Available online at www.sciencedirect.com
SCIENCE DIRECT
Insect Biochemistry and Molecular Biology 35 (2005) 553–560
Insect Biochemistry and Molecular Biology
www.elsevier.com/locate/ibmb

Unique features of the structural model of 'hard' cuticle proteins: implications for chitin–protein interactions and cross-linking in cuticle

Vassiliki A. Iconomidou[a], Judith H. Willis[b], Stavros J. Hamodrakas[a,*]

[a]Department of Cell Biology and Biophysics, Faculty of Biology, University of Athens, Panepistimiopolis, Athens 157 01, Greece
[b]Department of Cellular Biology, University of Georgia, Athens, GA 30602, USA

Received 7 December 2004; received in revised form 26 January 2005; accepted 27 January 2005

Abstract

Cuticular proteins are one of the determinants of the physical properties of cuticle. A common consensus region (extended R&R Consensus) in these proteins binds to chitin, the other major component of cuticle. We previously predicted the preponderance of β-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). Subsequently, we verified experimentally the abundance of antiparallel β-pleated sheet in the structure of cuticle proteins (Iconomidou et al., Insect Biochem. Mol. Biol. 31 (2001) 877). Homology modelling of soft (RR-1) cuticular proteins using bovine plasma retinol binding protein (RBP) as a template revealed an antiparallel β-sheet half-barrel structure as the basic folding motif (Hamodrakas et al., Insect Biochem. Molec. Biol. 32 (2002) 1577). The RR-2 proteins characteristic of hard cuticle, have a far more conserved consensus and frequently more histidine residues. Extension of modelling to this class of consensus, in this work, reveals in detail several unique features of the proposed structural model to serve as a chitin binding structural motif, thus providing the basis for elucidating cuticle's overall architecture and chitin–protein interactions in cuticle.
© 2005 Elsevier Ltd. All rights reserved.

Keywords: Hard cuticular proteins; Structural model; Sclerotization; Chitin-binding

1. Introduction

Cuticle is a composite material made primarily of chitin filaments embedded in a proteinaceous matrix. It provides structural and mechanical support by serving functionally as both skin and skeleton to arthropods (Neville, 1975; Vincent and Wegst, 2004). The mechanical properties of the cuticle are conferred by the proportion of chitin, by the degree of sclerotization and by the sequences of its proteins.

The precise nature of the interaction of cuticular proteins with chitin fibers and the detailed structure of insect cuticle have not yet been resolved. Certain sequence motifs occur in cuticular proteins from even distantly related species and such conserved motifs have common and important roles for the proper function of cuticle (Andersen et al., 1995). The most prevalent motif is the "R&R Consensus sequence" first identified by Rebers and Riddiford (1988): G-x(8)-G-x(6)-Y-x-A-x-E-x-G-Y-x(7)-P-x(2)-P or a modification of it: G-x(7)-[DEN]-G-x(6)-[FY]-x-A-[DGN]-x(2,3)-G-[FY]-x-[AP]-x(6) (Willis, 1999) (where x represents any amino acid, the values in parentheses indicate the number of residues and brackets include alternative amino acids at the site). An extension of this motif is a stretch of approximately 68 amino acids, the "extended R&R Consensus" that

*Corresponding author. Tel: +30 210 727 4931; fax: +30 210 727 4254.
E-mail addresses: vecomom@cc.uoa.gr (V.A. Iconomidou), shamodr@cc.uoa.gr (S.J. Hamodrakas).

0965-1748/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ibmb.2005.01.017

was recognized by several groups (reviewed in Willis et al., 2005). Andersen recognized and named three distinct forms of the "extended consensus" RR-1, RR-2 (Andersen, 1998) and RR-3 (Andersen, 2000). RR-1 bearing proteins have been isolated from flexible cuticles, while RR-2 proteins have been associated with hard cuticle. The RR-3 form of the consensus has been based on but five sequences from postecdysial cuticle of insects plus sequences from other arthropod classes. The assignment of RR-1 proteins to "soft" cuticle and RR-2 to "hard" is based on limited data (discussed in Willis et al., 2005). There is no agreed upon definiton of this classification of cuticle. "Hard" cuticles are generally sclerotized and mechanically stiff. "Soft" cuticles include, but are not restricted to, those that can expand within an instar due to growth by intussusception. Further work is needed to learn if the assignment of protein class to cuticle type is universal.

The prevalence of the R&R Consensus led several authors to postulate that it served an important function, quite possibly chitin binding (Bouhin et al., 1992; Charles et al., 1992; Andersen et al., 1995).

The involvement of the extended consensus in chitin binding has been confirmed by direct experimentation (Rebers and Willis, 2001; Togawa et al. 2004). There were earlier experimental findings and proposals that β-sheet should be involved in chitin–protein interactions (Fraenkel and Rudall, 1947; Atkins, 1985) and these have been amplified by secondary structure prediction and experimental data. This more recent work indicates that antiparallel β-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 most probably involved in β-sheet-chitin chain interactions of the cuticular proteins with chitin filaments (Iconomidou et al., 1999, 2001).

A more specific analysis of the nature of cuticular protein/chitin binding became possible when, unexpectedly, a distant (20%) sequence similarity was found between "soft" cuticle proteins and the crystallographically determined C-terminal (Zanotti et al., 1994), β-barrel portion, of bovine plasma retinol binding protein (RBP). When, following alignment, both conservative substitutions and identities were combined, the similarity rises to 60% of the total HCCP12 sequence (Hamodrakas et al., 2002), a representative member of the "soft" cuticle (RR-1) proteins (Binger and Willis, 1994; Iconomidou et al., 1999). This similarity allowed the construction, by "homology" (comparative) modelling of a structural model of the "extended R&R Consensus" of cuticle proteins (Hamodrakas et al., 2002). This modelling was successful even though it seems that RBP and the R&R Consensus-bearing cuticular proteins are not strictly homologous.

Furthermore, modelling of HCCP66 (Entrez accession number 1169133) and AGCP2b (Entrez accession number 2961110), two "hard" cuticle proteins, showed that the "extended R&R Consensus" (Iconomidou et al., 1999) not only of "soft" but also of "hard" cuticle proteins might easily adopt the proposed conformation (Hamodrakas et al., 2002).

The RR-2 bearing proteins, associated with hard cuticles, are of particular interest because the extended R&R Consensus is virtually invariant in length and in the identity of one third of its amino acids and very limited variation in another third. Furthermore, it is RR-2 proteins, far more than RR-1, which may have numerous histidine residues that could participate in cross-linking (Willis et al., 2005). In this work, we present in detail several unique features of the proposed structural model for 'hard' cuticle proteins to serve as a chitin binding structural motif, thus providing the basis for elucidating cuticle's overall architecture and chitin–protein interactions in cuticle.

2. Materials and methods

A sensitive alignment of a representative set of 44 'hard' cuticle protein sequences (Table 1) was produced with CLUSTAL W (Thompson et al., 1994). The BLOSUM 62 similarity matrix was used and all other parameters were the default parameters of CLUSTAL W (Thompson et al., 1994).

A structural model for 'hard' cuticle proteins was then derived by homology modelling, utilizing the program WHAT IF (Vriend, 1990), using as template the structural model proposed for 'soft' cuticle proteins (Hamodrakas et al., 2002). The model was regularized with the WHAT IF regularization options (Vriend, 1990) and optimized employing the GROMOS molecular dynamics software (Van Gunsteren and Berendsen, 1987).

The ribbon representation of the 'hard' cuticle protein model was displayed using GRASP (Nicholls et al., 1991). Docking experiments were performed utilizing the program GRAMM (Vakser, 1996).

3. Results

The extended RR-2 consensus region of 44 proteins with 36 different sequences were aligned (Fig. 1). What is extraordinary about the RR-2 consensus is its conservation across 14 species from six orders of insects. Only two single amino acid gaps are required to accommodate all 44 RR-2 sequences. Twenty-two of the 70 residues (31%) in the extended consensus are virtually invariant and an additional 23 are represented

Table 1
'Hard' cuticle proteins bearing the RR-2 motif

Protein	Entrez/SwissProt Accession Number	
Coleoptera		
	TM-LCP-A1A	1706191
	TM-LCP-A2B	1706192
	TM-LCP-A3A	1706194
	TMACP20	102879
	TMACP22	113012
Dictyoptera:		
	BC-NCP8	P82121
Diptera		
	AnGCP2a	2961109
	AnGCP2b,c,d	2961110, 2961111, 2961113
	DMCcp84Aa	4389433
	DMCcp84Ab	4389434
	DMCcp84Ac	4389435
	DMCcp84Ad	4389436
	DMCcp84Ae	4389437
	DMCcp84Af	4389438
	DMCcp84Ag	4389439
	DMEDG84	117640
	DMCry	22946279
	DS, DY	9966434, 9966436
Lepidoptera		
	BMEDG84A	3608259
	BMWCP1A	12862579
	BMWCP1B	12862581
	BMWCP2	12862583
	BMWCP3	12862585
	BMWCP4	12862587
	BMWCP5	12862589
	BMWCP6	12862591
	BMWCP7A,B	12862593, 12862595
	BMWCP8	12862597
	HCCP66	1169133
Orthoptera		
	LM-ACP7	998751
	LM-ACP8	84730
	LM-ACP19	1345864
	LM-ACP21	3287770
	LM-NCP19.8	P82166
Hemiptera		
	AGCP	29124934
	5 Aphid species	29124930, 29124932, 29124936, 29124938, 16798648

Protein sequences and additional annotation can be found at: http://www3.ncbi.nlm.nih.gov/Entrez/index.html. Sequences that have an identifier that begins with a letter can be found at: http://us.expasy.org/. They are also available at cuticleDB (http://bioinformatics.biol.uoa.gr/cuticleDB).

by a single amino acid in over half of the proteins (displayed by red and green color respectively in Fig. 1).

Based on the abundance of certain amino acids in each position, which correspond to 'any amino acid' (–x–), in the RR-2 consensus defined in Willis et al. (2005), the R&R Consensus was 'filled' and a 'hard' cuticle 'representative' protein sequence, the RR-2-Rep, was constructed (Fig. 1).

A structural model for RR-2-Rep (and consequently for several cuticular proteins of "hard" cuticles) was constructed, using as template the structural model proposed for 'soft' cuticle proteins (Hamodrakas et al., 2002) and utilizing the popular homology modelling software WHAT IF (Vriend, 1990). The model is presented in Fig. 2. It comprises 67 residues out of 68 residues of the "extended R&R Consensus", that is the RR-2 Consensus (Andersen, 2000), the evolutionarily conserved region of "hard" cuticle proteins.

Plots of the proposed model for "hard" cuticle proteins complexed with a *N*-acetyl glucosamine (NAG) tetramer in an extended conformation, are shown in Figs. 3A–C, respectively. They are the most favourable complexes, which were derived from a "high-resolution" docking experiment of a NAG tetramer to the 'hard' cuticle model, utilizing the docking program GRAMM (Vakser, 1996). The one in 3A has the NAG tetramer more or less parallel to the first β-strand of the RR-2-Rep half-β-barrel model, whereas that in 3B has the NAG tetramer more or less parallel to the last β-strand of the RR-2-Rep half-β-barrel model. The docking experiment in 3C clearly shows that the proposed model for cuticle proteins may, alternatively, accommodate, rather comfortably, at least one extended chitin chain, almost perpendicular to the β-strands.

4. Discussion

There are now 139 sequences available for what are known or postulated to be cuticle structural proteins (Willis et al., 2005). These numbers do not include almost 200 more that have been identified by protein prediction programs used to annotate the *D. melanogaster* and *A. gambiae* genomes. These have been omitted because their annotation is still in a state of flux.

The R&R Consensus is a common feature of cuticle structural proteins, from all six orders of insects examined to date (Fig. 1) and it has also been recognized in cuticular proteins from arachnids and crustaceans (reviewed in Willis, 1999). Although the evidence that the consensus binds to chitin is compelling, the precise nature of the interaction of insect cuticle proteins and chitin is still unknown. In a previous work, we proposed by homology modelling an antiparallel β-sheet half-barrel structure (Hamodrakas et al., 2002) as the basic folding motif of the "extended R&R Consensus" of cuticle proteins, using as template the HCCP12, a "soft" cuticle representative insect cuticular protein (bearing

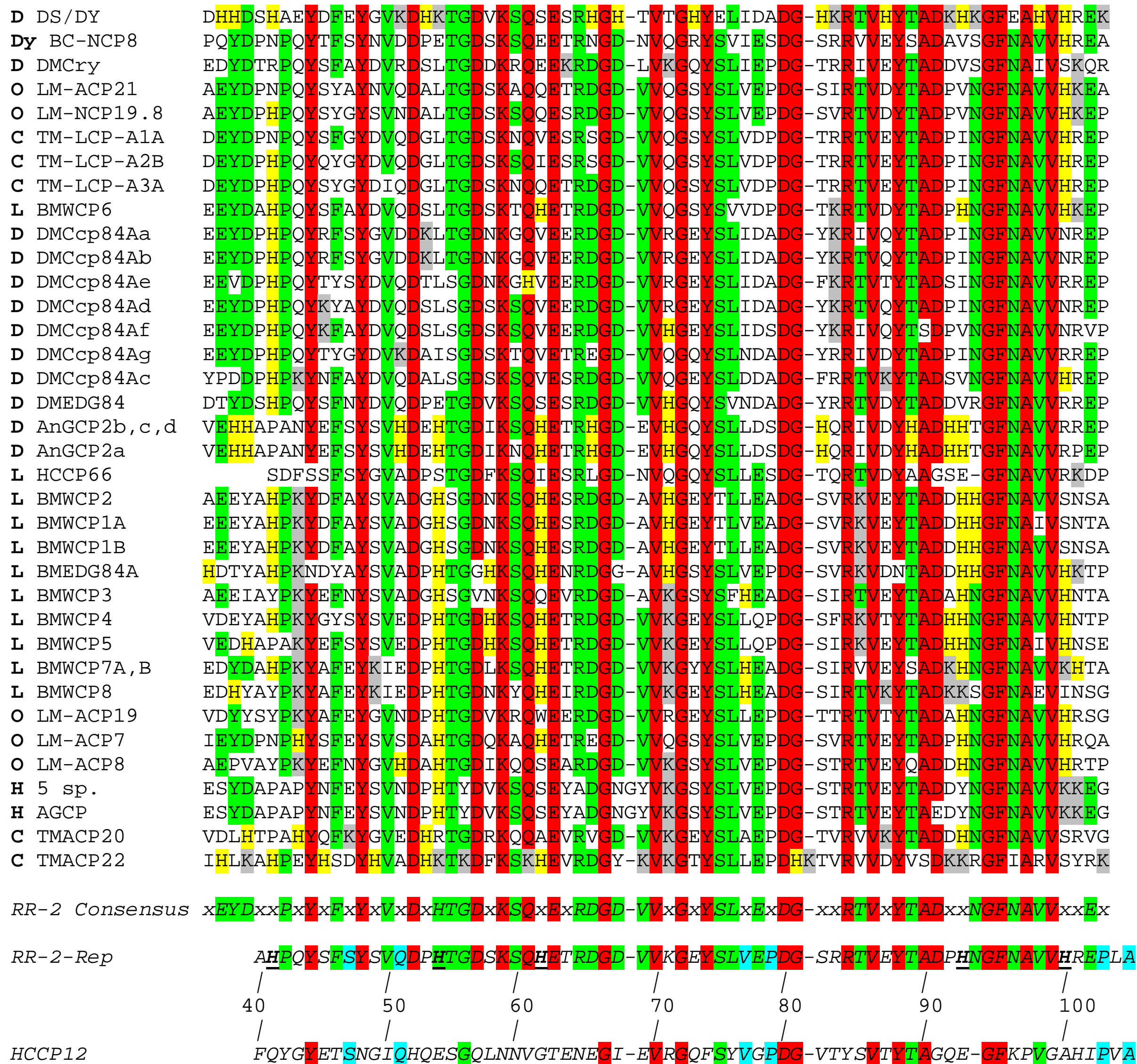

Fig. 1. Alignment of the pfam00379 region of 44 cuticular proteins with the RR-2 consensus. The pfam 00379 regions for RR-2 proteins were aligned with ClustalW [http://clustalw.genome.ad.jp/]. Orders of insects are: Coleoptera (**C**), Diptera (**D**), Lepidoptera (**L**), Orthoptera (**O**), Dictyoptera (**Dy**) and Hemiptera (**H**). All hemipteran proteins except the one from *A. gossypii* are indicated by [***H*** 5 sp]. Abbreviations for proteins as in Table 1. Red represents amino acids present in at least 95% of the proteins, green in the majority. Histidines are shown in yellow, lysines in light gray. The extended RR-2 consensus is shown under the protein set, together with a 'representative' hard-cuticle protein sequence, the RR-2-Rep. In RR-2-Rep, every -x-, "any amino acid", of the RR-2 consensus has been replaced with the most abundant amino acid of the same position in the sequences of the protein set. Important histidines, shown in Fig. 2 are bold and underlined. Alignment of the RR-2-Rep sequence with that of the "soft"-cuticle representative insect cuticular protein HCCP12 [ENTREZ accession number 1169129] is also shown, for comparison of the models constructed by homology modelling, both for the "hard" and "soft" cuticle proteins (see Fig. 2 and also Hamodrakas et al., 2002). The light blue color represents the extra conservative amino acids between the RR-2-Rep and the HCCP12 sequences. The numbering at the bottom is that of unprocessed HCCP12. This figure is modified from Fig. 2 in Willis et al. (2005) with permission.

the RR-1 motif), which exhibits significant sequence similarity to the C-terminal sequence part of retinol binding protein (PDB code: 1FEN; Zanotti et al., 1994).

In this paper, the tertiary structure of the extended consensus of RR-2 proteins has been modelled. The same proposed half-barrel model shown in Fig. 2, has several attractive features to act as a structural entity interacting with the chitin chains in hard cuticle as well.

It is an antiparallel β-sheet structure, in agreement with our earlier proposals (Iconomidou et al., 1999) and those of Atkins (1985), Hackman and Goldberg (1979) and Fraenkel and Rudall (1947) and with our recent experimental data (Iconomidou et al., 2001).

The proposed structure has a "cleft" containing conserved aromatic residues (mostly tyrosines and phenylalanines), which are seen to form "flat" hydrophobic

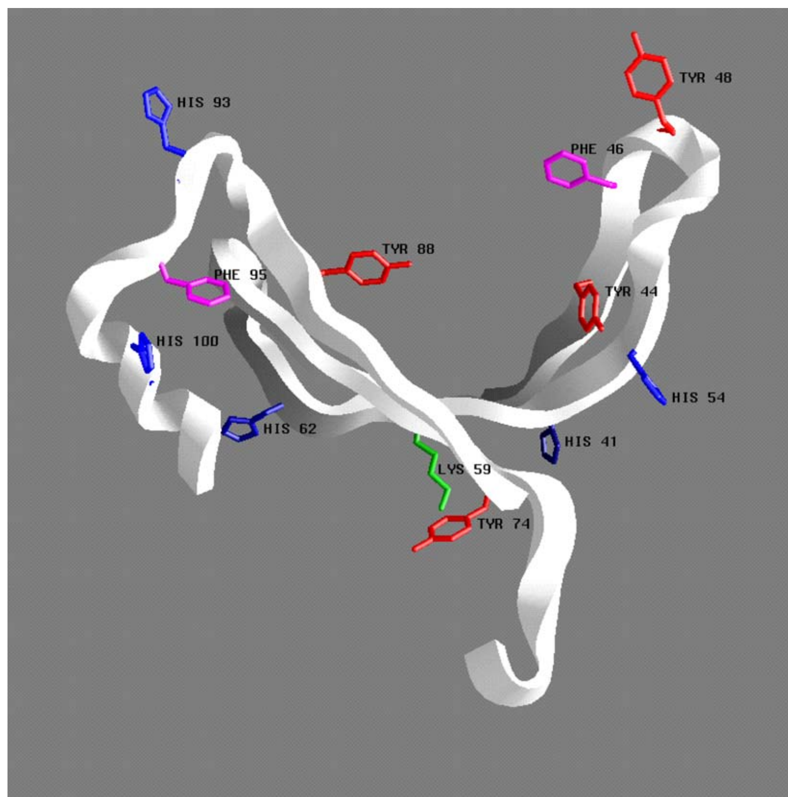

Fig. 2. A ribbon model of hard-cuticle protein structure, displayed using GRASP (Nicholls et al., 1991). The structure of the 'representative' hard-cuticle protein, RR-2-Rep (see Fig. 1), was modelled on that of HCCP12 (Hamodrakas et al., 2002), utilizing the program WHAT IF (Vriend, 1990) and the alignment details shown in Fig. 1. The side chains of several aromatic residues are shown as "ball and sticks" and numbered, following the numbering scheme of the unprocessed HCCP12 sequence, which begins at residue 17 as VPL. These are: H41, Y44, F46, Y48, H54, K59, H62, Y74, Y88, H93, F95 and H100. The model structure has a "cleft" full of aromatic residues, which form a "flat" surface of aromatic rings (upper side), ideally suited for cuticle protein–chitin chain interactions, and an outer surface (lower side) which should be important for protein–protein interactions in cuticle. Also, the model provides for the right positioning of histidine residues, so that these histidines might participate in cuticle sclerotization. It is seen that histidines (the bold and underlined histidines of Fig. 1) occupy "exposed" positions either in turns or at the "edges" of the half-β-barrel or its periphery, in excellent positions to be involved in cuticular sclerotization, readily reacting with quinones or quinone methides derived from *N*-acyldopamine residues (Kramer et al., 2001), or being involved in the variations of the water binding capacity of cuticle and the interactions of its constituent proteins, due to the fact that small changes of pH can affect the ionization of their imidazole group.

surfaces on one "face" of the model structure (Fig. 2). These are the side chains of Y44, F46 and Y48 on the right-hand side of the "cleft" and Y88 on the other. The aromatic rings of these residues could well stack against faces of the saccharide rings of chitin (poly *N*-acetyl glucosamine) chains. This type of interaction is fairly common in protein-saccharide complexes (Vyas, 1991; Hamodrakas et al., 1997; Tews et al., 1997). It is interesting to note that we had foreseen such interactions from secondary structure prediction alone, 5 years ago (Iconomidou et al., 1999). It is also of interest, that a triad of aromatic residues, comparable to those at 44, 46 and 48, is a common feature of RR-1 proteins even though their precise location in the extended consensus is somewhat variable (see, for example, residues 40, 42, 44 in HCCP12 in Fig. 1).

There is also an invariant lysine in all RR-2 sequences (K59, see Figs. 1 and 2) and several other positions seem to be favorable for either lysine or histidine residues. Over half of the RR-2 proteins have histidine as their final and/or penultimate C-terminal amino acid. Histidines are less abundant in the extended consensus of RR-1 proteins (Willis et al., 2005). In the 50 mature RR-1 proteins used in that analysis, the average number of histidines per protein was 3.1 (range of 0–13) and 3.7 lysines (range 0–21). By contrast, these 44 RR-2 proteins averaged 16.8 histidines (range of 0–58) and 6.8 lysines (range 3–12). The RR-2 histidine abundance is far more than can be explained by a 61% greater average length of the RR-2 proteins. Histidines and lysines are known to be reactive sites for sclerotizing agents (Kerwin et al., 1999; Kramer et al., 2001; Andersen, 2005), so it is provocative that so many proteins from "hard" sclerotized cuticles would have these amino acids in abundance (Willis et al., 2005)

All the bold and underlined histidines in RR-2-Rep (Fig. 1) occupy "exposed" positions either in turns (like H93), or at the "edges" of the half-β-barrel or its

periphery (like H41, H54, H62, H100), in excellent positions to be involved in cuticular sclerotization, readily reacting with catechol quinones and quinone methides (Kramer et al., 2001). Alternatively, histidines could be involved in the variations of the water binding capacity of cuticle and the interactions of its constituent proteins, because small changes of pH, within physiologically relevant ranges, can affect the ionization of their imidazole groups (Andersen et al., 1995).

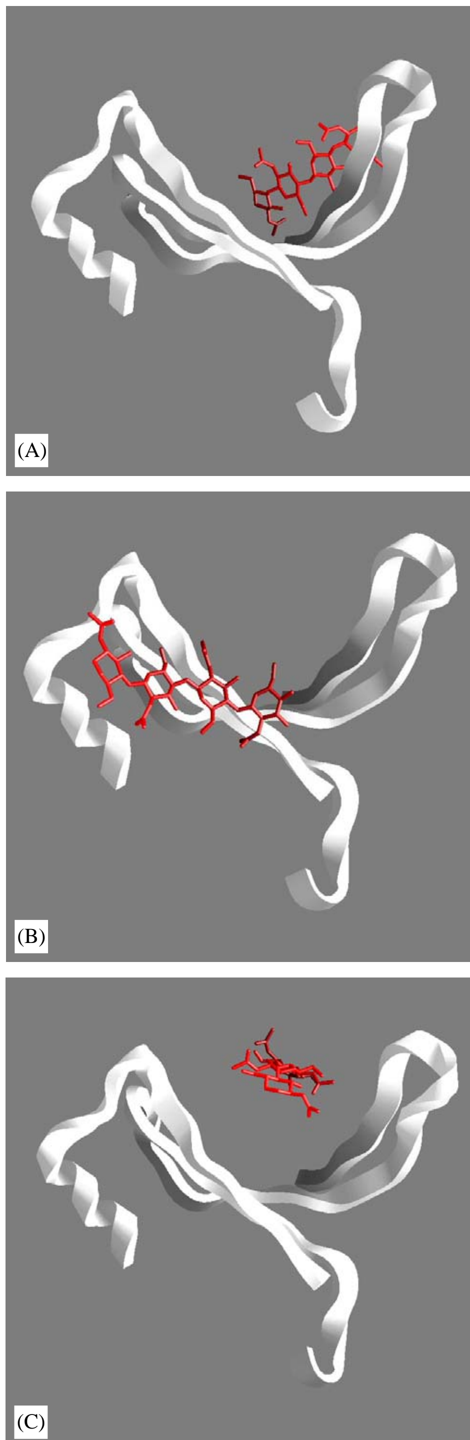

Fig. 3. Ribbon models of cuticular proteins derived from homology modelling and docking experiments with a *N*-acetyl glucosamine (NAG) tetramer in an extended conformation. (A)–(C) Possible complexes of RR-2-Rep with a NAG tetramer in an extended conformation derived from a "high resolution" docking experiment, utilizing the program GRAMM (Vakser, 1996) and the default parameters of the program for high resolution. The two models presented in (A) and (B) are the two "top on the list", most favorable complexes, whereas third on the list is a structure similar to that of (C). The one in (A) has the NAG tetramer more or less parallel to the first β-strand of the RR-2-Rep half-β-barrel model, whereas that in (B) has the NAG tetramer more or less parallel to the last β-strand of the RR-2-Rep half-β-barrel model. Note that, both in (A) and (B) the chitin chain runs parallel to the β-strands, whereas in (C) the chain is arranged perpendicular to the β-strands.

These observations are in excellent agreement with the predictions made several years ago for the role of histidines from secondary structure predictions (Iconomidou et al., 1999) and strengthen further the value of the models previously proposed both for "soft" and "hard" cuticle proteins (Hamodrakas et al., 2002). In fact, the favorable locations of all the side chains of the abundant histidines in "hard" cuticle proteins, in such excellent orientations to promote cross-linking, are the most eminent indication that the model should be correct. The presence and location of histidines is the main difference of the model for "hard" cuticle proteins from the model for "soft" cuticle proteins.

Furthermore, the invariant lysine (K59), in all 44 protein sequences from hard cuticular proteins of six different orders, deserves special attention. Its side chain protrudes away, almost perpendicular to the point with the highest curvature of the outer surface of the half-β-sheet barrel (Fig. 2) and is well positioned to serve in cross-linking (perhaps with the intervention of catechol quinones and quinone methides; Kramer et al., 2001) of successive half-barrels.

In addition, in several places along the sequences of the cuticular proteins from hard cuticle, two adjacent residues are either histidines (HH) or a histidine (H) together with a positively charged residue, a lysine (K) or an arginine (R). The most characteristic examples are positions 92 and 93 or 93 and 94 (Fig. 1). Ten of the sequences have two adjacent histidines, HH, and an additional six have KH or KK in that region. A close examination of the proposed model shows that these positions are found at a turn of the model (top left hand corner of Fig. 2), in ideal positions to serve for cross-linking or being exposed to the solvent. In all other appearances of two successive histidine occurrences (or two successive positively charged residues) these, again, occupy favorable exposed positions (data not shown), which strongly supports the proposed model.

Docking experiments of an extended *N*-acetylglucosamine tetramer to the model of RR-2-Rep, utilizing the

docking program GRAMM (Vakser, 1996) suggest certain important modes for chitin–protein possible interactions (Fig. 3): First, a high-resolution experiment, utilizing the default parameters of the program, indicates that chitin protein chains may run parallel to the β-strands of the half-β-barrel (Figs. 3A and B). Thus, β-barrels of cuticle proteins may intervene between the long chitin chains, in cuticle without disrupting continuity. This parallel arrangement of cuticle protein β-strands with the chitin chains agrees with observations made by Atkins some 20 years ago (Atkins, 1985) from X-ray diffraction patterns. Alternatively, it is seen that the proposed model for 'hard' cuticle proteins may accommodate, rather comfortably, at least one extended chitin chain (Fig. 3C) perpendicularly to the half-barrel β-strands.

It is interesting to note that, recently, docking software was also used to model insect protein–chitin interactions, although in this case it was mainly focused on the role of a putative active site tryptophan residue in catalysis, caused by the insect chitinase catalytic domain (Huang et al., 2000).

It should be mentioned that, recent systematic and thorough work has resulted in the formation of an extensive and complete database of carbohydrate-binding modules (CBMs) (http://afmb.cnrs-mrs.fr/CAZY), which contains all modes of protein–carbohydrate interactions deduced up to date from X-ray crystallography and NMR studies. This work is summarized in recent reviews (Boraston et al., 2004 and references therein). One observation that can be made examining this database is that, in most cases, a twisted β-sheet is a domain involved in protein–carbohydrate interaction, usually via a number of aromatic residues that interact with the polysaccharide rings (see also Katouno et al., 2004; Watanabe et al., 2003; Suetake et al., 2000). Another rather intriguing observation is that in most cases in this database, the polysaccharide chains usually run perpendicular to the β-strands of the β-sheets. In our proposed model (Fig. 2) both observations are satisfied. However, initial attempts made do not classify our model with certainty to any of the structural motifs contained in the CBMs database.

At this point, is should be emphasized that despite the high conservation of the extended RR-2 consensus, the position of the consensus and the sequences on the amino and carboxyl sides of the consensus are not conserved (data not shown). This set of proteins (Fig. 1) had from 112 to 457 amino acids with an average length of 195. The consensus occurred as proximal as the N-terminus of the mature protein or not until the final third of the sequence. It occupied from 15% to 56% of the sequence, with an average of 38%. The surrounding regions may have multiple AAP(A/V) repeats, may be enriched in glycine residues, may have other common motifs, and may together have from 0% to 36% histidines.

Therefore, it should be stressed that, the model shown in Fig. 2, although it represents the basic building motif of cuticle proteins with the RR-2 (or the RR-1) consensus, it is accompanied by other structural domains formed by the remainder of the proteins. The exact structure of these domains, their relation with the model shown in Fig. 2 and their possible interactions with chitin or other proteins, remain to be elucidated by future work. Also, more elaborate modelling and experimental future work is needed both to correlate the models of chitin–protein interaction proposed here with the model proposed by Blackwell and Weih (1980) for chitin fiber (2.8 nm chitin crystallites composed of several chitin chains)–protein interactions and also to reveal cuticle's architecture in detail.

Twenty years ago it was proposed that the helicoidal architecture of silkmoth chorion should be based on simple sterochemical rules of packing of twisted β-pleated sheets (Hamodrakas, 1984). Whether the helicoidal architecture of cuticle, which is analogous to chorion in many respects, will follow similar rules also awaits further refined modelling and experimental work. However, looking at the model presented in Fig. 2, it is clear that even in the case of cuticle, the rules of packing of twisted β-pleated sheets should be dominant in dictating the formation of its helicoidal structure.

Acknowledgements

This work was supported in part by grant AI055624 from the US National Institutes of Health. We thank the University of Athens for financial support. We also thank Drs. G. Vriend, I. Vakser and A. Nicholls for providing us with their computer programs, and the anonymous referees for their constructive criticism.

References

Andersen, S.O., 1998. Amino acid sequence studies on endocuticular proteins from the desert locust, *Schistocerca gregaria*. Insect Biochem. Molec. Biol. 28, 421–434.

Andersen, S.O., 2000. Studies on proteins in post-ecdysial nymphal cuticle of locust, *Locusta migratoria*, and cockroach, *Blaberus cranifer*. Insect Biochem. Molec. Biol. 30, 569–577.

Andersen, S.O., 2005. Cuticular sclerotization and tanning. In: Gilbert, L., Iatrou, K., Gill, S. (Eds.), Comprehensive Molecular Insect Science, vol. 4. Elsevier, Oxford, pp. 145–170 (Chapter 4).

Andersen, S.O., Højrup, 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. Proceedings of the International Symposium on Biomolecular Structure 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

silkmoth *Hyalophora cecropia*. Insect Biochem. Mol. Biol. 24, 989–1000.

Blackwell, J., Weih, M.A., 1980. Structure of chitin–protein complexes: ovipositor of the ichneumon fly Megarhyssa. J. Mol. Biol. 137, 49–60.

Boraston, A.B., Bolam, D.N., Gilbert, H.J., Davies, G.J., 2004. Carbohydrate-binding modules: fine-tuning polysaccharide recognition. Biochem. J. 382, 769–781.

Bouhin, H., Charles, J.-P., Quennedey, B., Courrent, A., Delachambre, J., 1992. Characterization of a cDNA clone encoding a glycine-rich cuticular protein of *Tenebrio molitor*: developmental expression and effect of a juvenile hormone analogue. Insect Mol. Biol. 1, 53–62.

Charles, J.-P., Bouhin, H., Quennedey, B., Courrent, A., Delachambre, J., 1992. cDNA cloning and deduced amino acid sequence of a major, glycine-rich cuticular protein from the coleopteran *Tenebrio molitor*. Temporal and spatial distribution of the transcript during metamorphosis. Eur. J. Biochem. 206, 813–819.

Fraenkel, G., Rudall, K.M., 1947. The structure of insect cuticles. Proc. Roy. Soc. B 34, 111–143.

Hackman, R.H., Goldberg, M., 1979. Some conformational studies of larval cuticular proteins from *Calliphora vicina*. Insect Biochem. 9, 557–561.

Hamodrakas, S.J., 1984. Twisted β-pleated sheet: the molecular conformation which possibly dictates the formation of the helicoidal architecture of several proteinaceous eggshells. Int. J. Biol. Macromol. 6, 51–53.

Hamodrakas, S.J., Kanellopoulos, P.N., Pavlou, K., Tucker, P.A., 1997. The crystal structure of the complex of concanavalin A with 4-methylumbelliferyl-α-D-glucopyranoside. J. Struct. Biol. 118, 23–30.

Hamodrakas, S.J., Willis, J.H., Iconomidou, V.A., 2002. A structural model of the chitin-binding domain of cuticle proteins. Insect Biochem. Mol. Biol. 32, 1577–1583.

Huang, X., Zhang, H., Zen, K.C., Muthukrishnan, S., Kramer, K.J., 2000. Homology modeling of the insect chitinase catalytic domain-oligosaccharide complex and the role of a putative active site tryptophan in catalysis. Insect Biochem. Mol. Biol. 30, 107–117.

Iconomidou, V.A., Willis, J.H., Hamodrakas, S.J., 1999. Is ß-pleated sheet the molecular conformation which dictates the formation of the helicoidal cuticle? Insect Biochem. Mol. Biol. 29, 285–292.

Iconomidou, V.A., Chryssikos, G.D., Gionis, V., Willis, J.H., Hamodrakas, S.J., 2001. "Soft"-cuticle protein secondary structure as revealed by FT-Raman, ATR FT-IR and CD spectroscopy. Insect Biochem. Mol. Biol. 31, 877–885.

Katouno, F., Taguchi, M., Sakurai, K., Uchiyama, T., Nikaidou, N., Nonaka, T., Sugiyama, J., Watanabe, T., 2004. Importance of exposed aromatic residues in chitinase B from *Serratia marcescens* 2170 for crystalline chitin hydrolysis. J. Biochem. (Tokyo) 136 (2), 163–168.

Kerwin, J.L., Turecek, F., Xu, R., Kramer, K.J., Hopkins, T.L., Gatlin, C.L., Yates III, J.R., 1999. Mass spectrometric analysis of catechol-histidine adducts from insect cuticle. Anal. Biochem. 268, 229–237.

Kramer, K.J., Kanost, M.R., Hopkins, T.L., Jiang, H., Zhu, Y.C., Xu, R., Kerwin, J.L., Turecek, F., 2001. Oxidative conjugation of catechols with proteins in insect skeletal systems. Tetrahedron 57, 385–392.

Neville, A.C., 1975. Biology of the Arthropod Cuticle. Springer, New York.

Nicholls, A., Sharp, K.A., Honig, B., 1991. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins: Struct. Funct. Genet. 11, 281–296.

Rebers, J.E., Riddiford, L.M., 1988. Structure and expression of a *Manduca sexta* larval cuticle gene homologous to *Drosophila* cuticle genes. J. Mol. Biol. 203, 411–423.

Rebers, J.E., Willis, J.H., 2001. A conserved domain in arthropod cuticular proteins binds chitin. Insect Biochem. Mol. Biol. 31, 1083–1093.

Suetake, T., Tsuda, S., Kawabata, S.I., Miura, K., Iwanaga, S., Hikichi, K., Nitta, K., Kawano, K., 2000. Chitin-binding proteins in invertebrates and plants comprise a common chitin-binding structural motif. J. Biol. Chem. 275, 17929–17932.

Tews, I., Scheltiga, T., Perrakis, A., Wilson, K.S., Dijkstra, B.W., 1997. Substrate-assisted catalysis unifies 2 families of chitinolytic enzymes. J. Am. Chem. Soc. 119, 7954–7959.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucl. Acids Res. 22, 4673–4680.

Togawa, T., Nakato, H., Izumi, S., 2004. Analysis for the chitin recognition mechanism of cuticle proteins from the soft cuticle of the silkworm, *Bombyx mori*. Insect Biochem. Mol. Biol. 34, 1059–1067.

Van Gunsteren, W.F., Berendsen, H.J., 1987. BIOMOS, Biomolecular software. Laboratory of Physical Chemistry, University of Groningen, The Netherlands.

Vakser, I.A., 1996. Low-resolution docking: prediction of complexes for undetermined structures. Biopolymers 39, 455–464.

Vincent, J.F.V., Wegst, U.G.K., 2004. Design and mechanical properties of insect cuticle. Arthropod Struct. Dev. 33, 187–199.

Vriend, G., 1990. WHAT IF: a molecular modeling and drug design package. J. Mol. Graph. 8, 52–56.

Vyas, N.K., 1991. Atomic features of protein–carbohydrate interactions. Curr. Opin. Struct. Biol. 1, 732–740.

Watanabe, T., Ariga, Y., Sato, U., Toratani, T., Hashimoto, M., Nikaidou, N., Kezuka, Y., Nonaka, T., Sugiyama, J., 2003. Aromatic residues within the substrate-binding cleft of *Bacillus circulans* chitinase A1 are essential for hydrolysis of crystalline chitin. Biochem. J. 376, 237–244.

Willis, J.H., 1999. Cuticular proteins in insects and crustaceans. Am. Zool. 39, 600–609.

Willis, J.H., Iconomidou, V., Smith, R.F., Hamodrakas, S.J., 2005. Cuticular proteins. In: Gilbert, L., Iatrou, K., Gill, S. (Eds.), Comprehensive Molecular Insect Science, vol. 4. Elsevier, Oxford, pp. 79–110 Chapter 2.

Zanotti, G., Marcello, M., Malpeli, G., Folli, C., Sartori, G., Berni, R., 1994. Crystallographic studies on complexes between retinoids and plasma retinol-binding protein. J. Biol. Chem. 269, 29,613–29,620.