (27) are still large: the apparent transition point is located at $c_{\rm app} \simeq 0.3\ c_*$ whereas the theoretical transition point is c_* (see the curve (c) in Fig. 5). At $E_{\rm tr} \sim 0.01$ ($\varepsilon_{\rm tr} \simeq 4.6$) the f_β *vs.* c curve reveals the characteristic two stage behavior with semi-linear initial part (Eq. (26)) and hyperbolic shape at high concentrations (Eq. (27)). For $E_{\rm tr} \lesssim 10^{-3}$ ($\varepsilon_{\rm tr} \gtrsim 7$) its behavior becomes very close to the asymptotic one, and the errors in experimental determination of the parameter $c_* \equiv E_\beta/(E_{\rm tr} v_0)$ from the apparent transition concentration is less than 10% (see the curves (e) and (f) in Fig. 5).

Now we can try to fit the spectroscopic experimental data, Figure 3, to our theory supposing that only single tapes are allowed. The β-sheet band values $[\theta]_{214}$ (Fig. 6) do show a characteristic two-stage behavior with an apparent transition at $[c]_* \simeq 70\ \mu$M. Note that it is the molar concentration, the corresponding number concentration being $c_* \simeq 70 \times c_\mu \simeq 4.2 \times 10^{-8}$ Å^{-3}:

$$c \equiv [c] \times c_\mu; \quad c_\mu \equiv \frac{N_A}{10^3\ \mathrm{cm}^3} \simeq 6.02 \times 10^{-10}\ \text{Å}^{-3}, \quad (30)$$

where $[c]$ is measured in μM. However, the initial slope is not too small yet, this suggesting that $E_{\rm tr}$ is not very small either. The best fit for the spectroscopy data yields $E_{\rm tr} \simeq 0.01$ (see Fig. 6). This corresponds to the tape length $\langle m \rangle \sim 15$ for $[c] \simeq 200\ \mu$M which is much smaller than the observed lengths

$$\langle m\,([c] \simeq 200\ \mu\mathrm{M}) \rangle \geq 1000 \quad (31)$$

(see Fig. 2b). That long tapes are predicted only for $E_{\rm tr} < 10^{-6}$ (dashed line in Fig. 6, cp. Eq. (28)) in drastic disagreement with the value $E_{\rm tr} \simeq 0.01$ obtained by fitting the spectroscopy data (Fig. 6).

It is easy to show that any other model, based on the assumption that only *one* type of tapes is possible, produces similar results, *i.e.* these models *fail* to explain the spectroscopic data $[\theta](c)$ and the TEM photographs simultaneously: the molecular parameters extracted from the $[\theta](c)$ curve imply very short tape lengths in contradiction with the photographs.

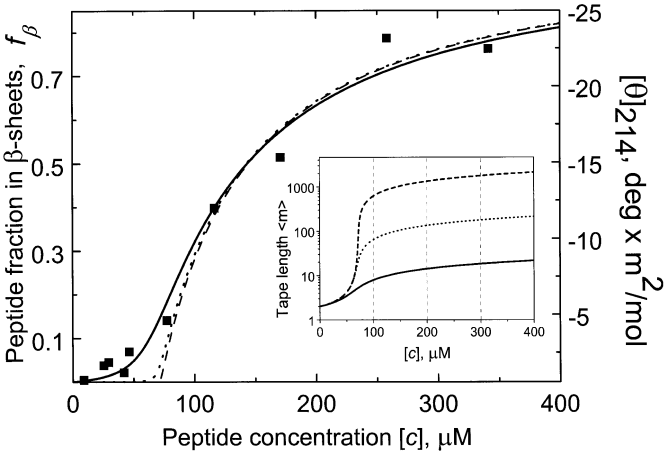

Fig. 6. Theoretical fits for the tape fraction f_β and the number average tape length $\langle m \rangle = (2 - \Lambda)/(1 - \Lambda)$ (the inset) concentration dependencies, based on the model of section 4.1 (Eqs. (25, 28)), when only single tapes are allowed. The far-UV CD data for the values $[\theta]_{214}$ of the negative band intensity (as in Fig. 3) are shown with boxes related to the right ordinate axis. These values are assumed to be a linear function of f_β, see equation (1). The theoretical curves are drawn with lines related to the left axes: the best CD fit ($E_{\rm tr} = 0.01$; $E_\beta/v_0 = 4.8 \times 10^{-10}$) being drawn with solid lines, the fit ($E_{\rm tr} = 10^{-4}$; $E_\beta = 4.35 \times 10^{-12}$) being drawn with dotted lines, and the fit ($E_{\rm tr} = 10^{-6}$; $E_\beta = 4.35 \times 10^{-14}$) — with dashed lines. The latter set of parameters implies long enough tapes to satisfy the condition (31), however the CD fitting is poor in this case.

4.2 Monomers, single tapes and double tapes (two types of tapes are possible)

Now we proceed to consideration of the situations when more thick tapes are allowed in addition to single types. We start with the case of two different types of tapes. For the case of DN1 peptide we expect that these are single tapes (the classic β-sheet structures) and double tapes (see Sect. 2). Then the basic composition equation (10) takes the form:

$$c \equiv \frac{N}{V} = \frac{E_\beta}{E_{\rm tr} v_0}\Lambda + \frac{E_\beta}{v_0}\Lambda^2 \frac{2 - \Lambda}{(1 - \Lambda)^2} + \frac{2E_\beta^2 v_2}{v_0^2} \frac{(\Lambda/E_2)^4}{(1 - \Lambda)^4 \left(1 - (\Lambda/E_2)^2\right)} \times \left[\frac{2 - (\Lambda/E_2)^2}{\left(1 - (\Lambda/E_2)^2\right)} + \frac{2\Lambda}{1 - \Lambda} \right] \quad (32)$$

with notations defined in equations (7, 9). Here

$$E_2 \equiv \exp(-\varepsilon_2) \equiv \exp(-\varepsilon_{\rm dbl}/2), \quad (33)$$

ε_2 is free energy gain (per peptide) due to aggregation of single β-sheets into double tapes (the total energy of side bonds is thus $2m\varepsilon_2$, where m is the number of peptides in each single tape), $\varepsilon_{\rm dbl} = 2\varepsilon_2$ is the cross-tape linkage (side

bond) free energy. Note that now the composition of the solution is determined by four independent parameters: E_β/v_0, $E_{\rm tr}$, v_2 and E_2.

The critical value of the chemical activity is $\Lambda_* = E_2$, see equation (13). The first term in equation (32) always dominates over the second one since $E_{\rm tr}(1-E_2)^{-2} \ll 1$.

As Λ approaches E_2 from below, the concentration of monomer peptides and the number of peptides in single tapes (given by the first two terms in the sum (32) respectively) tend to the asymptotic values:

$$c^{(0)} \simeq c_* \equiv K_1 = \frac{E_\beta E_2}{E_{\rm tr} v_0} \equiv \tilde{\tilde{K}}_1;$$
$$c^{(1)} \simeq c_*^{(1)} \equiv \tilde{\tilde{K}}_1 E_2 E_{\rm tr} \frac{2-E_2}{(1-E_2)^2}\,. \tag{34}$$

Hence the tape fraction is determined by equation (17):

$$f_\beta \simeq 1 - \frac{c_*}{c}, \quad \text{if } \frac{c}{c_*} - 1 \gg \Delta, \tag{35}$$

$$\Delta \equiv \max\left\{\frac{E_{\rm tr}E_2}{(1-E_2)^2}, K_3^{1/3}\right\}, \quad K_3 \equiv \frac{E_{\rm tr}E_\beta v_2}{v_0 \Lambda_* (1-\Lambda_*)^4}\,. \tag{36}$$

In the limit of high concentrations, the typical lengths of single and double tapes are

$$\left\langle m^{(1)}\right\rangle = 1 + \frac{1}{1-\Lambda} \simeq 1 + \frac{1}{1-E_2} \tag{37}$$

$$\left\langle m^{(2)}\right\rangle = 1 + \frac{1}{1-(\Lambda/E_2)^2} \simeq 1 + \left[\frac{1}{2K_3}\left(\frac{c}{c_*}-1\right)\right]^{1/2}, \quad \text{if } \frac{c}{c_*} - 1 \gg \Delta \tag{38}$$

(cp. Eqs. (20, 21)).

Note that the region of validity of the asymptotic equations (34, 35, 37, 38) extends almost down to $c = c_*$ since

$$E_{\rm tr} \ll 1, \quad E_\beta v_2/v_0 \ll 1. \tag{39}$$

Indeed, the tape lengths observed in TEM (Fig. 2b) correspond to condition (31). From equations (37, C.1) it is clear that single tapes cannot be so long, hence it is double tapes that are visible in Figure 2b, *i.e.*

$$\left\langle m^{(2)}([c] = 200\ \mu\text{M})\right\rangle \geq 10^3. \tag{40}$$

From Figure 3 it is clear that $[c]_* \sim 70\ \mu$M, and hence equations (38, 40) imply that

$$K_3 \lesssim 10^{-6}. \tag{41}$$

Below (Eq. (44)) we show that $E_{\rm tr} \sim 0.01 \div 0.1$ and $E_2 \sim 1/2$. Therefore, $E_\beta v_2/v_0 \lesssim 10^{-4}$, *i.e.* the inequalities (39) are verified.

Note that in the region $c < c_*$ the third term in the r.h.s. of equation (32) is much smaller than the second term which, in turn, is much smaller than the first one (see uneq. (39)). Hence, a more universal dependence of f_β *vs.* c in this region:

$$f_\beta \simeq E_{\rm tr} \frac{\frac{c}{c_*}\left(\frac{2}{E_2} - \frac{c}{c_*}\right)}{\left(\frac{1}{E_2} - \frac{c}{c_*}\right)^2}, \quad \text{if } c < c_*. \tag{42}$$

We are now in a position to fit the far-UV CD spectroscopy data, $[\theta]_{214}$ from Figure 3, with equation (32) using $E_{\rm tr}$, E_β/v_0, E_2 and v_2 as fitting parameters. Inequalities (39) ensure that the whole dependence $f_\beta(c)$ is quite accurately represented by the two asymptotics, equations (35, 42). Hence one can extract the combination $c_* \equiv \tilde{\tilde{K}}_1$ (34) from the high concentration part of the experimental data set, $[\theta](c)$ for $c > c_*$. The result is

$$[c_*] \equiv \tilde{\tilde{K}}_1/c_\mu \simeq 73 \pm 4, \tag{43}$$

with a good ($\pm 5\%$) accuracy. With this $\tilde{\tilde{K}}_1$, the values of $E_{\rm tr}$ and E_2 can be obtained using the low concentration part ($c < c_*$) of the same data. The best fit yields

$$E_{\rm tr}^{\rm bf} \simeq 0.05, \quad E_2^{\rm bf} \simeq 0.55;$$
$$(E_\beta/v_0)^{\rm bf} \equiv \exp(-\varepsilon_\beta)/v_\beta \simeq 4 \times 10^{-9}\ \text{Å}^{-3}. \tag{44}$$

However, reasonable fittings are achieved also for other $E_{\rm tr}$ and E_2 ranging in the interval:

$$\text{from} \quad E_{\rm tr} \simeq 0.012;\ E_2 \simeq 0.9;\ E_\beta/v_0 \simeq 5.9 \times 10^{-10}\ \text{Å}^{-3} \tag{45}$$

$$\text{to} \quad E_{\rm tr} \simeq 0.2;\ E_2 \simeq 0.2;\ E_\beta/v_0 \simeq 4.4 \times 10^{-8}\ \text{Å}^{-3} \tag{46}$$

(see Fig. 7). The value of v_2 nearly does not affect the predicted dependence $f_\beta(c)$ provided that K_3 is small. Condition (40) implies that K_3 must be smaller than the corresponding threshold $K_3^{\max}$ which depends on the choice of energetic parameters, see Figure 7:

$$(K_3^{\max})^{\rm bf} \simeq 7.1 \times 10^{-7}, \quad K_3^{\max} \simeq 2.4 \times 10^{-7}, \quad K_3^{\max} \simeq 7.6 \times 10^{-7} \tag{47}$$

respectively for the best fit (44), and for the boundaries (45, 46). The volume v_2 hence must be smaller than the threshold volume $v_2^{\max}$ which is determined by equation (36) for $K_3^{\max}$. With the values $K_3^{\max}$ from equation (47) we have

$$(v_2^{\max})^{\rm bf} \simeq 80\ \text{Å}^3, \quad v_2^{\max} \simeq 3\ \text{Å}^3, \quad v_2^{\max} \simeq 7\ \text{Å}^3 \tag{48}$$

respectively (for the best fit (44), and for the boundaries (45, 46)). One can easily check that the range of the parameters, equations (44–48), is in accordance with the estimates done in Appendixes B and C, equations (B.2, C.1, C.2). Finally, for the primary molecular parameters of the DN1 peptide tapes (the β-sheet scission energy ε_β,

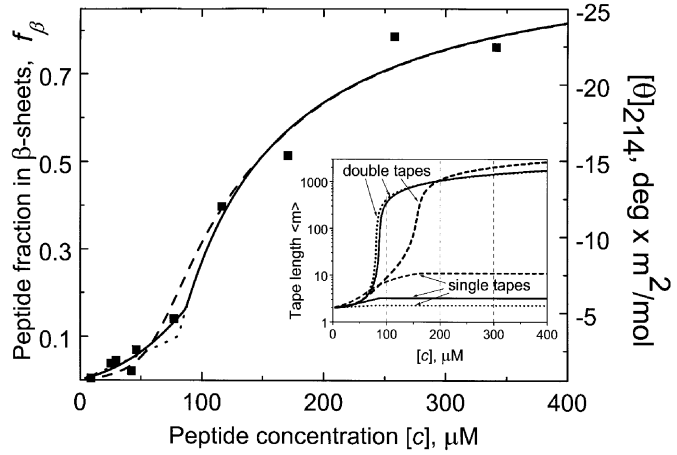

Fig. 7. Theoretical fits for the tape fraction $f_\beta(c)$ *vs.* data points $[\theta]_{214}(c)$ and the tape lengths $\langle m\rangle\,(c)$ (inset) for single and double tapes as the functions of peptide concentration, based on the model of Section 4.2 (Eqs. (11, 32)), when both single and double tapes are allowed (*cf.* Fig. 6). The values of the molecular parameters are: $E_\beta/v_0 = 5.9\times10^{-10}$; $E_{\rm tr} = 0.012$; $E_2 = 0.9$; $v_2 = 3$ Å^3 ($K_3 \simeq 2.36\times10^{-7}$) (Eq. (45), dashed line), $E_\beta/v_0 = 4.0\times10^{-9}$; $E_{\rm tr} = 0.05$; $E_2 = 0.55$; $v_2 = 80$ Å^3 ($K_3 \simeq 7.1\times10^{-7}$) (the best fit: Eq. (44), solid line), $E_\beta/v_0 = 4.4\times10^{-8}$; $E_{\rm tr} = 0.2$; $E_2 = 0.2$; $v_2 = 7$ Å^3 ($K_3 = 7.6\times10^{-7}$) (Eq. (46), dotted line). The choice of values of v_2 (or K_3) practically do not affect the shape of f_β *vs.* c curves (if $K_3 \ll 1$), however they do affect $\langle m\rangle\,(c)$ dependences. The values v_2 indicated above are the highest ones compatible with the condition (40).

the transformation energy $\varepsilon_{\rm tr}$ and the energy of the face-to-face attraction inside double tapes ε_2) we get:

$$-\varepsilon_\beta - \ln\left(v_\beta/\text{Å}^3\right) \simeq -21.3 \div -16.9\ (-19.3);$$
$$-\varepsilon_{\rm tr} \simeq -4.4 \div -1.6\ (-3.0);$$
$$\varepsilon_2 \simeq 0.1 \div 1.6\ (0.6) \qquad (49)$$

where the left values correspond to equation (45), the right ones — to equation (46) and the values in the brackets — to the best fit, equation (44). With *e.g.* $v_\beta = 0.002$ Å^3 the best fit gives $\varepsilon_\beta \simeq 25.6$ and with $v_\beta = 0.015$ Å^3 it gives $\varepsilon_\beta \simeq 23.5$ (see Eq. (B.4)).

4.3 Monomers and three types of tapes: single, double tapes and 2p-fold fibrils

Two different kinds of long self-assembling structures in solutions of DN1 peptide are visible in TEM photographs, Figure 2. In the region of relatively low concentrations semi-flexible thinner chain-like objects are formed, Figure 2b, whereas at higher concentrations thicker rod-like fibrils emerge, Figure 2a. We expect that the former (semi-flexible) structures are the double tapes, and the latter (rod-like) ones are stacks formed by several double tapes stuck together by weak attraction between the outer sides of double tapes (see Sect. 2).

This attraction between the outer sides of the double tapes is weak enough, hence fibrils cannot be formed at low concentrations. However at higher concentrations (when double tapes become very long, $m \gg 1$), it can lead to fibrillization since the total energy of attraction between long double tapes $m\varepsilon_{\rm fib}$ becomes large. In this section we take fibrils into account and estimate $\varepsilon_{\rm fib}$ (the net free energy gained during fibrillization, per a pair of interacting peptides).

Now the basic composition equation (10) takes the form, cp. equation (32):

$$c \equiv \frac{N}{V} = \frac{E_\beta}{E_{\rm tr}v_0}\Lambda + \frac{E_\beta}{v_0}\Lambda^2\frac{2-\Lambda}{(1-\Lambda)^2} + \frac{2E_\beta^2 v_2}{v_0^2}\frac{(\Lambda/E_2)^4}{(1-\Lambda)^4\left(1-(\Lambda/E_2)^2\right)} \times\left[\frac{2-(\Lambda/E_2)^2}{\left(1-(\Lambda/E_2)^2\right)} + \frac{2\Lambda}{1-\Lambda}\right] + v_2^p v_{\rm fib}^{p-1}\left(\frac{E_\beta}{v_0}\right)^{2p}\frac{2p\,(\Lambda/E_{2p})^{4p}}{(1-\Lambda)^{4p}\,(1-(\Lambda/E_{2p})^{2p})} \times\left[\frac{2-(\Lambda/E_{2p})^{2p}}{(1-(\Lambda/E_{2p})^{2p})} + \frac{2\Lambda}{1-\Lambda}\right] \qquad (50)$$

with notations (7),

$$E_2 \equiv \exp(-\varepsilon_2), \quad E_{2p} \equiv E_2 E_{\rm net},$$
$$E_{\rm net} \equiv \exp\left(-\tfrac{p-1}{2p}\varepsilon_{\rm fib}\right), \qquad (51)$$

where ε_2 is the free energy gain (per peptide) due to aggregation of two single β-sheets into a double tape, $\varepsilon_{\rm fib}$ is the net free energy gain due to sticking between the outer sides of double tapes (per pair of interacting peptides) ($\varepsilon_{2p} \equiv \varepsilon_2 + \frac{p-1}{2p}\varepsilon_{\rm fib}$, see definition (4)), p is the number of double tapes in fibrils ($p \geq 2$); v_2 is the effective attraction volume of the tapes inside a double tape and $v_{\rm fib}$ is a similar volume for double tapes inside a fibril (see Appendix B for details). In the sum (50) the terms describe the monomers, the single, the double and the $2p$-folded tapes (fibrils) correspondingly. Note that now the composition of the solutions is determined by seven independent parameters: p, E_{2p} (or $\varepsilon_{\rm fib}$) and $v_{\rm fib}$ are involved in addition to E_β/v_0, $E_{\rm tr}$, v_2, E_2 (which were present in the model of Sect. 4.2).

The critical value of the chemical activity is now

$$\Lambda_* = E_{2p}, \qquad (52)$$

which is slightly less than $E_2 < 1$ (cp. with Sect. 4.2). As before, the first (monomer) term in equation (50) always dominates over the second term (describing single tapes), since $E_{\rm tr} \ll 1$ and $\Lambda < E_{2p} \lesssim 1/2$. Hence, it is double and $2p$-fold tapes (two last terms in Eq. (50)) that are dominant β-sheet structures in the region $c/c_* - 1 \gtrsim E_{\rm tr}$.

The last term ($2p$-fold fibrils) in the r.h.s. of equation (50) is dominant at high enough concentrations. Indeed, as Λ approaches $\Lambda_* = E_{2p}$ from below, the first

three terms in the sum (50) (representing peptide concentrations in monomers, in single and in double tapes) tend to their respective asymptotic values, equation (16):

$$
\begin{aligned}
c^{(0)} &\simeq c_* = \tilde{c}_*^{(0)} \equiv \frac{E_\beta E_{2p}}{E_{\rm tr} v_0} \equiv \tilde{K}_1 \equiv \tilde{\tilde{K}}_1 E_{\rm net};\\
c^{(1)} &\simeq \tilde{c}_*^{(1)} \equiv \tilde{c}_*^{(0)} E_{\rm tr}\ E_{2p} \frac{2 - E_{2p}}{(1-E_{2p})^2};\\
c^{(2)} &\simeq \tilde{c}_*^{(2)} \simeq 2\tilde{c}_*^{(0)} \tilde{K}_3 \frac{E_{\rm net}^4 \left(2 - E_{\rm net}^2\right)}{\left(1 - E_{\rm net}^2\right)^2},\\
\tilde{K}_3 &\equiv \frac{E_\beta E_{\rm tr} v_2}{E_{2p} \left(1 - E_{2p}\right)^4 v_0}\,.
\end{aligned}
\tag{53}
$$

At the same time, the peptide concentration in fibrils increases as $\mathcal{O}\left(\left[1 - \left(\Lambda/E_{2p}\right)^{2p}\right]^{-2}\right) \sim \mathcal{O}\left(\Delta\Lambda^{-2}\right)$, equations (14, 15), where

$$
\Delta\Lambda \equiv \Lambda_* - \Lambda. \tag{54}
$$

Let us consider below the case when

$$
\varepsilon_{\rm fib} \ll 1, \tag{55}
$$

then (Eq. (51))

$$
1 - \frac{E_{2p}}{E_2} \equiv 1 - E_{\rm net} \simeq \frac{p-1}{2p}\varepsilon_{\rm fib} \ll 1, \tag{56}
$$

and $\tilde{K}_1 \simeq \tilde{\tilde{K}}_1$, $\tilde{K}_3 \simeq \tilde{\tilde{K}}_3 \equiv \left(E_\beta E_{\rm tr}/E_2\right)\left(1 - E_2\right)^{-4}\left(v_2/v_0\right)$, where $\tilde{\tilde{K}}_1$ and $\tilde{\tilde{K}}_3$ were used in Section 4.2, equations (34, 36). If Λ is close to Λ_* ($\Delta\Lambda \ll \Lambda_* \equiv E_{2p} \simeq E_2$), the mass ratio of double tapes to monomers is

$$
\frac{c^{(2)}}{c^{(0)}} \simeq \frac{2K_3}{\left(1 - \left(\Lambda/E_2\right)^2\right)^2} \simeq \frac{2K_3}{\left(\frac{p-1}{p}\varepsilon_{\rm fib} + 2\frac{\Delta\Lambda}{\Lambda_*}\right)^2}; \tag{57}
$$

$$
\frac{c^{(2)}}{c^{(0)}} \to \frac{\tilde{c}_*^{(2)}}{\tilde{c}_*^{(0)}} \equiv \left(\frac{p}{p-1}\right)^2 \frac{2K_3}{\varepsilon_{\rm fib}^2}, \quad \text{if } \Delta\Lambda \to 0; \tag{58}
$$

and the mass ratio of fibrils to double tapes is (see Eq. (50))

$$
\begin{aligned}
\frac{c^{(2p)}}{c^{(2)}} &\simeq p K_4^{p-1} \left(\frac{1 - \left(\Lambda/E_2\right)^2}{1 - \left(\Lambda/E_{2p}\right)^{2p}}\right)^2\\
&\simeq \frac{K_4^{p-1}}{p}\left(1 + \frac{p-1}{2p}\varepsilon_{\rm fib}\frac{\Lambda_*}{\Delta\Lambda}\right)^2,
\end{aligned}
\tag{59}
$$

$$
K_4 \equiv \frac{E_\beta^2 v_2 v_{\rm fib}}{\left(1 - E_2\right)^4 v_0^2}\,. \tag{60}
$$

We now suppose that

$$
K_3 \ll 1 \quad \text{and} \quad K_4 \ll 1 \tag{61}
$$

(note these unequalities are satisfied well for the DN1 peptide, as follows from Eqs. (44–48, B.3)) and that

$$
\varepsilon_{\rm fib} \lesssim \frac{p}{p-1}\left(2K_3\right)^{1/2} \tag{62}
$$

(note that it ensures that $\varepsilon_{\rm fib} \ll 1$, cp. condition (55)). Then, with increase of peptide concentration, as the chemical activity Λ approaches Λ_* from below, the dominant structure transforms first from monomers to double tapes at

$$
\Delta\Lambda \sim \Lambda_{\rm dbl} \equiv \left(\frac{K_3}{2}\right)^{1/2} \Lambda_* \sim K_3^{1/2} E_2, \tag{63}
$$

and then from double tapes to $2p$-fibrils at

$$
\Delta\Lambda \sim \Lambda_{\rm fib} \equiv K_4^{(p-1)/2}\left(\frac{p-1}{2p^{3/2}}\right)\varepsilon_{\rm fib}\Lambda_* \sim \frac{\varepsilon_{\rm fib}}{\sqrt{p}} K_4^{(p-1)/2} E_2 \tag{64}
$$

(note that

$$
\Lambda_{\rm fib} \ll \Lambda_{(2)} \ll \Lambda_{\rm dbl} \ll \Lambda_*, \quad \text{where } \Lambda_{(2)} \equiv \Lambda_* \varepsilon_{\rm fib} \frac{p-1}{2p} \tag{65}
$$

is the point at which the ratio (57) saturates). Double tapes emerge at concentration

$$
c_{\rm dbl} = \tilde{c}_*^{(0)} + \tilde{c}_*^{(1)}, \tag{66}
$$

and fibrils appear at

$$
c_{\rm fib} = \tilde{c}_*^{(0)} + \tilde{c}_*^{(1)} + \tilde{c}_*^{(2)}, \tag{67}
$$

the gap between these transitions being relatively wide due to $\tilde{c}_*^{(2)} \gtrsim c_*^{(0)} \gg c_*^{(1)}$ (see Eqs. (53, 58, 62)). Both transitions reminiscent of micelle formation in surfactant solutions: the mass concentrations of monomers and single tapes saturate at the point (66); similarly, the peptide concentration in double tapes becomes constant at $c_{\rm fib}$, where fibrils emerge.

This typical behavior (with two transitions) of the solution composition with increase of the concentration can be recognized in Figure 8a, which is plotted in accordance with the exact formula (50) using the molecular parameters from the best fit for DN1, equations (44, 48), and with $\varepsilon_{\rm fib}$ consistent with the condition (62).

Using equations (11, 55), we get for the number average lengths of the single, double and $2p$-folded tapes and their asymptotics in the region of high concentrations:

$$
\begin{aligned}
\left\langle m^{(1)}\right\rangle &= 1 + \frac{1}{1-\Lambda} \to 1 + \frac{1}{1-E_2},\\
\left\langle m^{(2)}\right\rangle &= 1 + \frac{1}{1 - \left(\Lambda/E_2\right)^2} \to \frac{E_2}{2\ \Delta\Lambda + \varepsilon_{\rm fib} E_2 (1 - p^{-1})},\\
\left\langle m^{(2p)}\right\rangle &= 1 + \frac{1}{1 - \left(\Lambda/E_{2p}\right)^{2p}} \to \frac{E_2}{2p\ \Delta\Lambda}\,.
\end{aligned}
\tag{68}
$$

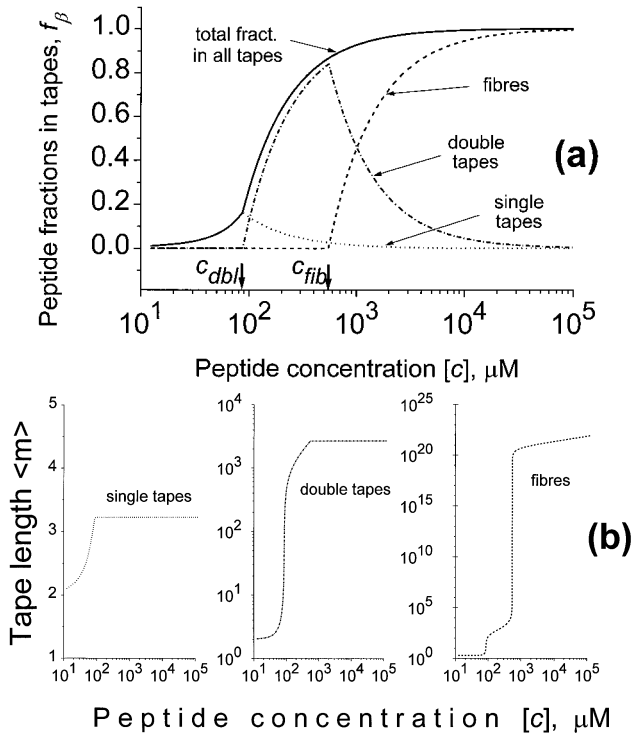

Fig. 8. (a): Concentration dependencies of the peptide fractions f_β related to various conformations: the total peptide fraction in all kinds of β-sheeted tapes (solid line) and peptide fractions in the aggregates of particular kinds (other lines) for the model of Section 4.3 (single, double and $2p$-tapes), for $p = 4$. The peptide fractions in single (classical) β-sheets is shown with a dotted line, the peptide fraction in double β-sheet tapes is shown with a dashed-and-dotted line and the peptide fraction in fibrils ($2p$-tapes) — with a dashed line. The values of the molecular parameters are: $E_\beta/v_0 = 4.0\times10^{-9}$; $E_{\rm tr} = 0.05$; $E_2 = 0.55$; $v_2 = 50$ Å^3; $\varepsilon_{\rm fib} = 5\times10^{-4}$; $v_{\rm fib} = 100$ Å^3. (b): Concentration dependencies of the average tape lengths $\langle m\rangle$, for the same system as in (a).

In particular, if conditions (61, 62) are valid: for double tapes

$$\left\langle m^{(2)}\right\rangle \simeq \frac{E_2}{2\Delta\Lambda} \simeq \left[\frac{1}{2K_3}\left(\frac{c}{c_{\rm dbl}}-1\right)\right]^{1/2}, \quad \text{if } \Delta\Lambda_{(2)} \lesssim \Delta\Lambda \ll \Lambda_* \tag{69}$$

$$\left\langle m^{(2)}\right\rangle \simeq \left(\frac{p}{p-1}\right)\frac{1}{\varepsilon_{\rm fib}}, \quad \text{if } \Delta\Lambda \lesssim \Delta\Lambda_{(2)}, \ i.e. \ \text{if } c > c_{\rm fib} \tag{70}$$

(the latter is the limiting length of double tapes, cp. Eq. (20)); and for fibrils

$$\left\langle m^{(2p)}\right\rangle \simeq \left[\frac{p}{(p-1)^2K_4^{p-1}\varepsilon_{\rm fib}^2}\left(\frac{c}{c_{\rm fib}}-1\right)\right]^{1/2}, \quad \text{if } c > c_{\rm fib} \tag{71}$$

(cp. with Fig. 8b).

Now let us compare again the results listed above with the experimental data on the DN1 peptide solutions. As before, we assume that all energetic parameters including those related to the weakest interactions (like $\varepsilon_{\rm fib}$) are virtually concentration independent.

It is known from the experiments on the DN1 solutions that in the "lower concentration region" (namely, for *ca.* $100 \div 600$ μM) it is double tapes that are visible by TEM, and at higher concentrations the fibrils are the most common structure, see Section 2. This implies the validity of the conditions (61, 62) which ensure that fibrils are unimportant in the low concentration region $c < c_{\rm fib}$. Hence, the treatment of Section 4.2 is applicable in this region. In particular, equations (44–49) defining the parameters E_β/v_0, $E_{\rm tr}$, E_2 and v_2 are still valid. Hence, from the best fit values, equation (44), we get: $[\tilde c_*^{(0)}] \simeq 73$ μM, $[\tilde c_*^{(1)}] \simeq 14$ μM, $[c_{\rm dbl}] \simeq 87\mu$M (see Eqs. (30, 43, 53, 56, 66), Fig. 8a).

From the apparent width ($D \simeq 8\div10$ nm) of the fibrils, Figure 2a, and the width of double tapes ($a \simeq 2\div2.5$ nm, see Sect. 2 or Eq. (B.1)) one can estimate p:

$$p = D/a \sim 4. \tag{72}$$

Double tapes were observed to dominate over fibrils in the concentration range up to 600 μM where fibrils emerge, this means that $c_{\rm fib} \sim 600$ μM. Hence (Eq. (67)) $[\tilde c_*^{(2)}] \sim 500$ μM, and finally (see Eq. (58)):

$$\varepsilon_{\rm fib} \simeq \frac{p}{p-1}\left(2K_3\frac{\tilde c_*^{(0)}}{\tilde c_*^{(2)}}\right)^{1/2} \sim 0.7\sqrt{K_3} < 6\times10^{-4} \tag{73}$$

where we use the threshold best fit value $\left(K_3^{\max}\right)^{\rm bf}$ for K_3, equation (47). Condition (73) ensures that double tapes dominate over the fibrils for the concentration range up to 600 μM, and that double tapes are long enough (condition (40)). Note that $\varepsilon_{\rm fib}$ is equal to the net free energy gained per an extra inter-peptide contact when fibrils are formed from double tapes; $\varepsilon_{\rm fib}$ is thus the face-to-face attraction energy (gained in such a contact) less the elastic energy cost (lost due to distortions of the straight twisted primary conformation of double tapes when they enter the fibril, cp. Ref. [10]) and less the entropic cost ($\varepsilon_{\rm entr} \sim \bar\varepsilon_{\rm att}^{(\min)}$, cp. Appendix C):

$$\varepsilon_{\rm fib} = \bar\varepsilon_{\rm att} - \varepsilon_{\rm elas} - \varepsilon_{\rm entr}. \tag{74}$$

5 Discussion and conclusions

In this paper we consider a family of self-assembling structures formed in solutions of peptide molecules forming β-sheet tapes. In addition to ordinary living polymer structure ("single chain", *i.e.* linear sequence of connected units – peptides, Fig. 4a), many-folded sheafs (fibrils) of such "single chains" (tapes) are also formed due to face-to-face attraction between the tapes, Figure 4b.

The main energetic parameters which control the behavior of the system are: the transformation energy $\varepsilon_{\rm tr}$, the β-sheet scission energy ε_β and the fibrillization energies ε_p (ε_p depends on the number of the chains in the "sheaf", p). The scission energy ε_β is the total energy of bonds between two neighboring peptides in a primary single tape. The transformation energy $\varepsilon_{\rm tr}$ is the difference in internal free energies, between a lone free peptide (an equilibrium state in solution in the limit of extremely low concentrations) and a peptide in the "rod-like" configuration which is required in order to be incorporated into a chain (tape). When a long single chain is broken into two halves, the energy increment is equal to ε_β. However, if a "bead" is cut from the end of such chain, the free energy increment is smaller: it is equal to $(\varepsilon_\beta - \varepsilon_{\rm tr})$. The fibrillization energy ε_p is the typical gain in free energy for a "bead" inside p-fibril compared to a "bead" in single chain. Thus our model includes new parameters ($\varepsilon_{\rm tr}$ and ε_p) to cover a broader range of systems, compared to classical theories of living polymers, reference [11].

The composition of such solution of self-assembling "beads" (peptides) is implicitly defined in equations (7–10). At extremely low concentrations ($c \to 0$) only lone peptides and a small amount of dimers are present (see Sect. 3.4). On the other hand, at high concentrations ($c \gtrsim c_{\rm fib} \gg c_*$), the most favourable self-assembling structures are fibrils (consisting of $p = p_*$ single tapes); their lengths increase with concentration as $c^{1/2}$, see Section 3.3. The fibrils with different aggregation numbers (p) can compete with each other at intermediate concentrations ($c_* < c < c_{\rm fib}$).

The self-assembling behaviour of the system is more rich than that of classical living polymers even if fibrils are not allowed: formation of single tapes is strongly affected by the transformation energy, $\varepsilon_{\rm tr}$ (see Sect. 4.1). For $\varepsilon_{\rm tr} \lesssim 1.5$ this transition (from lone peptides to tapes) is smooth, yet it becomes very sharp if $\varepsilon_{\rm tr} \gtrsim 4.5$: both tape fraction and length increase considerably in a very narrow concentration region, $c \simeq c_*$ (Fig. 5). The chain length increases as $(c/c_* - 1)^{1/2} \exp(\varepsilon_{\rm tr}/2)$ for $c \gtrsim c_*$ (Eq. (28)). Hence, many living polymer systems are characterized by more abrupt transitions and longer chain lengths than those predicted by classical theories based on the assumption that $\varepsilon_{\rm tr} = 0$ (see *e.g.* Ref. [11]).

If thicker aggregates (fibrils or doubled chains) are allowed in addition to single chains, then the transition from lone monomer "phase" into fibril "phase" becomes even more pronounced (see Sect. 4.2 where the case of double tapes is considered). For $c < c_*$ the fraction f_β of aggregated peptide units is low: it is proportional to $\exp(-\varepsilon_{\rm tr}) \ll 1$ for a given ratio $c/c_* < 1$; the aggregates are mostly short single tapes in this region. The situation changes drastically above c_*: here double tapes come into play. Both their concentration and their typical length increase rapidly as c increases. The length is proportional to $\exp\left(\frac{\varepsilon_{\rm tr}+\varepsilon_\beta}{2}\right)$ for a given $c/c_* > 1$, *i.e.* double tapes are much longer than single tapes which are subdominant in this region.

If three types of tapes are possible (single, double and $2p$-fold tapes, see Sect. 4.3) two consequent transitions are predicted in case when attraction between double tapes $\varepsilon_{\rm fib}$ is very weak (condition (62)). In addition to a transition at $c_{\rm dbl}$ where double tapes emerge, a second transition from double tapes to $2p$-tapes (fibrils) is predicted at $c_{\rm fib} > c_{\rm dbl}$, equations (66, 67). The latter transition is qualitatively similar to the former one. In both cases the transition region is very narrow even for $\varepsilon_{\rm tr} = 0$, if scission energy ε_β is high. The transition widths are determined respectively by the values of of the parameters $K_3 \propto \exp(-\varepsilon_\beta)$ and $(K_4^{p-1}\varepsilon_{\rm fib}^2) \propto \exp(-2(p-1)\varepsilon_\beta)$, the same parameters determine characteristic lengths of the aggregates above the transition points, see equations (63, 64, 69, 71). Concentration dependencies of the typical lengths and relative amounts of single, double tapes and $2p$-fibrils are shown in Figure 8. The width of the "window" between $c_{\rm dbl}$ and $c_{\rm fib}$ increases as attraction $\varepsilon_{\rm fib}$ between double tapes becomes weaker and/or when either scission or transformation energy decreases (see Eq. (58)).

The general approach developed in this paper is applied to a particular system — aqueous solution of DN1 peptide molecules. DN1 molecule is a short 11-residue peptide that was rationally designed in order to facilitate its self-assembling ability. The peptide tends to form long β-sheet structures (tapes) where it adopt stretched (rod-like) conformation. One side of a DN1 tape is rather hydrophobic, thus favouring tape stacking into double tapes, see Figures 1a-c. At higher concentrations DN1 forms thicker aggregates (fibrils), *i.e.* stacks of several double tapes which attract each other by their outer sides. These stacks have a finite well-defined diameter, being stabilized by the competition between the face-to-face attraction and the distortion to the equilibrium twisted configuration of the primary β-sheets (see Introduction for details).

Our theoretical analysis of available far-UV CD spectroscopy data and TEM images allows us to estimate the main microscopic energetic parameters of the system and it also yields a conclusion that the long semi-flexible chains observed in DN1 solutions at *ca.* $100 \div 600$ μM are in fact *doubled* tapes: single tapes are too short to be possibly observed in this regime.

The following values of the primary energetic parameters of DN1 water systems were obtained (see Eqs. (49)). The transformation energy $\varepsilon_{\rm tr}$ turned out to be about $3k_{\rm B}T$, hence a sharp transition associated with tape formation. The attraction between single tapes (which drives double tape formation) is rather strong as well: it amounts to about $1.2k_{\rm B}T$ per pair of interacting peptides. Hence single tapes are suppressed in favour of double tapes; this also leads to a more narrow monomer $\rightarrow$ tape transition region. What is even more important, the scission energy ε_β turned out to be extremely high: $23 \div 26k_{\rm B}T$, implying that the energetic cost of breaking a double tape is *ca.* $46 \div 52k_{\rm B}T$, hence extremely long life times of the aggregates and very slow rates of structure transformations observed in the DN1 systems (see Introduction). High scission and face-to-face attraction energies, together with

the twisted nature of the tapes and fibrils, all ensure that formation and dissociation of these aggregates must be both energetically and geometrically hindered. A more detailed discussion of the kinetics in the DN1 systems and other synthetic and native peptide systems related to them [10], including those responsible for the so called amyloidosis syndromes (incurable disorders like Alzheimer's diseases, Parkinsonism, prions, arthritises, haemodialysis, ... (Ref. [5]) will be a subject of a separate study.

The work was supported by the UK Engineering and Physical Sciences Research Council (GR/L37694 and GR/L34983). One of us (AA) wishes to thank the Royal Society for the award of a Dorothy Hodgkin Fellowship.

Appendices: Estimations of molecular parameters of DN1 peptide tapes

Appendix A: Persistence lengths of tapes and fibrils

We can estimate the persistence length $P_{(2d)}$ of worm-like objects on a plane surface as $P_{(2d)} \simeq 2x/\gamma(x)^2$ (where γ^2 is the mean square angle between tangential vectors at the points separated by the contour distance x). $P_{(2d)}$ is twice the persistence length of the same objects in three dimensions: $P_{(3d)} = P_{(2d)}/2$ (see *e.g.* [12]). From the TEM image, Figure 2a, we get: $P_{(2d)} \simeq 40 \div 135$ μm for thick aggregates (fibrils) with an average of about 70 μm; hence we estimate the persistence length of the fibrils as

$$P_{\rm fib} \sim 35\ \mu{\rm m} \quad (20 \div 70\ \mu{\rm m}). \tag{A.1}$$

Similarly, from Figure 2b the persistence length of the double tapes is estimated as

$$P_{\rm dbl} \sim 0.9\ \mu{\rm m} \quad (0.4 \div 1.8\ \mu{\rm m}). \tag{A.2}$$

We expect that the persistence length of a single β-sheet tape, $P_{\rm sng}$, should be at least twice lower. If side-chains recognize the positions of each particular side group in the neighbouring β-sheet forming a double tape, then elastic modulus of a double tape is four times that of a single tape, and thus $P_{\rm sng} \sim P_{\rm dbl}/4$. So we estimate the persistence length of a single β-sheet as

$$P_{\rm sng} \sim 0.2 \div 0.5\ \mu{\rm m}. \tag{A.3}$$

Appendix B: Effective attraction volumes/bond volumes

The peptide sizes inside β-sheet are

$$a_1 \simeq 10 \div 12.5\ \text{Å}; \quad b_1 \simeq 37\ \text{Å}; \quad b_2 \simeq 4.7\ \text{Å} \tag{B.1}$$

which are correspondingly the thickness of the β-sheet (a_1), the peptide rod length (*i.e.* the tape width) (b_1) and the periodicity along the β-sheet tape (b_2). We expect that attraction volume v_2 (which describes how freely two β-sheets combined into a double tape can move around) may be in the range of

$$v_2 \sim 1 \div 300\ \text{Å}^3 \tag{B.2}$$

assuming that $v_2 \simeq \delta a_1^{(2)} \times \delta b_1^{(2)} \times \delta b_2^{(2)}$ where $\delta a_1^{(2)} \sim 1 \div 10$ Å (the width of the attraction gap between two β-sheets), $\delta b_1^{(2)} \sim 1 \div 6.7$ Å (taking into account that the distance between the Trp and Phe side-chains responsible for the hydrophobicity of the corresponding sides of the primary DN1 β-sheets is about 6.7 Å), $\delta b_2^{(2)} \sim 1 \div 4.7$ Å. The attraction volume $v_{\rm fib}$ is a measure of the freedom available for a double tape inside a fibril, hence with a similar argument we get

$$v_{\rm fib} \sim 1 \div 600\ \text{Å}^3 \tag{B.3}$$

(note that δb_1 may be larger in this case: $\delta b_1^{(\rm fib)} \sim 1 \div 13$ Å, whereas $\delta a_1^{(\rm fib)} \sim \delta a_1^{(2)}$, $\delta b_2^{(\rm fib)} \sim \delta b_2^{(2)}$).

The volume v_β characterizes the rigidity of the bonds responsible for the β-sheet structure. These bonds also define the tape rigidity and hence the persistence length, equation (A.3). The typical bending angle between neighbouring peptides in a β-sheet is $\gamma_c \simeq (2b_2/P_{\rm sng})^{1/2} \sim 0.04 \div 0.07$, this corresponds to a side shift deformation $\delta a_1^{(\beta)} \simeq \gamma_c b_2 \sim 0.2 \div 0.3$ Å and to an average bond distortion $\delta b_2^{(\beta)} \sim \gamma_c a_1/2\sqrt{3} \sim 0.1 \div 0.16$ Å on the gap between the peptides. So the bond volume is

$$v_\beta \sim \delta a_1^{(\beta)} \times \delta b_1^{(\beta)} \times \delta b_2^{(\beta)} \sim 0.002 \div 0.015\ \text{Å}^3 \tag{B.4}$$

with $\delta b_1^{(\beta)}$ is assumed to be of the same order as $\delta a_1^{(\beta)}$ and $\delta b_2^{(\beta)}$.

Appendix C: Stability thresholds for the face-to-face attraction energy

We assume that many-folded tape structure (double tape, fibril) is sheaf-like. For the case of DN1 this assumption is supported by TEM images, Figure 2. It means that there are practically no splitting defects in the middle of the fibrils. To ensure this, the energy of the face-to-face attraction (per effective rigidity segment) stabilizing the fibril, should be high enough to counter-balance the entropic repulsion. Let us consider the corresponding stability thresholds for the attraction energies $\bar{\varepsilon}$.

The double tapes are stabilized from splitting into single β-sheets by the attraction energy $\bar{\varepsilon}_2$. The effective width of this attraction is $\delta a_2 \sim 1 \div 10$ Å (see above), the persistence length of each β-sheet is $P_{\rm sng}$, equation (A.3). For a β-sheet fragment of contour length s, the typical cotangent angle deviation is $\gamma_s \simeq (2s/P_{\rm sng})^{1/2}$, the typical shift into the side direction is $\delta a_s \sim \gamma_s s \sim \left(2s^3/P_{\rm sng}\right)^{1/2}$,

hence the typical distance between the ends of two β-sheets (which form a common double tape at the starting point) is $\sqrt{2}\delta a_s \sim 2\left(s^3/P_{\rm sng}\right)^{1/2}$. The latter value should be compared with δa_2, the corresponding threshold contour length is $s_a \sim \left(\delta a_2^2 P_{\rm sng}/4\right)^{1/3}$. Finally, the value of the face-to-face attraction per threshold fragment s_a ($\bar{\varepsilon}_2 k_{\rm B}T$ per peptide, the number of peptides is $2s_a/c$) should be compared with $k_{\rm B}T$, the result is

$$2\bar{\varepsilon}_2^{(\min)} \sim \frac{c}{\left((\delta a_2)^2 P_{\rm sng}/4\right)^{1/3}} \sim 0.1 \div 0.6. \qquad \text{(C.1)}$$

The lower value (0.1) in equation (C.1) corresponds to $\delta a_2 = 10$ Å, $P_{\rm sng} = 0.5$ μm, the higher one (0.6) — to $\delta a_2 = 1$ Å, $P_{\rm sng} = 0.2$ μm. For lower $\bar{\varepsilon}_2 < \bar{\varepsilon}_2^{(\min)}$ the structure of the doubled tape would be defective, occasionally splitting into separate single tapes. Note that $\bar{\varepsilon}_2$ is the genuine attraction energy per a peptide. The net (effective) attraction energy $\varepsilon_2 = \varepsilon_{\rm dbl}/2$ used in the main text (which is the difference between the free energies of two single tapes and a double tape of the same length divided by the number of peptides involved) includes the entropic contribution which was estimated above:

$$\varepsilon_2 = \bar{\varepsilon}_2 - \varepsilon_2^{\rm entr}, \quad \varepsilon_2^{\rm entr} \sim \bar{\varepsilon}_2^{(\min)}. \qquad \text{(C.2)}$$

The DN1 fibrils are stabilized from splitting into separate double tapes by the attraction energy $\bar{\varepsilon}_{\rm att}$. The net fibrillization energy $\varepsilon_{\rm fib}$ (Eq. (51)) is this attraction energy less elastic and entropy costs of the fibril formation. The elastic energy cost was calculated in a separate paper, reference [10]. Let us estimate the entropic energy. The effective width of this attraction is $\delta a_{\rm fib} \sim 1 \div 10$ Å (see above), the persistence length of the double tapes is $P_{\rm dbl}$, equation (A.2). Similarly to what was done above for double tapes, one can get for the stability threshold value for the face-to-face attraction between double tapes forming a *thick* ($p \gg 1$) fibril:

$$\bar{\varepsilon}_{\rm att}^{(\min)} \sim \frac{c}{\left((\delta a_{\rm fib})^2 P_{\rm dbl}/2\right)^{1/3}} \sim 0.07 \div 0.3. \qquad \text{(C.3)}$$

(Note that for the case of fibrils, $\bar{\varepsilon}_{\rm att}$ is the energy per contact between double tapes, cp. Eqs. (50, 51) and Eq. (32). Also note that when a double tape separates from a thick fibril, the latter can be considered as nearly a straight line.)

References

1. T.E. Creighton, *Proteins. Structures and molecular properties* (W.H. Freeman and Co: N.Y., 1993).
2. M.T. Krejchi, E.D.T. Atkins, A.J. Waddon, *et al.*, Science **265**, 1427 (1994).
3. M.R. Ghadiri, J.R. Granja, R.A. Milligan, *et al.*, Nature **366**, 324 (1993).
4. S. Zhang, M. Altman, Reactive Functional Polymers **41**, 91-102 (1999).
5. S.Y. Tan, M.B. Pepys, Histopathology **25**, 403-414 (1994).
6. J.A. Zarutskie, A.K. Sato, M.M. Rushe, *et al.*, Biochemistry **38**, 5878-5887 (1999).
7. A. Aggeli, M. Bell, N. Boden, *et al.* Nature **386**, 259-262 (1997).
8. A. Aggeli, M. Bell, N. Boden, *et al.*, J. Mater. Chem. **7**, 1135-1145 (1997).
9. A. Aggeli, I.A. Nyrkova, M. Bell, *et al.*, Proc. Natl. Acad. Sci. USA (2000), submitted.
10. I.A. Nyrkova, A.N. Semenov, A. Aggeli, N. Boden, Eur. Phys. J. B **17**, 481 (2000).
11. J.N. Israelachvili, *Intermolecular and Surface Forces* (Academic Press, NY, 1985).
12. A.Yu. Grosberg, A.R. Khokhlov, *Statistical Physics of Macromolecule* (New York: AIP Press, 1994).