



Pergamon

Insect Biochemistry and Molecular Biology 31 (2001) 877–885

*Insect  
Biochemistry  
and  
Molecular  
Biology*

www.elsevier.com/locate/ibmb

## “Soft”-cuticle protein secondary structure as revealed by FT-Raman, ATR FT-IR and CD spectroscopy

Vassiliki A. Iconomidou <sup>a</sup>, Georgios D. Chryssikos <sup>b</sup>, Vassilis Gionis <sup>b</sup>, Judith H. Willis <sup>c</sup>, Stavros J. Hamodrakas <sup>a,\*</sup>

<sup>a</sup> Department of Cell Biology and Biophysics, Faculty of Biology, University of Athens, Athens 157 01, Greece

<sup>b</sup> Theoretical and Physical Chemistry Institute, National Hellenic Research Foundation, Athens 116 35, Greece

<sup>c</sup> Department of Cellular Biology, University of Georgia, Athens, GA 30602, USA

Received 15 September 2000; received in revised form 20 December 2000; accepted 21 December 2000

### Abstract

The nature of the interaction of insect cuticular proteins and chitin is unknown even though about half of the cuticular proteins sequenced thus far share a consensus region that has been predicted to be the site of chitin binding. We previously predicted the preponderance of a  $\beta$ -pleated sheet in the consensus region and proposed its responsibility for the formation of helicoidal cuticle (Iconomidou et al., *Insect Biochem. Mol. Biol.* 29 (1999) 285). In this study, we examined experimentally the secondary structure of intact and guanidine hydrochloride extracted cuticle and the cuticular protein extract. The studied cuticle came from the larval dorsal abdomen of the lepidopteran *Hyalophora cecropia*, a classical example of “soft” cuticle. Analysis with FT-Raman, ATR FT-IR and CD spectroscopy indicates that antiparallel  $\beta$ -pleated sheet is the predominant molecular conformation of “soft-cuticle” proteins both in situ in the cuticle and following extraction. It seems that this conformation dictates the modes of chitin–protein interaction in cuticle, in agreement with earlier proposals (Atkins, *J. Biosci.* 8 (1985) 375). © 2001 Elsevier Science Ltd. All rights reserved.

**Keywords:** “Soft-cuticle” proteins; Chitin; Antiparallel  $\beta$ -pleated sheet; FT-Raman; ATR FT-IR; CD

### 1. Introduction

The insect exoskeleton or cuticle is a composite of proteins and chitin that provides protective, locomotive and structural functions. Little information, however, is available about the molecular structure of this important complex that exhibits a helicoidal architecture (Neville, 1975).

Extracts of a single region of cuticle may contain 50 or more electrophoretic bands. More than 100 cuticular proteins of known sequences are available to date from insects of five different orders as well as an arachnid and diverse crustaceae (Willis, 1999). Over half bear a motif similar to that first identified by Rebers and Riddiford (1988) in seven cuticular proteins: G-x(8)-G-x(6)-Y-x-

A-x-E-x-G-Y-x(7)-P-x(2)-P (where x represents any amino acid, and the value in parentheses indicates the number of residues). Additional sequences revealed minor differences and new conserved residues in this R&R Consensus. A modification of this consensus, G-x(7)-[DEN]-G-x(6)-[FY]-x-A-[DGN]-x(2,3)-G-[FY]-x-[AP]-x(6), is present in about 40 unique insect cuticular proteins (Willis, 1999). None of these shows identity in the consensus region. This motif has been used to identify at least 50 more putative cuticular protein sequences from the *Drosophila* genome project (Rubin et al., 2000) and from a *Bombyx* EST project (<http://www.ab.a.u-tokyo.ac.jp/silkbase/>). These sequences can be obtained at the following websites, using CUTICLE as the search protein domain: <http://www.fruitfly.org/annot/> and <http://www.ab.a.u-tokyo.ac.jp/silkbase/>.

Rebers and Riddiford (1988) suggested that the consensus would turn out to be a region of structural importance. Subsequently, Andersen et al. (1995) postulated that the motif might be involved in protein–chitin inter-

\* Corresponding author. Tel.: +30-1-727-4545; fax: +30-1-727-4742.

E-mail address: shamodr@cc.uoa.gr (S.J. Hamodrakas).

action. The region N-terminal to the R&R Consensus is enriched in hydrophilic amino acids (Lampe and Willis, 1994; Andersen et al., 1995). In total, a stretch of approximately 68 amino acids appears to be conserved, an “extended R&R Consensus” (Iconomidou et al., 1999). This extended consensus is sufficient to confer chitin binding to a fusion protein made with glutathione *S*-transferase (J.F. Rebers and J.H. Willis, personal communication).

Previous experimental studies concerning the conformation of the cuticular proteins were not conclusive (Hackman and Goldberg, 1979; Hillerton and Vincent, 1979; Stephens and Vincent, 1976). Recently, we presented secondary structure prediction data indicating that a  $\beta$ -pleated sheet is most probably the underlying molecular conformation of a large part of this extended R&R consensus, especially the part which contains the R&R consensus itself. We also proposed that this conformation is essential in defining cuticle’s helicoidal architecture and most probably is involved in specific interactions of the cuticular proteins with the chitin filaments. We also suggested that perhaps a universal architectural plan for helicoidal extracellular structures exists which is based primarily on the packing interactions of  $\beta$ -pleated sheets (Iconomidou et al., 1999). To date, there has been no systematic experimental analysis of the secondary structure of cuticular proteins. In this study, we present experimental data for the molecular conformation of proteins from “soft” cuticle utilizing FT-Raman, ATR FT-IR and CD spectroscopy, which suggest that the dominant conformation of “soft”-cuticle proteins is the *antiparallel*  $\beta$ -pleated sheet.

## 2. Materials and methods

### 2.1. Cuticular proteins—cuticle samples

The protein HCCP12 chosen for inclusion in the secondary structure prediction analysis is a “soft”-cuticle representative insect cuticular protein. It is a major constituent of cuticle extracts and has been associated with cuticles of all three metamorphic stages, L(larva), P(pupa) or A(adult) of the giant silkworm *Hyalophora cecropia* [1169129 (ENTREZ)] (Binger and Willis, 1994).

Two mid-fifth instar larvae of *H. cecropia* were frozen, thawed slightly, then the dorsal abdominal cuticle removed, cleaned of all adhering cells and hard tubercles, and lyophilized. An aliquot was saved as unextracted cuticle. Cuticle proteins were extracted by vigorous shaking with an extraction buffer (8 M guanidine hydrochloride, 1 mM PMSF, 1% MSH and 100 mM Tris, pH 8.9) for 2 h at 60°C. This procedure removed 44% of the dry weight of the cuticle. Comparable extracts have revealed 51 electrophoretic bands on isoe-

lectric focusing gels, 13 of them strong, with HCCP12 one of the strongest (Cox and Willis, 1985). The extract was dialyzed thoroughly against water and spun in a clinical centrifuge. The supernatant (“extracted” proteins) and precipitated flocculent material (“water-insoluble material”) were lyophilized separately. The “extracted” cuticle was rinsed in water and lyophilized. The cuticle proteins used in the FT-Raman, ATR FT-IR and CD spectroscopy experiments mentioned below were the “extracted” proteins. The “water-insoluble material” gave almost identical FT-Raman and ATR FT-IR spectra to the “extracted” proteins (data not shown).

### 2.2. Secondary structure prediction

Secondary structure prediction was done as described previously (Iconomidou et al., 1999).

### 2.3. FT-Raman spectroscopy

Sample preparation for Raman measurements involved pressing pieces of cuticle samples (intact and “extracted”) into the 2-mm cavity of a standard aluminum holder. The same procedure was followed for the lyophilized extracted “soft”-cuticle proteins.

When performing the experiments for the intact and “extracted” cuticle, great care was taken to illuminate the inner part of the cuticle (endocuticle) and not the outer surface of the cuticle closest to the epicuticle (which is sclerotized and contains proteins covalently bonded; these, presumably, are not removed by the extraction procedure described above).

Raman spectra were obtained on a Fourier-transform instrument (Bruker RFS 100) employing for excitation ca. 400 mW of the Nd:YAG 1064-nm line in a back-scattering geometry. The samples are relatively poor Raman scatterers and often exhibit strong fluorescence upon laser excitation. Working in the near-infrared greatly reduces the fluorescence of the sample and eliminates the need for prolonged laser annealing. The resolution was 8  $\text{cm}^{-1}$ , a zero-filling factor of 2 was employed, and the total acquisition time was ca. 10 h (10 000 scans). Attempts to collect data at higher resolution resulted in the deterioration of the S/N ratio, without significant improvement in band resolution. The interferograms were Fourier-transformed in 1-h acquisition time segments, in order to allow for the detection of time-dependent phenomena (sample degradation, luminescence, bleaching, etc.) and averaged. Subsequently, the standard deviation ( $\sigma$ ) of the experiment was calculated. All spectral data are shown with a  $2\sigma$  error bar.

### 2.4. Attenuated total reflectance infrared spectroscopy (ATR-IR)

Infrared spectra were obtained at a resolution of 2  $\text{cm}^{-1}$  and a zero-filling factor of 2, on a Fourier-trans-

form vacuum instrument (Bruker IFS 113v) equipped with a 45° KRS5 Attenuated Total Reflectance accessory. Internal reflection spectroscopy presents several advantages compared to the more common KBr dispersion technique (de Jongh et al., 1996). Having a penetration depth of ca. 2  $\mu\text{m}$  (1000  $\text{cm}^{-1}$ , KRS5, 45°), ATR is free of saturation effects which may be present in the transmission spectra of thicker samples. Furthermore, the choice of the ATR technique was dictated by the need to exclude any possible spectroscopic and chemical interactions between the sample and the dispersing medium. In the literature, ATR-IR has been employed to investigate solutions of proteins, or thin films deposited directly on the ATR crystal by evaporation of the solvent (de Jongh et al., 1996; Fu et al., 1994; Fink and Gendreau, 1984; Singh and Fuller, 1991; Jakobsen and Wasacz, 1990; Kim et al., 1998). In the present investigation, pieces of lyophilized “soft”-cuticle protein were placed on the surface of the KRS5 element and kept in place by tightening the metal plate holder. Fifteen 400-scan spectra were collected and averaged to improve the S/N ratio. The spectra were corrected for the effect of wavelength on the penetration depth ( $p.d. \propto \lambda$ ). The corresponding effect of the (frequency-dependent) refractive index ( $n$ ) of cuticular protein samples was not taken into account due to the lack of relevant data.

#### 2.4.1. Post-run computations of the spectra

The Raman scattering and infrared ATR absorption peak maxima were determined from the minima in the second derivative of the corresponding spectra. Derivatives were computed analytically using routines of the OPUS/OS2 software and included 9-point smoothing by the Savitzky–Golay algorithm (Savitsky and Golay, 1964). Taking into consideration the resolution and zero-filling-factors employed, smoothing was effective over frequency ranges of ca. 20  $\text{cm}^{-1}$  (Raman) and 5  $\text{cm}^{-1}$  (ATR infrared).

#### 2.5. Circular dichroism (CD) spectroscopy

Samples of *H. cecropia* larval-extracted cuticular proteins were solubilized in distilled water (pH 5.5) and measured at 4°C on a Jasco J-710 spectropolarimeter in a cell with a 0.1-cm path length. Fractions of secondary structure were estimated using the secondary structure estimation program (SSEAX) of Jasco Corp. Spectra of four standard proteins were provided for the calculation of the basis spectra. Each observed spectrum was expressed as a linear combination of reference spectra of four types of secondary structure:  $\alpha$ -helix ( $\alpha$ ),  $\beta$ -sheet ( $\beta$ ),  $\beta$ -turn (t) and coil (c, random). The linear coefficients representing secondary structure estimates were calculated by a linear least-squares method. The reference spectra of the four types of secondary structure were calculated from the CD spectra of four proteins

with crystallographically known secondary structures: lysozyme (46%  $\alpha$ , 19%  $\beta$ , 23% t, 12% c), papain (28%  $\alpha$ , 29%  $\beta$ , 18% t, 25% c), ribonuclease A (23%  $\alpha$ , 46%  $\beta$ , 21% t, 10% c) and  $\alpha$ -chymotrypsin (11%  $\alpha$ , 50%  $\beta$ , 25% t, 14% c). The percentages of observed (crystallographically) secondary structure were taken from Levitt and Greer (1977). This set gave the best fit to the data.

### 3. Results

#### 3.1. Secondary structure prediction

The results of secondary structure prediction on the representative “soft”-cuticular protein HCCP12 are shown in Fig. 1. Prediction of secondary structure for HCCP12 was made according to Hamodrakas, 1988. The black shadowed parts of the protein sequence correspond to predicted  $\beta$ -strands. The R&R consensus is identified by the arrow below the sequence of HCCP12. The protein has 89 nominal residues in total. The R&R Consensus begins at nominal position 56.

The result of prediction of another popular prediction algorithm, PHD (Rost and Sander 1993, 1994; Rost, 1996), on HCCP12 is also displayed below the protein sequence for comparison. The symbol E represents the predicted secondary structure of a  $\beta$ -pleated sheet in this case, whereas gaps correspond to random coils or  $\beta$ -turns/loops. This algorithm claims to achieve prediction of protein secondary structure at better than 70% accuracy (Rost and Sander, 1993).

Our prediction package and the PHD package clearly suggest that not only the extended R&R domain (Iconomidou et al., 1999) but also the whole sequence of HCCP12 has a considerable proportion of  $\beta$ -pleated sheet structure and total absence of  $\alpha$ -helix.

The  $\beta$ -strands are predicted at nominal positions 10–13, 19–21, 24–28, 32–35, 41–44, 52–55, 58–61 and 66–73 (relevant amino acid sequence of these segments: QILK, IGV, FQYGY, NGIQ, QLNN, IEVR, FSYV and VTYSVTYT, respectively).

#### 3.2. FT-Raman spectroscopy

The Raman spectra obtained from *H. cecropia* larval dorsal abdomen (LDA) intact, “extracted” cuticle, as well as and the corresponding weighted difference spectrum, are shown in Figs. 2–4, respectively. The spectrum of the proteins extracted from *H. cecropia* LDA cuticle is shown in Fig. 5. The band maxima of the cuticle and protein spectra, determined on the basis of their second derivatives, are compiled in Table 1, together with the tentative assignments of the main spectral features.

The Raman spectra of both the intact and the “extracted” *H. cecropia* LDA cuticle are dominated by

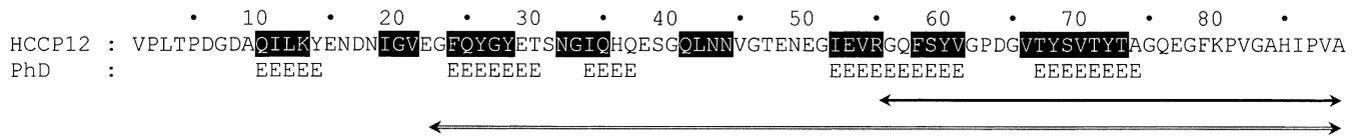


Fig. 1. Predicted secondary structure of the “soft”-cuticle representative insect cuticular protein HCCP12. This protein has been associated with cuticles of all three metamorphic stages, L(larva), P(pupa) or A(adult) of *Hyalophora cecropia* (ENTREZ accession number 1169129) (Binger and Willis, 1994). It is definitely a major cuticular protein of larval dorsal abdomen (LDA) of *Hyalophora cecropia*. The sequence of the protein is given in the one-letter code. It has 89 nominal positions marked at the top (Binger and Willis, 1994). Black-boxed residues denote predicted  $\beta$ -strands according to Hamodrakas (1988). Predicted secondary structure according to another popular prediction algorithm, PHD (Rost and Sander 1993, 1994; Rost, 1996), is also displayed below the sequence for comparison. The symbol E represents predicted secondary structure of  $\beta$ -strands in this case, whereas gaps correspond to random coil or  $\beta$ -turns/loops. The R&R consensus (Rebers and Riddiford, 1988) is underlined, whereas the “extended” R&R consensus (Iconomidou et al., 1999) is doubly underlined below the sequence.

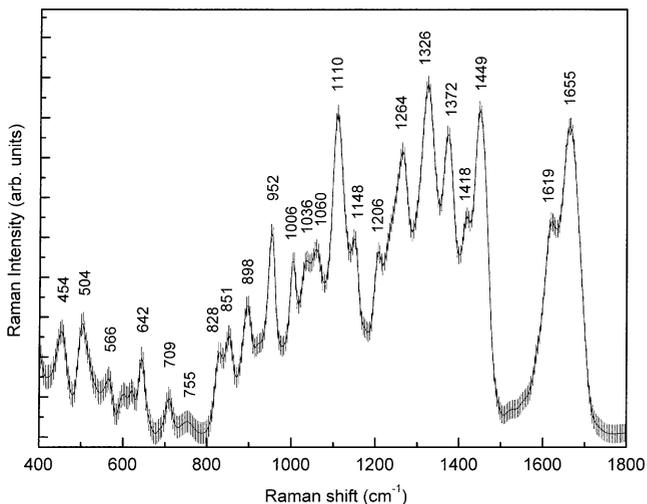


Fig. 2. FT-Raman spectrum of *Hyalophora cecropia* larval dorsal abdomen (LDA) intact cuticle.

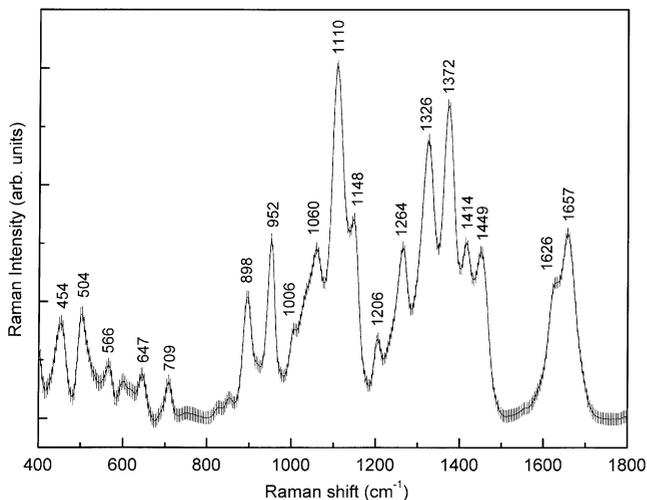


Fig. 3. FT-Raman spectrum of *Hyalophora cecropia* larval dorsal abdomen (LDA) cuticle from which proteins have been extracted to a great extent (see Section 2). The cuticle is called “extracted” for brevity. The spectrum results almost entirely from chitin. Protein contribution is negligible.

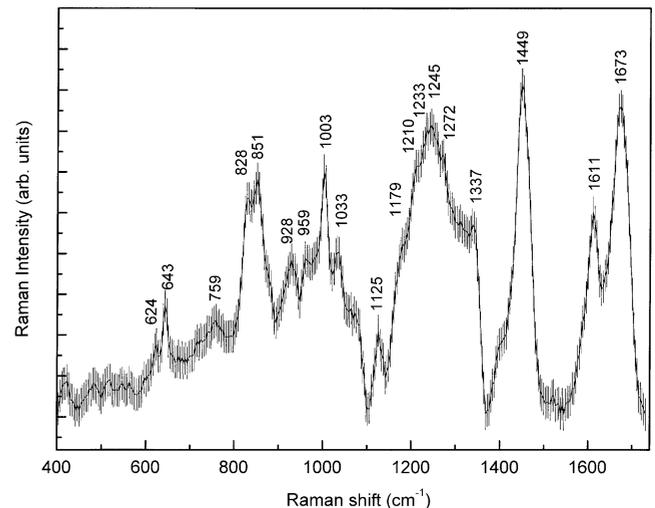


Fig. 4. Difference spectrum of the Raman spectra presented in Figs. 2 and 3. The spectra prior to subtraction were normalized to the 952  $\text{cm}^{-1}$  band intensity due to chitin. An almost identical difference spectrum resulted when normalization prior to subtraction was done, based on the intensity of the 709  $\text{cm}^{-1}$  band, which is also due to chitin. The net result, in both cases, is a difference spectrum almost entirely due to protein (compare with Fig. 5).

the contribution of bands due to chitin (Figs. 2 and 3 and Table 1).

However, certain features of the FT-Raman spectrum of intact cuticle signify the presence of proteins. The bands at 755  $\text{cm}^{-1}$  due to Trp and the Tyr doublet at 828, 851  $\text{cm}^{-1}$  are characteristic. Also, the band at 642  $\text{cm}^{-1}$  is mostly due to Phe.

All other low-wavenumber modes in the spectra of the intact and “extracted” LDA cuticle at 454, 504, 566, 647 and 709  $\text{cm}^{-1}$  can be attributed mainly to  $\alpha$ -chitin (Figs. 2 and 3, Table 1; Mikkelsen et al., 1997, and references cited therein). The band at 898  $\text{cm}^{-1}$  is known to be sensitive to the anomeric configuration (typical for  $\beta$ -linked carbohydrates) (Cael et al., 1974). Several of the bands in the two spectra (952, 1060, 1110, 1148, 1206, 1264, 1326, 1372, 1414 and 1449  $\text{cm}^{-1}$ ) also have a major contribution from  $\alpha$ -chitin vibrations.

In the carbonyl region we find a broad Raman band with a center at  $\sim 1656 \text{ cm}^{-1}$  and a shoulder at  $\sim 1620$

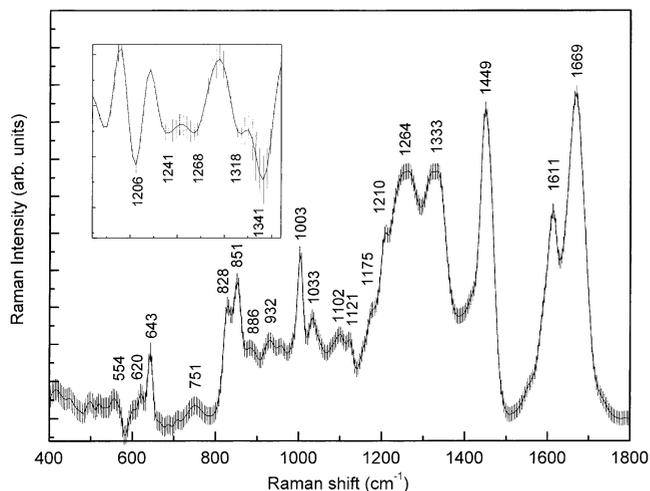


Fig. 5. FT-Raman spectrum of the cuticular proteins extracted from *Hyalophora cecropia* larval dorsal abdomen (LDA) cuticle (see Section 2). The second derivative of the spectrum in the amide III region is included in the inset.

$\text{cm}^{-1}$ . This split of the secondary amide I band into two components is presumably caused by the statistical mixture of the hydrogen bonding pattern of the primary hydroxyl group in  $\alpha$ -chitin (Minke and Blackwell, 1978). The consequence of the proposed structure of  $\alpha$ -chitin (Minke and Blackwell, 1978) is two different types of amide groups: one which is only involved in one intermolecular hydrogen bond ( $\text{C}=\text{O}\cdots\text{H}-\text{N}$ ) and is responsible for the  $\sim 1656 \text{ cm}^{-1}$  mode, and one which is bifurcated by forming an additional intramolecular hydrogen bond to the primary hydroxyl group leading to a lowering of the amide I frequency (Mikkelsen et al., 1997, and references cited therein). However, the amide I region in the spectrum of the intact cuticle (Fig. 2) has a contribution from vibrations due to the protein components of LDA cuticle as well.

The protein contribution to the spectrum of intact cuticle can be revealed by subtracting the spectrum of the “extracted” cuticle after scaling the two spectra in the regions where chitin is the only vibrationally active component. The difference spectrum is shown in Fig. 4. The spectral subtraction is based on the assumption that the chitin of the cuticle does not undergo structural changes during the protein extraction treatment. The validity of this assumption is supported by the absence of chitinic spectral bands in the spectrum of Fig. 4, at least within the resolution of our experiment (see also Table 1). The comparison of the spectra of Figs. 4 and 5 (protein by spectral subtraction and extraction, respectively) reveals striking similarities, suggesting that the former gives a reliable physical picture of the cuticle protein vibrations in the native (in vivo) state.

No significant bands below  $500 \text{ cm}^{-1}$  were observed in these protein spectra (Figs. 4 and 5). Several bands are clearly due to side-chain vibrations of amino acid

residues, especially to the aromatic ring-containing Tyr, Trp and Phe. More specifically, the bands at  $1033$ ,  $1003$ , and  $624(620) \text{ cm}^{-1}$  are ascribable to Phe, whereas the bands at  $643$ ,  $828$ ,  $851$ ,  $1179$ ,  $1210$  and  $1611 \text{ cm}^{-1}$  are most probably due to Tyr (Frushour and Koenig, 1975; Yu, 1977; Spiro and Gaber, 1977). The intensity ratio of the tyrosine doublet at  $851$  and  $828 \text{ cm}^{-1}$ ,  $R=I_{853}/I_{830}$ , is sensitive to the nature of hydrogen bonding, or to the state of the ionization of the phenolic hydroxyl group, and has been used to identify “buried” and “exposed” Tyr moieties (Carey, 1982, and references cited therein). The  $643$  and  $1210 \text{ cm}^{-1}$  bands may also involve vibrations of Phe. Also, the band at  $1611 \text{ cm}^{-1}$  may hide components due to Phe ( $1609 \text{ cm}^{-1}$ ) and Trp. Bands at  $759$  ( $751 \text{ cm}^{-1}$ ) are ascribable to Trp (Frushour and Koenig, 1975; Yu, 1977; Spiro and Gaber, 1977). The absence of a band at  $1361 \text{ cm}^{-1}$  suggests that the Trp side chains are not “exposed” (Yu, 1977).

Bands in the  $500\text{--}550 \text{ cm}^{-1}$  region are typically associated with the S–S stretching mode of the C–C–S–S–C–C structural unit of disulfide bonds (Frushour and Koenig, 1975; Yu, 1977; Spiro and Gaber, 1977). Some minor features are seen in this region in the Raman spectra of *H. cecropia* LDA cuticle proteins. We have not attempted to assign them to specific S–S stretching modes, as it is known that cuticle proteins are devoid of S–S bonds (Andersen et al., 1995).

Armed with the identification of the vibrational signatures of the aromatic residues, we can now focus our attention on the “amide” band envelopes, which are well-known indicators of secondary structure features in proteins and polypeptides. Several papers review the relevant assignments of the Raman (Frushour and Koenig, 1975; Yu, 1977; Spiro and Gaber, 1977; Carey and Salares, 1980; Carey, 1982) and IR spectra (Surewicz et al., 1993; Haris and Chapman, 1995; Jackson and Mantsch, 1995; Griebenow et al., 1999) in the amide I and III regions.

In the amide I region ( $1600\text{--}1700 \text{ cm}^{-1}$ ) the Raman spectrum of the extracted *H. cecropia* LDA cuticle proteins exhibits a well-defined maximum at  $1669 \text{ cm}^{-1}$ , typical of  $\beta$ -sheet structure. Its half-width of  $\sim 35 \text{ cm}^{-1}$  perhaps reflects the uniformity or narrow distribution of chain conformations of the  $\beta$ -sheets in cuticular proteins. (Note the presence of a similarly sharp band at  $1673 \text{ cm}^{-1}$  in the difference spectrum of Fig. 4.) The absence of bands at ca.  $1650 \text{ cm}^{-1}$  indicates that  $\alpha$ -helical structures are not favored, or remain below the detection limit of our techniques.

The amide III range ( $1230\text{--}1320 \text{ cm}^{-1}$ ) is relatively free from side-group vibrations and, thus, highly diagnostic of secondary structure. The amide III band of the FT-Raman spectrum of *H. cecropia* LDA extracted proteins shows a doublet at  $1241$  and  $1268 \text{ cm}^{-1}$  (Fig. 5, inset; Table 1). The  $1241 \text{ cm}^{-1}$  is assigned to a  $\beta$ -sheet, and the  $1268 \text{ cm}^{-1}$  band to  $\beta$ -turns or coil. The differ-

Table 1

Main Raman (1800–400  $\text{cm}^{-1}$ ) peak maxima of (a) *H. cecropia* larval dorsal abdomen intact soft cuticle, (b) soft cuticle after protein extraction, (c) the difference spectrum (a)–(b) and (d) the extracted proteins. Tentative assignments are included. For details see text

Intact cuticle	Cuticle after protein extraction	Difference	Extracted protein	Assignments
1655	1657	1673	1669	$\beta$ -sheet
1619	1626 sh			<i>chitin</i>
		1611	1611	Tyr, Phe
1449	1449	1449	1449	$\text{CH}_2$ deformation
1418	1414			<i>chitin</i>
1372	1372			<i>chitin</i>
		1337	1341	CH deformation
1326	1326			<i>chitin</i>
			1318	CH deformation
		1272	1268	$\beta$ -turns? coil?
1264	1264			<i>chitin</i>
		1245, 1233	1241	$\beta$ -sheet
		1210	1206	Tyr, Phe
1206	1206			Tyr, Phe/ <i>chitin</i>
		1179	1175	Tyr
1148	1148			<i>chitin</i>
		1125	1121	C–N stretching
1110	1110			<i>chitin</i>
			1102	C–N stretching
1060	1060			C–N stretching
1036		1033	1033	Phe
1006		1003	1003	Phe
		959		
952	952			<i>chitin</i>
		928	932	C–C stretching
898	898			<i>chitin</i>
			886	Trp
851		851	851	Tyr
828		828	828	Tyr
755		759	751	Trp
709	709			<i>chitin</i>
642	647	643	643	Tyr, Phe
		624	620	Phe
566	566			<i>chitin</i>
			554	
		520	520	
504	504			

ence spectrum at the amide III region exhibits three components at 1233, 1245 and 1272  $\text{cm}^{-1}$ . The splitting of the  $\beta$ -sheet mode at ca. 1240  $\text{cm}^{-1}$  (1233, 1245  $\text{cm}^{-1}$ ) is not uncommon (Singh et al., 1990). The band at 1272  $\text{cm}^{-1}$  is associated with the presence of unordered structure or  $\beta$ -turns. A list of amide I and III wavenumbers of some protein and polypeptide structures (fibrous and globular) containing  $\beta$ -sheets is given by Hamodrakas et al. (1982).

### 3.3. ATR FT-IR spectroscopy

The ATR FT-IR spectrum of the cuticular proteins extracted from *H. cecropia* LDA cuticle shown in Fig. 6 suggests clearly that a  $\beta$ -sheet is the predominant structure of “soft”-cuticle proteins in the extracted dry

state. The strong amide I band at 1624  $\text{cm}^{-1}$  and strong amide III band at 1227  $\text{cm}^{-1}$  are definitely due to a  $\beta$ -sheet (Krimm and Bandekar, 1986; Surewicz et al., 1993; Haris and Chapman, 1995; Jackson and Mantsch, 1995; Cai and Singh, 1999). Shoulders at 1655  $\text{cm}^{-1}$  in the amide I, 1542  $\text{cm}^{-1}$  in the amide II and 1267  $\text{cm}^{-1}$  in the amide III bands are probably due to a random coil structure (references as above). There is some disagreement in the literature about the origin ( $\beta$ -turns or high-frequency component of antiparallel  $\beta$ -sheet) of bands in the 1670–1695  $\text{cm}^{-1}$  region (Haris and Chapman, 1995, and references cited therein). The high-frequency 1691  $\text{cm}^{-1}$  band (seen as a shoulder in the ATR spectrum) requires special attention. This band is either due to  $\beta$ -turns (probably joining strands of  $\beta$ -sheets, consequently part of the  $\beta$ -sheets themselves) or can be

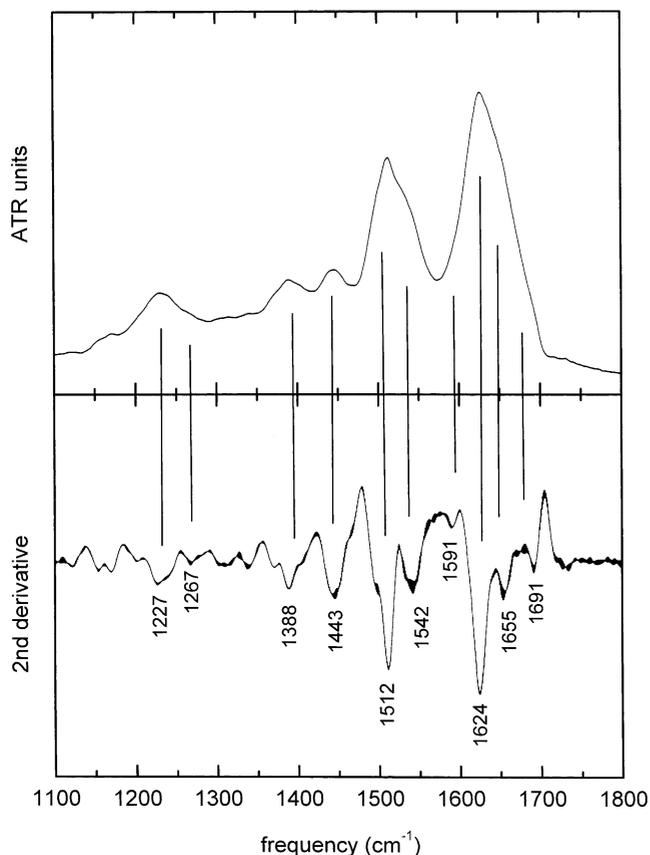


Fig. 6. Attenuated total reflectance FT-infrared (ATR FT-IR) spectrum of the cuticular proteins extracted from *Hyalophora cecropia* larval dorsal abdomen (LDA) cuticle. The second derivative spectrum is included.

assigned to an antiparallel  $\beta$ -sheet. In proteins containing antiparallel  $\beta$ -sheets, a high-frequency  $\beta$ -sheet component that arises from transition dipole coupling is usually found at 50–70  $\text{cm}^{-1}$  higher than the main  $\beta$ -sheet component (Jackson and Mantsch, 1995, and references cited therein). Bands at ca. 1510  $\text{cm}^{-1}$  are usually assigned to Tyr ring vibrations. However, it is interesting to note that the strong observed amide II band for poly-Gly I, an antiparallel  $\beta$ -pleated sheet structure, appears at 1517  $\text{cm}^{-1}$  (Krimm and Bandekar, 1986). The strong band at 1512  $\text{cm}^{-1}$  probably has contributions from both Tyr and  $\beta$ -sheet. Overall, the results obtained from the analysis of the ATR FT-IR spectra are in good agreement with the results obtained from the FT-Raman spectra of the extracted *H. cecropia* LDA cuticle proteins in the dry state.

### 3.4. CD spectroscopy

The CD spectrum of *H. cecropia* extracted LDA cuticular proteins solubilized in distilled water, pH 5.5 at 4°C, is shown in Fig. 7. It exhibits a large negative band at approximately 206 nm and other prominent negative bands at 216, 221 and 235 nm.

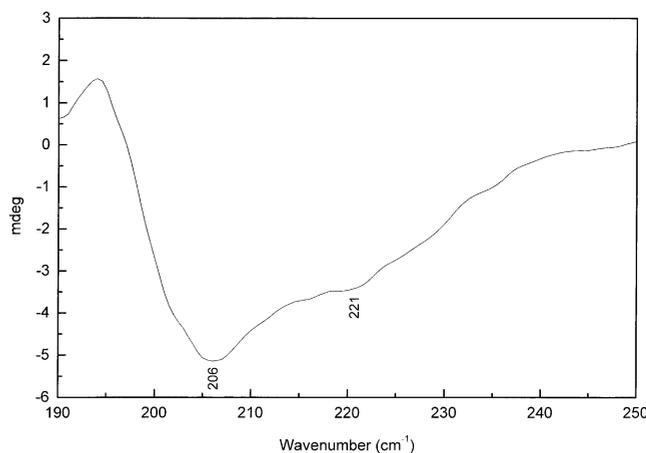


Fig. 7. CD spectrum of the cuticular proteins extracted from *Hyalophora cecropia* larval dorsal abdomen (LDA) cuticle, solubilized in distilled water (pH 5.5) at 4°C. Instrumental conditions: path length 0.01 cm, scanning speed 50 nm/min, sensitivity 5 mdeg, bandwidth 1 nm, resolution 0.5 nm.

It shows similarities with CD spectra obtained from a distinct class of all  $\beta$ -proteins (Manavalan and Johnson, 1983) and especially with the spectrum of  $\alpha$ -chymotrypsin. Some similarities to spectra obtained from  $\alpha$ + $\beta$  proteins (such as ribonuclease A) where  $\beta$ -sheets prevail are also seen (Manavalan and Johnson, 1983). Fractions of cuticular protein secondary structure were estimated from analysis of the CD spectrum as described in Section 2. The analysis indicates a high percentage (54%) of  $\beta$ -sheet conformation with a small contribution of  $\alpha$ -helix (ca. 13%). The contributions of  $\beta$ -turns/loops and random coil are estimated as 24% and 9%, respectively.

## 4. Discussion

Spectral acquisition by ATR FT-IR and FT-Raman spectroscopy has been shown to yield rich information on the secondary structure of cuticle proteins, without the drawbacks associated with the more conventional vibrational techniques. Our previous attempts to obtain useful laser-Raman spectra from insect cuticles utilizing the conventional form of this spectroscopy were not successful mainly because of the high fluorescent background in the spectra.

Arthropod cuticle is a natural composite of chitin filaments embedded in a protein matrix, built by self-assembly mechanisms, of critical importance for the survival of most arthropods (Neville, 1975). The structural properties of its component proteins seem to dictate the ways they interact with each other and with chitin. Fortunately, due to sustained efforts primarily by Svend Andersen and his group (for reviews see Andersen et al., 1995; Willis, 1999), the amino acid sequences of several cuticle proteins are now known. However, to date there has been no systematic structural analysis of these pro-

teins and no substantial suggestion as to how they interact three-dimensionally with chitin (Iconomidou et al., 1999; Andersen, 2000). The efforts have long been hampered by the complexity of cuticle (Neville, 1975). As a first step towards unraveling the secrets of cuticle's architecture we have undertaken to study the secondary structure of its proteins: protein–chitin interactions (molecular recognition between cuticular proteins and chitin) are definitely based on interactions of cuticular protein secondary structure elements with the chitin filaments. Our results here show conclusively that the main secondary structure element of cuticle proteins is the antiparallel  $\beta$ -pleated sheet. Also, “soft”-cuticle proteins both in the dried (lyophilized) state and in solution are dominated by  $\beta$ -sheet conformation, most probably an antiparallel  $\beta$ -sheet. Our conclusion is supported by recent experimental evidence (Krejchi et al., 1997; Parkhe et al., 1998), where the salient spectroscopic fingerprint of an antiparallel  $\beta$ -pleated sheet is given and discussed. Although lyophilization might increase the  $\beta$ -sheet content of proteins (Griebenow et al., 1999 and references cited therein), the CD experiments verify that for “soft”-cuticle proteins this conformation persists in solution as well. Furthermore, cuticular proteins in the intact larval dorsal abdomen (LDA) cuticle of *H. cecropia* are found to have the antiparallel  $\beta$ -sheet as the predominant molecular conformation. The results agree with our previous proposal (Iconomidou et al., 1999) and secondary structure prediction (Fig. 1 and data not shown) on full amino acid sequences of “soft”-cuticle proteins. These observations are also in agreement with the conclusion of Frankel and Rudall some 50 years ago (Frankel and Rudall, 1947), based on X-ray diffraction data that dry blowfly larval cuticular protein is in the  $\beta$ -form. Hackman and Goldberg (1979) have also shown that the cuticular proteins from larvae of *Caliphora vicina* in the solid state adopt a  $\beta$ -sheet conformation. The latter authors seem to have misinterpreted their ORD data and came to the conclusion that the same cuticular proteins in solution are in a disordered conformation. Their ORD spectrum (Fig. 3 of Hackman and Goldberg, 1979) exhibits profound similarities with our CD spectrum presented in Fig. 7 and probably arises mostly from an antiparallel  $\beta$ -sheet. Therefore, chitin–protein interactions are probably based on *antiparallel*  $\beta$ -sheet–chitin chain interactions. It is very interesting to note that 15 years ago, Atkins (1985) proposed that the antiparallel  $\beta$ -pleated sheet part of cuticular proteins binds to  $\alpha$ -chitin. His proposal was based mainly on a 2D lattice matching between the 010 surface of  $\alpha$ -chitin (with a lattice of 0.47 nm $\times$ 2.07 nm) and the antiparallel  $\beta$ -pleated sheet structure of cuticular proteins.

Our conclusion and the proposal of Atkins (1985) are in good agreement with recent experimental findings at atomic resolution (from the solution of the structures of tachycitin and hevein, an invertebrate chitin-binding pro-

tein and a plant-chitin binding protein, respectively), that chitin-binding proteins in invertebrate and plants comprise a common chitin-binding structural motif, which contains a two-stranded  $\beta$ -sheet (Suetake et al., 2000).

It is also interesting to note that in the case of the bacterial chitinases, three-domain enzymes which bind and hydrolyze chitin, an immunoglobulin-like module (domain), an antiparallel  $\beta$ -sheet barrel is postulated to play an important role in “holding” (interacting with) the chitin chain in place to facilitate catalysis. Four conserved tryptophans on the surface of this  $\beta$ -sheet domain are assumed to interact firmly with chitin, “guiding” the long chitin chains towards the catalytic “groove” (Perrakis et al., 1997).

Therefore, it seems that there are at least three distinct solutions in nature whereby chitin binds to protein and in all these cases,  $\beta$ -sheet and surface aromatic residues appear to be significant (see also Iconomidou et al., 1999). In the third case, the case of insect cuticle, the precise way(s) of interaction remains to be established by more refined theoretical and experimental work.

## Acknowledgements

We thank Dr A. Paipetis for his help with the experiments. Financial support of this work was provided by the University of Athens and by the National Hellenic Research Foundation.

## References

- Andersen, S.O., 2000. Studies in post-ecdysial nymphal cuticle of locust, *Locusta migratoria*, and cockroach, *Blaberus craniifer*. *Insect Biochem. Mol. Biol.* 30, 569–577.
- Andersen, S.O., Hojrup, P., Roepstorff, P., 1995. Insect cuticular proteins. *Insect Biochem. Mol. Biol.* 25, 153–176.
- Atkins, E.D.T., 1985. Conformations in polysaccharides and complex carbohydrates. *Proc. Int. Symp. Biomol. Struct. Interactions, Suppl. J. Biosci.* 8 (1, 2), 375–387.
- Binger, L.C., Willis, J.H., 1994. Identification of the cDNA, gene and promoter for a major protein from flexible cuticles of the giant silkworm *Hyalophora cecropia*. *Insect Biochem. Mol. Biol.* 24, 989–1000.
- Cael, J.J., Koenig, J.L., Blackwell, J., 1974. Infrared and raman spectroscopy of carbohydrates. 3. Raman spectra of the polymorphic forms of amylose. *Carbohydr. Res.* 29, 129–134.
- Cai, S., Singh, B.R., 1999. Identification of  $\beta$ -turn and random coil amide III infrared bands for secondary structure estimation of proteins. *Biophys. Chem.* 80, 7–20.
- Carey, P.R., 1982. *Biochemical Applications of Raman and Resonance Raman Spectroscopies*. Academic Press, New York.
- Carey, P.R., Salares, V.R., 1980. Raman and resonance Raman studies of biological systems. In: Clark, R.J.H., Hester, R.E. (Eds.), *Advances in Infrared and Raman Spectroscopy*, vol. 7. Heyden, London, pp. 1–58.
- Cox, D.L., Willis, J.H., 1985. The cuticular proteins of *Hyalophora cecropia* from different anatomical regions and metamorphic stages. *Insect Biochem. Mol. Biol.* 15, 349–362.

- de Jongh, H.H.J., Goormaghtigh, E., Ruyschaert, J.M., 1996. The different molar absorptivities of the secondary structure types in the Amide I region: An attenuated total reflection infrared study on globular proteins. *Anal. Chem.* 242, 95–103.
- Fink, D.J., Gendreau, R.M., 1984. Quantitative surface studies of protein adsorption by infrared spectroscopy. I. Correction for bulk concentrations. *Anal. Biochem.* 139, 140–148.
- Frankel, G., Rudall, K.M., 1947. The structure of insect cuticles. *Proc. R. Soc. London, Ser. B.* 34, 111–143.
- Frushour, B.J., Koenig, J.L., 1975. Raman spectroscopy of proteins. In: Clark, R.J.H., Hester, R.E. (Eds.), *Advances in Infrared and Raman Spectroscopy*, vol. 1. Heyden, London, pp. 35–97.
- Fu, F.-N., De Oliveira, D.B., Trumble, W.R., Sarkar, H.K., Singh, B.R., 1994. Secondary structure estimation of proteins using the Amide III region of Fourier transform infrared spectroscopy: Application to analyze calcium-binding-induced structural changes in calsequestrin. *Appl. Spectrosc.* 48 (11), 1432–1441.
- Griebenow, K., Santos, A.M., Carrasquillo, K.G., 1999. Secondary structure of proteins in the amorphous dehydrated state probed by FTIR spectroscopy. Dehydration-induced structural changes and their prevention. *Int. J. Vib. Spectrosc.* 3 (1), 1–34.
- Hackman, R.H., Goldberg, M., 1979. Some conformational studies of larval cuticular proteins from *Caliphora vicina*. *Insect Biochem.* 9, 557–561.
- Hamodrakas, S.J., 1988. A protein secondary structure prediction scheme for the IBM PC and compatibles. *CABIOS* 4, 473–477.
- Hamodrakas, S.J., Asher, S.A., Mazur, G.D., Regier, J.C., Kafatos, F.C., 1982. Laser-Raman studies of protein conformation in the silkworm chorion. *Biochim. Biophys. Acta* 703, 216–222.
- Haris, P.I., Chapman, D., 1995. The conformational analysis of peptides using Fourier transform IR spectroscopy. *Biopolymers (Peptide Sci.)* 37, 251–263.
- Hillerton, J.E., Vincent, J.F.V., 1979. The stabilization of insect cuticles. *J. Insect Physiol.* 25, 957–963.
- Iconomidou, V.A., Willis, J., Hamodrakas, S.J., 1999. Is  $\beta$ -pleated sheet the molecular conformation which dictates the formation of the helicoidal cuticle? *Insect Biochem. Mol. Biol.* 29, 285–292.
- Jackson, M., Mantsch, H.H., 1995. The use and misuse of FTIR spectroscopy in the determination of protein structure. *Crit. Rev. Biochem. Mol. Biol.* 30 (2), 95–120.
- Jakobsen, R.J., Wasacz, F.M., 1990. Infrared spectra–structure correlations and absorption behavior for helix proteins. *Appl. Spectrosc.* 44 (9), 1478–1490.
- Kim, H.S., Hartgerink, J.D., Ghadiri, M.R., 1998. Oriented self-assembly of cyclic peptide nanotubes in lipid membranes. *J. Am. Chem. Soc.* 120, 4417–4424.
- Krejchi, M., Cooper, S.J., Deguchi, Y., Atkins, E.D.T., Fournier, M.J., Mason, T.L. et al., 1997. Crystal structures of chain-folded antiparallel  $\beta$ -sheet assemblies from sequence-designed periodic polypeptides. *Macromolecules* 30, 5012–5024.
- Krimm, S., Bandekar, J., 1986. Vibrational spectroscopy and conformation of peptides, polypeptides, and proteins. *Adv. Protein Chem.* 38, 181–386.
- Lampe, D.J., Willis, J.H., 1994. Characterization of a cDNA and gene encoding a cuticular protein from rigid cuticles of the giant silkworm, *Hyalophora cecropia*. *Insect Biochem. Mol. Biol.* 24, 419–435.
- Levitt, M., Greer, J., 1977. Automatic identification of secondary structure in globular proteins. *J. Mol. Biol.* 114, 181–293.
- Manavalan, P., Johnson, W.C. Jr., 1983. Sensitivity of circular dichroism to protein tertiary structure class. *Nature* 305, 831–832.
- Minke, R., Blackwell, J., 1978. The structure of alpha-chitin. *J. Mol. Biol.* 120, 167–181.
- Mikkelsen, A., Engelsen, S.B., Hansen, H.C.B., Larsen, O., Skibsted, L.H., 1997. Calcium carbonate crystallization in the  $\alpha$ -chitin matrix of the shell of pink shrimp, *Pandalus borealis*, during frozen storage. *J. Cryst. Growth* 177, 125–134.
- Neville, A.C., 1975. *Biology of the Arthropod Cuticle*. Springer-Verlag, New York.
- Parkhe, A.D., Cooper, S.J., Atkins, E.D.T., Fournier, M.J., Mason, T.L., Tirrell, D.A., 1998. Effect of local sequence inversions on the crystalline antiparallel  $\beta$ -sheet lamellar structures of periodic polypeptides: implications for chain-folding. *Int. J. Biol. Macromol.* 23, 251–258.
- Perrakis, A., Ouzounis, C., Wilson, K.S., 1997. Evolution of immunoglobulin-like modules in chitinases: their structural flexibility and functional implications. *Folding Design* 2, 291–294.
- Rebers, J.F., Riddiford, L.M., 1988. Structure and expression of a *Manuca sexta* larval cuticle gene homologous to *Drosophila* cuticle genes. *J. Mol. Biol.* 203, 411–423.
- Rost, B., 1996. PHD: Predicting one-dimensional protein structure by profile based neural networks. *Methods Enzymol.* 266, 525–539.
- Rost, B., Sander, C., 1993. Prediction of protein secondary structure at better than 70% accuracy. *J. Mol. Biol.* 232, 584–599.
- Rost, B., Sander, C., 1994. Combining evolutionary information and neural networks to predict protein secondary structure. *Proteins* 19, 55–55.
- Rubin, G.M., Yandell, M.D., Wortman, J.R., Miklos, G.L.G., Nelson, C.R., Hariharan, I.K. et al., 2000. Comparative genomics of the eukaryotes. *Science* 287, 2204–2215.
- Savitsky, A., Golay, M.J.E., 1964. Smoothing and differentiation of data by simplified least-squares procedures. *Anal. Chem.* 36, 1627–1639.
- Singh, B.R., Fuller, M.P., 1991. FT-IR in combination with the attenuated total reflectance technique: A very sensitive method for the structural analysis of polypeptides. *Appl. Spectrosc.* 45 (6), 1017–1021.
- Singh, B.R., Fuller, M.P., Schiavo, G., 1990. Molecular structure of tetanus neurotoxin as revealed by Fourier transform infrared and circular dichroic spectroscopy. *Biophys. Chem.* 46, 155–166.
- Spiro, T.G., Gaber, B.P., 1977. Laser-Raman scattering as a probe of protein structure. *Annu. Rev. Biochem.* 46, 553–572.
- Stephens, R.M., Vincent, J.F.V., 1976. Infrared spectroscopy of locust extensible intersegmental cuticle and its components. *J. Insect Physiol.* 22, 601–605.
- Suetake, T., Tsuda, S., Kawabata, S., Miura, K., Iwanaga, S., Hikichi, K. et al., 2000. Chitin-binding proteins in invertebrates and plants comprise a common chitin-binding structural motif. *J. Biol. Chem.* 275 (24), 17929–17932.
- Surewicz, W.K., Mantsch, H.H., Chapman, D., 1993. Determination of protein secondary structure by Fourier transform infrared spectroscopy: A critical assessment. *Biochemistry* 32 (2), 389–394.
- Yu, N.T., 1977. Raman spectroscopy: A conformational probe in biochemistry. *CRC Crit. Rev. Biochem.* 4, 229–280.
- Willis, J.H., 1999. Cuticular proteins in insects and crustaceans. *Am. Zool.* 39, 600–609.