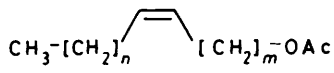
Chain-elongated Analogues of a Pheromone Component of the Turnip Moth, *Agrotis segetum*. A Structure–Activity Study using Molecular Mechanics

Tommy Liljefors * and Bernt Thelin
Department of Organic Chemistry 3, Chemical Centre, University of Lund, P.O. Box 124, S-221 00 Lund, Sweden
Jan N. C. van der Pers and Christer Löfstedt
Department of Animal Ecology, Ecology Building, University of Lund, S-223 62 Lund, Sweden

Chain-elongated analogues of (*Z*)-dec-5-enyl acetate, a pheromone component of the turnip moth, *Agrotis segetum*, have been studied. The conformational energies required for the analogues to mimic spatial relationships in the parent molecule, assumed to be crucial for the receptor interaction, were calculated by molecular mechanics (MM2). The calculated energies show a striking correlation with measured single-cell electrophysiological activities. The results indicate that an elongated alkyl chain is conformationally rearranged when the analogue is bound to the receptor, and that the biological activity is determined by the corresponding conformational energy.

The great majority of sex pheromone components so far identified from noctuid moths are straight-chain mono-olefinic acetates with a *Z*-double bond.[1] The pheromone blend produced by the female moth may include several components of this type differing in chain length and/or double-bond position. Electrophysiological studies have shown that each of the male olfactory receptor cells generally responds with high selectivity to one of the components in the pheromone blend.[2,3] The receptor cells are thus able very efficiently to discriminate between components which differ in length of the alkyl chains (*m* and *n*) separated by the *Z*-double bond [structure (A)].



(A)

Receptor cells specific for straight-chain mono-olefinic acetates in different noctuid moth species respond very similarly to structural variations of their natural stimulus molecule.[4] This indicates that the decisive properties of this class of receptors are similar for related species, and that studies on receptor cell responses for one moth species may have more general validity.

Discussions on the mode of interaction between the stimulus molecule and its receptor have been centred around three prominent parts of the pheromone component molecule: the terminal methyl group, the double bond, and the polar functional group.[5–7] It has then been assumed that the biological receptor includes binding sites complementary to these parts of the substrate molecule. The lengths and the conformations of the alkyl chains connecting the three molecular parts determine their spatial relationships. A study of the receptor response to chain-elongated or chain-shortened analogues may thus give information on the sensitivity of the substrate–receptor interaction to variations in the mutual relationships in space of the three groups. This knowledge may be used to deduce the effective dimensions of the receptor region that interacts with the stimulus molecule, and give insight into the basis for the receptor selectivity at a molecular level.

Straight-chain mono-olefinic acetates are flexible molecules. The possibility of internal rotations about carbon–carbon single bonds with low energy barriers gives rise to a very large number of conformations with different geometrical relationships between the terminal methyl group, the double bond and the acetate group. Bestmann *et al.* recognize this flexibility and assume a flexible insertion of the stimulus molecule into the entire receptor region.[7,8] The thermodynamic and kinetic aspects of such a dynamic model (the 'zipper' model) have been discussed in connection with studies on peptide hormones and the double-helix formation of nucleic acids.[9,10] This model implies that an elongated alkyl chain may be accommodated in the receptor region by 'winding' of the chain. A conformational rearrangement of the alkyl chain(s) may make it possible for the acetate group, the double bond, and the terminal methyl group in chain-elongated analogues to interact optimally with their complementary receptor sites. The equilibrium constant and thus the free energy of binding of the stimulus molecule to the receptor, or the formation of a substrate–receptor 'activated complex,' will then depend on the free energy of the 'biologically active' conformation of the molecule. The activity should decrease with increasing conformational energy required to obtain a 'correct' fit between critical parts of the stimulus molecule and their receptor sites.

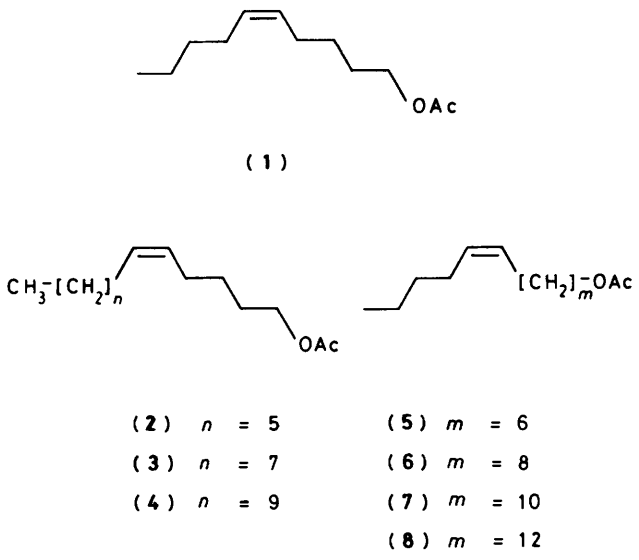
The purpose of the study presented in this paper is to investigate if the conformational energies of chain-elongated analogues, conformationally rearranged in order to mimic geometrical features of the natural pheromone component, are correlated with observed biological activities. We report electrophysiological single-cell measurements and conformational analysis of analogues of (*Z*)-dec-5-enyl acetate, a pheromone component of the turnip moth, *Agrotis segetum* (Schiff.) (Lepidoptera: Noctuidae). Receptor cells selectively tuned to this compound have previously been identified in the male antennae of this species and are readily accessible for single-cell recordings.[3,11,12]

The compounds studied have up to eight methylene groups added to the natural pheromone component (**1**) and can be divided into two groups as shown in Scheme 1. Compounds (**2**)—(**4**) have the alkyl chain between the *Z*-double bond and the terminal methyl group elongated, while compounds (**5**)—(**8**) have the corresponding variations in the alkyl chain between the double bond and the acetate group.

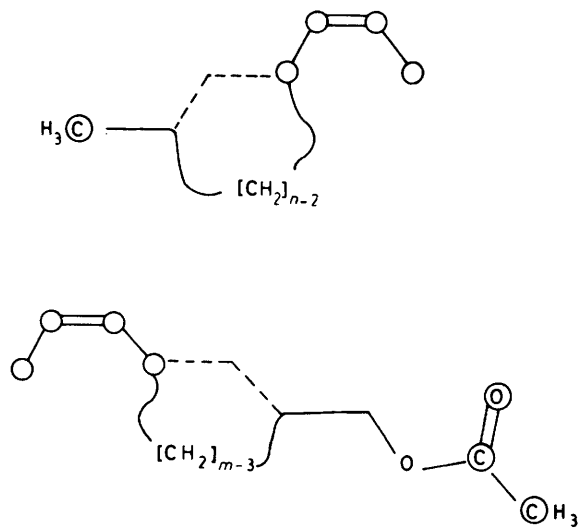
By using a simple model for the interaction of the substrate with the receptor, the geometries and conformational energies of the deduced biologically active conformations were calculated employing the molecular mechanics method (MM2).[13,14]

Materials and Methods

Chemicals.—(*Z*)-Dec-5-enyl acetate (**1**), (*Z*)-dodec-7-enyl acetate (**5**), and (*Z*)-tetradec-9-enyl acetate (**6**) were synthesized



Scheme 1.



Scheme 2. Constrained atoms (encircled) in the calculations of structures and energies for biologically active conformations. Dashed lines indicate the alkyl chains in the natural pheromone component (1)

and purified as previously described.[15] Compounds (2)—(4), (7), and (8) were purchased from the Institute for Pesticide Research, Wageningen, The Netherlands. All compounds were at least 98% pure with respect to positional and geometrical isomers.

Calculations.—Energy-minimized geometries and conformational energies were calculated using the molecular mechanics program MM2 developed by Allinger and his co-workers.[13,14] For computational expediency only those parts of the molecules which are changed in a series were included in the calculations. The parts not included are constant within a series and furthermore do not significantly influence the conformational properties of included parts. Thus, compounds (2)—(4) are modelled by homologues of (*Z*)-hept-2-ene and compounds (5)—(7) by homologues of (*Z*)-hept-5-enyl acetate.

Electrophysiology.—The electrophysiological activities of the compounds were determined by measuring the electrical responses of single receptor cells on the antennae of male *A. segetum* moths during stimulation with compounds (1)—(8).

Freshly excised antennae of two- to four-day-old moths were used. The electrophysiological recording technique was essentially the same as applied by van der Pers and den Otter.[16] Electrical contact was made with receptors highly selective to (*Z*)-dec-5-enyl acetate (1).[3,12,17] Stimulation of these receptors was achieved by mixing 2 ml from the gaseous contents of disposable syringes into the air stream (1 m s^{-1}) flowing over the antennal preparation. The syringes contained pieces of filter paper on which 10^{-3} to 10^{3} μg of the compounds had been applied. The response of the receptor was defined as the number of action potentials generated by the receptor cell during 1 s of stimulation. Each compound was tested using at least two different stimulus concentrations.

The relative activities were calculated from dose–response curves and expressed as the relative quantities required to elicit the same response of the receptor cell.

The Substrate–Receptor Interaction Model

The model used in the present work assumes geometrically well defined receptor sites complementary to the terminal methyl group, the *Z*-double bond and the acetate group. The spatial relationship between the double bond and the preferred position of the methyl group when bound to the receptor site is defined by the relative positions in space of the corresponding molecular parts in the energy-minimized geometry of (*Z*)-hept-2-ene. The preferred position of the acetate group is correspondingly determined by the geometry of (*Z*)-hept-5-enyl acetate. The fully extended all-*anti*-conformation of the alkyl chains was assumed in both cases. Structure–activity studies on analogues of (1) with an extra *E*-double bond in different positions to simulate an *anti*-conformation about that bond indicate that this is the biologically active conformation of the alkyl chains in (1).[18] Furthermore, the all-*anti*-conformation is the thermodynamically most stable one (see later).

We further assume that a chain-elongated analogue interacts with the receptor with the terminal methyl group, the double bond, and the acetate group in the relative positions in space as already defined. This means that the analogue must rearrange conformationally in order to place the crucial molecular parts in the required positions in space.

Computationally this is accomplished by restricting the encircled atoms in Scheme 2 to fixed positions during the energy-minimization procedure, while the remaining alkyl-chain atoms and other non-restricted atoms are allowed to find positions that minimize the total energy of the molecule.

In terms of substrate–receptor binding energies this model implies that all analogues have closely similar interaction energies with the assumed receptor sites. The different abilities of compounds (2)—(8) to bind to the receptor, or to form an 'activated complex' with the receptor, are then, according to the model, due to different conformational energies of the biologically active conformations.

To find possible conformations of the alkyl chains which give the acetate or the terminal methyl group the required position in space relative to the double bond, we employed a tetrahedral lattice (diamond lattice) built around the energy-minimized geometries of (*Z*)-hept-2-ene and (*Z*)-hept-5-enyl acetate, respectively. Pathways on the lattice linking the double bond with the terminal methyl group and the acetate group, respectively, were selected as trial conformations for the energy-minimization procedure. These pathways are characterized by perfect staggering at each carbon–carbon bond in the alkyl chain.

For a 'correct' binding to the terminal methyl site, the direction of the C–CH_3 bond vector in the elongated-chain

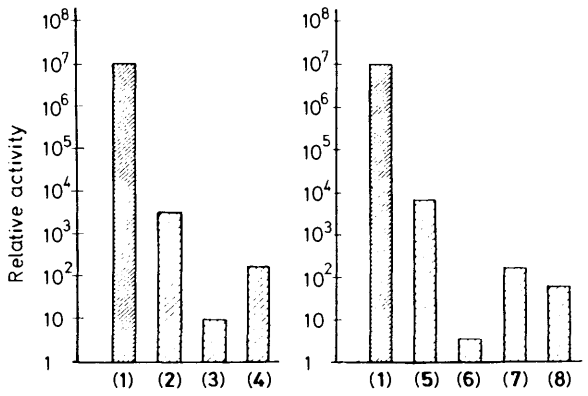


Figure 1. Relative single-cell activities for compounds (**1**)—(**8**). The activities are expressed as the reciprocal of the relative number of moles required to elicit the same receptor response

analogues should be as close as possible to the corresponding vector in the natural pheromone component. This is ensured by requiring the carbon atom next to the terminal methyl carbon to be included in an acceptable lattice pathway (see Scheme 2). With this choice we also make sure that the chain does not have serious repulsive steric interactions with the methyl group site in the receptor region. For the homologues of (*Z*)-hept-5-enyl acetate the same requirement was imposed on the carbon next to the CH_2–O bond (see Scheme 2). In this way the carbonyl group may bind to its assumed complementary receptor site without steric hindrance by the 'winding' chain.

All conformations on the tetrahedral lattice which are in accordance with the foregoing geometrical requirements were determined and subjected to energy-minimization by MM2 with the constraints on the movements of the encircled atoms in Scheme 2.

The construction of the tetrahedral lattice and the trial conformations was greatly facilitated by the use of computer graphics employing the MOLBUILD part of the interactive molecular graphics system MIMIC.[19]

In a second series of calculations the global unconstrained energy minima of the same molecules were calculated. The conformational energy of the calculated biologically active conformation could then be evaluated for each compound by taking the calculated energy difference between the lowest energy conformationally rearranged structure which fits the model and the thermodynamically most stable one.

Results and Discussion

Receptor Cell Responses.—The measured electrophysiological single cell activities of compounds (**2**)—(**4**) and (**5**)—(**8**) relative to the natural pheromone component (**1**) are shown in Figure 1. The addition of two and four methylene groups to compound (**1**) results in both series in a monotonic decrease of the electrophysiologal activity, as previously observed in other moth species.[2,4,20–22] However, elongation of the chains by six methylene groups [compounds (**4**) and (**7**)] increases the activity as compared with four methylene units of elongation. In the second series a further addition of two methylene units [compound (**8**)] produces only a minor change of activity. There is thus a pronounced minimum in the measured electrophysiological activity corresponding to four methylene units of elongation in both series. Since other studies have only considered up to four methylene units of elongation, such a minimum has not previously been observed.

Compounds (**1**)—(**8**) have significantly different vapour pressures. Since the measurements shown in Figure 1 refer to amounts of stimulus molecule loaded on a filter paper (see Materials and Methods section), differences in volatility should be taken into account. The vapour pressures of straight-chain olefinic acetates are an approximately linear function of the number of carbon atoms; additions of two methylene groups decreases the vapour pressure by a factor of 5—8.[15] If we assume that the relative amounts of the stimulus compounds in the air stream for a fixed quantity loaded on the filter paper are determined by their relative saturated vapour pressures, the activities shown in Figure 1 should be increased by this factor with respect to the next lower homologue in the series and decreased by the same factor with respect to the next higher one. These corrections do not change the relative orders of the electrophysiological activities shown in Figure 1. However, they somewhat accentuate the minimum of activity for (**3**) and (**6**).

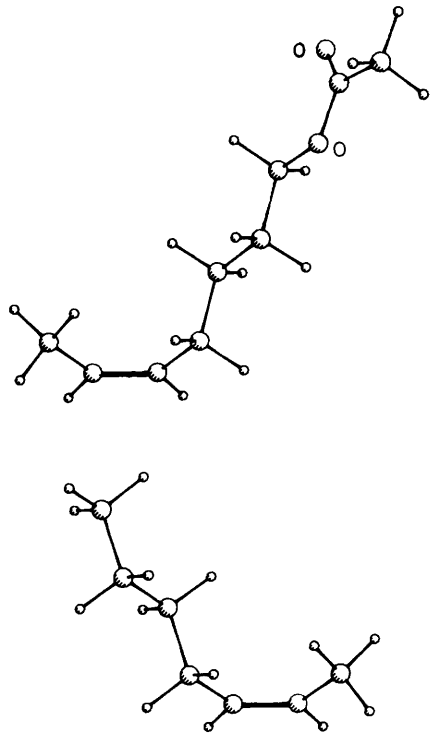


Figure 2. Calculated lowest energy conformations of (*Z*)-hept-2-ene and (*Z*)-hept-5-enyl acetate, model compounds for (*Z*)-dec-5-enyl acetate (**1**)

Calculated Structures and Conformational Energies.—The calculated global energy minima for (*Z*)-hept-2-ene and (*Z*)-hept-5-enyl acetate are shown in Figure 2. These molecules are calculated to prefer strongly the skew conformation. There are no experimental data on the structures and conformations of these molecules to compare with, but for the analogous (*Z*)-pent-2-ene the skew conformer has been found to be the only one significantly populated.[23,24] In both molecules in Figure 2 the alkyl chains prefer the all-*anti*-conformation which is also the case for the chain-elongated analogues.

The calculated differences in conformational energies between the lowest energy structures suitable for receptor interaction and the global energy minima, according to our model, are shown in Figure 3. These energies correspond to the energies (enthalpies) required to bring the molecules from their preferred conformations to their calculated biologically active conformations. The conformational energy increases with the addition of two and four methylene groups to the reference compounds. However, when one more methylene group is

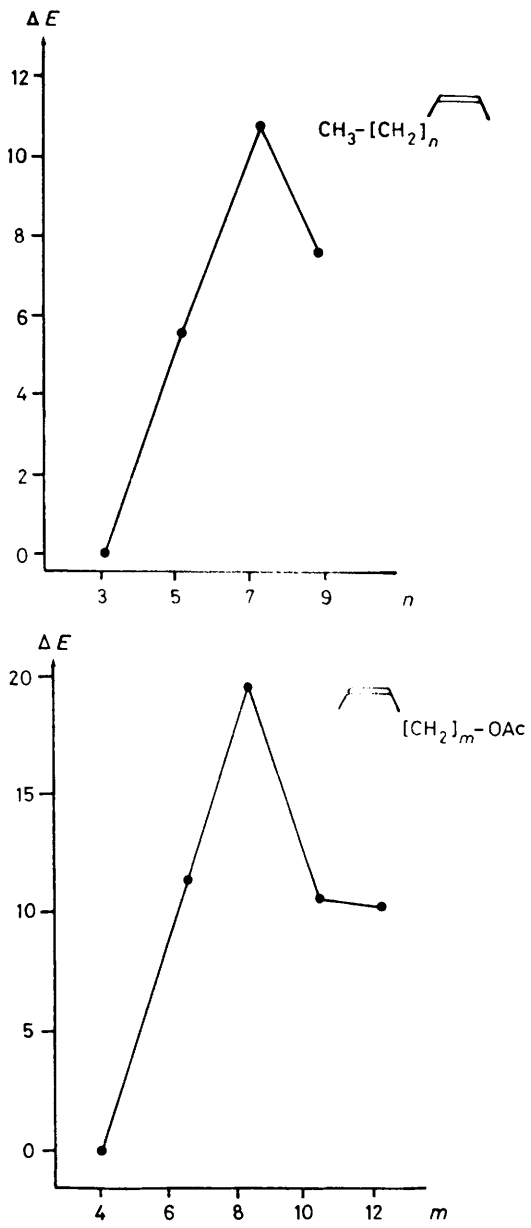


Figure 3. Calculated energy differences in kcal mol^{-1} between the biologically active and the all-*anti*-conformations for different chain lengths

added [$n = 9$ and $m = 10$, corresponding to (**4**) and (**7**), respectively] the conformational energy decreases and stays approximately constant with the further addition of two methylene groups [$m = 12$, corresponding to (**8**)]. In comparing the results in Figures 1 and 3, there is a striking correspondence between the energy required to bring a chain-elongated analogue into the biologically active conformation, according to our model, and the observed electrophysiological activity. An increase of conformational energy corresponds to a decrease of measured activity.

The use of conformational enthalpies in the foregoing discussion assumes that it is possible to neglect the entropy contributions. The translational and rotational contributions to the entropy should be closely similar for the compounds studied. Since the addition of two methylene groups to a chain leads to two more 'frozen' internal rotations in the calculated 'active' structures, the conformational entropy term should be increasingly negative with increasing chain-elongation. From studies on cyclizations of hydrocarbon chains, an entropy loss of *ca.* 4 cal mol^{-1} K^{-1} for 'freezing' one internal rotation has been inferred.[25] This value may be taken as an upper limit in the present case. If we adopt this value, the calculated conformational energies in Figure 3 should be increased by *ca.* 1 kcal mol^{-1} for each additional methylene group to obtain the corresponding free energies. However, these entropy contributions should be largely compensated by increasing hydrophobic binding (*ca.* 0.9 kcal mol^{-1} per methylene group[26]).

Kafka and Neuwirth proposed a model for the interaction between a pheromone component and its receptor in which the decrease in activity for various analogues is a consequence of displacements of the double bond and/or the acetate group from the positions with maximum binding to the assumed complementary receptor sites.[5] This model predicts a monotonic decrease of activity with alkyl-chain elongation until it levels off for sufficiently long chains, in disagreement with the foregoing results. The receptor interaction model used in the present work, however, explains the experimental observations very satisfactorily. It thus seems clear that the flexibility of alkyl chains must be taken into account in attempts to model substrate–receptor interactions for moth sex pheromone components and their analogues. This flexibility may make it possible for an analogue to 'mimic' closely the important features of the natural pheromone component and thus interact with the receptor with the critical molecular parts in 'correct' positions. The probability of binding to the receptor then depends on the energy of the conformationally rearranged structure relative to that of the thermodynamically preferred one.

The energy-minimized structures corresponding to the lowest energy biologically active conformations of the chains with a terminal methyl group are shown in Figure 4. Since the number of methylene groups involved in 'chain winding' is the same for the corresponding compounds in both series (see Scheme 2), the structures of the acetate substituted chains are very similar.

The large conformational energies calculated for the active conformations of the chains with $n = 7$ and $m = 8$ (Figure 3), corresponding to compounds (**3**) and (**6**), are due to the restricted ability of these alkyl chains to relieve repulsive steric interactions. To fit the receptor interaction model, the alkyl chains in the compounds studied must form loops on the tetrahedral lattice with five ($n = 5$; $m = 6$), seven ($n = 7$; $m = 8$), nine ($n = 9$; $m = 10$), or eleven carbon atoms ($m = 12$) including the two loop-end atoms. These conformations are characterized by a gauche(+)–gauche(−) arrangement about two adjacent bonds which causes severe repulsive interactions between hydrogen atoms pointing 'inwards' into the loops and competing for the same position on the tetrahedral lattice (Figure 4). The number of such interactions increases with increasing chain-length. However, the larger, more flexible loops ($n > 7$ or $m > 8$) are able efficiently to relieve a substantial part of the steric strain imposed by the receptor-interaction model.

The experimental receptor responses shown in Figure 1 indicate that the sensitivities of the biological activity to chain elongation are similar for the two series of compounds. This has also been shown in studies on other moth species.[2,19,21] In some investigations it has been concluded that the activity is somewhat less sensitive to changes in the chain connecting the double bond and the acetate group.[22] In contrast, the calculated conformational energies for the alkyl chains in compounds (**5**)–(**8**) (Figure 3) predict a much greater dependence of electrophysiological activity on elongation of the acetate-substituted chain than on a corresponding methylene-group addition to the non-polar end of the molecule. Inspection of the calculated

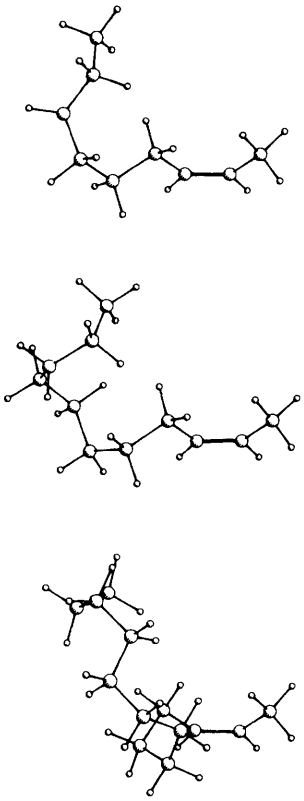


Figure 4. Calculated biologically active structures for the non-polar end of compounds (**2**)—(**4**)

geometries of the biologically active conformations of compounds (**5**)—(**8**) shows that the difference in conformational energies between corresponding compounds in the two series is mainly due to deformations of the acetate group as a result of the imposed constraints (see Scheme 2). This indicates that the constraints imposed on the relaxation of the acetate group are probably too severe. This group most likely binds to the receptor through electrostatic forces and/or hydrogen bonds. The terminal methyl group can only interact with the receptor through dispersion forces, which fall off with the inverse sixth power of the distance between interacting centres. A close fit is thus required for the methyl group–receptor site interaction. In contrast, the interactions between the acetate group and its receptor site are much less distance-dependent, and sufficiently good interaction may be obtained with the acetate group at somewhat different positions in the receptor. A refined model should thus allow for more freedom to the acetate group than was the case in the present work.

The foregoing analysis is only applicable to analogues in which an even number of methylene groups has been added to the natural pheromone component; only such alkyl chains fit into the tetrahedral lattice in the model used in this study. However, it is of interest to investigate the extension of the model to include analogues in which an odd number of methylene groups has been added. Since such analogues have not yet been tested on the (*Z*)-dec-5-enyl acetate receptor of *Agrotis segetum* we will only consider the addition of one methylene group to the non-polar end of the pheromone component. In other moth species such chain-elongated analogues have been shown to be more active than an analogue with two additional methylene groups.[4,22]

The geometrical constraints defined above were imposed on (*Z*)-oct-2-ene, used as model for (*Z*)-undec-5-enyl acetate, and

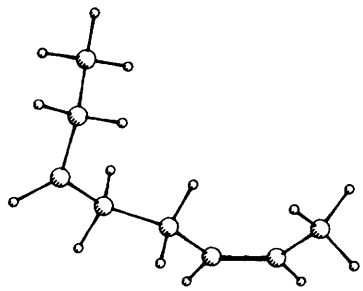


Figure 5. Calculated biologically active structure for the non-polar end of (*Z*)-undec-5-enyl acetate

the lowest energy conformation of the four-membered chain that fits the receptor model was calculated. The resulting energy-minimized structure is shown in Figure 5. This conformation is calculated to be 4.3 kcal mol^{-1} higher in energy than the corresponding all-*anti*- global minimum conformation. The conformational energy is thus lower than that for (*Z*)-non-2-ene ($n = 5$ in Figure 3), which means that (*Z*)-undec-5-enyl acetate is predicted to be significantly more active than compound (**2**). This agrees with experiments for other moth species already referred to. The model used in the present work thus seems to work equally well for odd-numbered and even-numbered elongations of alkyl chains.

Conclusions

The flexibility of the alkyl chains in mono-olefinic straight-chain acetates makes it possible for the terminal methyl group, the double bond, and the acetate group in chain-elongated analogues to occupy positions in space very similar to those of the corresponding groups in the parent molecule. A simple receptor-interaction model based on this possibility satisfactorily explains observed biological activities for this type of analogues. This indicates that the elongated chains are conformationally rearranged when the analogues are bound to the pheromone receptor. The energy required for this rearrangement is the determining factor for the relative biological activities. The results obtained imply that the receptor region binding the stimulus molecule has a restricted extension in space, corresponding to the molecular extension of the natural substrate.

Acknowledgements

We thank Dr. Robert E. Carter for linguistic criticism. This work was supported by grants to Dr. Jan Löfqvist and Dr. Tommy Liljefors from the Swedish Research Councils (N.F.R., F.R.N., and S.J.F.R.).

References

1 W. Steck, E. W. Underhill, and M. D. Chisholm, *J. Chem. Ecol.*, 1982, **8**, 731.
2 E. Priesner in 'Chemical Ecology: Odour Communication in Animals,' ed. F. J. Ritter, Elsevier/North Holland, Amsterdam, 1979, p. 57.
3 C. Löfstedt, J. N. C. van der Pers, J. Löfqvist, B. S. Lanne, M. Appelgren, G. Bergström, and B. Thelin, *J. Chem. Ecol.*, 1982, **8**, 1305.
4 E. Priesner, *Ann. Zool. Ecol. Anim.*, 1979, **11**, 533.
5 W. A. Kafka and J. Neuwirth, *Z. Naturforsch., Teil C*, 1975, **30**, 278.
6 W. L. Roelofs and A. Comeau, *J. Insect Physiol.*, 1971, **17**, 435.

7 H. J. Bestmann and O. Vostrowsky in 'Olfaction and Endocrine Regulations,' ed. W. Breipohl, IRL Press, London, 1982, p. 253.
8 H. J. Bestmann, P. Rosel, and O. Vostrowsky, *Liebigs Ann. Chem.*, 1979, 1189.
9 A. S. V. Burgen, G. C. K. Roberts, and J. Feeney, *Nature (London)*, 1975, **253**, 753.
10 J. Engel and D. Winklmair, *FEBS Symp.*, 1972, **29**, 29.
11 E. Hallberg, *Cell Tissue Res.*, 1981, **218**, 209.
12 J. N. C. van der Pers, and C. Löfstedt, *Physiol. Entomol.*, 1983, **8**, 203.
13 N. L. Allinger and Y. H. Yuh, Quantum Chemistry Program Exchange, 1980, vol. 11, program no. 395.
14 U. Burkert and N. L. Allinger, 'Molecular Mechanics,' American Chemical Society, Washington, D. C., 1982.
15 A.-M. Olsson, J. Å. Jönsson, B. Thelin, and T. Liljefors, *J. Chem. Ecol.*, 1983, **9**, 375.
16 J. N. C. van der Pers and C. J. den Otter, *J. Insect Physiol.*, 1978, **24**, 337.
17 J. N. C. van der Pers and C. Löfstedt, Proceedings of the NSF-NATO Seminar on Mechanisms of Perception and Orientation to Insect Olfactory Signals, Oxford, 1984, in the press.
18 M. Bengtsson, T. Liljefors, and B. Hansson, in preparation.
19 T. Liljefors, *J. Mol. Graphics*, 1983, **1**, 111.
20 E. Priesner, *Z. Naturforsch., Teil C*, 1983, **38**, 874.
21 E. Priesner in 'Insect Neurobiology and Insecticide Action (Neurotox 1979),' Society of Chemical Industry, London, 1980, p. 359.
22 E. Priesner, M. Jacobson, H. J. Bestmann, *Z. Naturforsch., Teil C*, 1975, **30**, 283.
23 B. P. van Eijk, *J. Mol. Spectrosc.*, 1981, **85**, 189.
24 J. H. M. Ter Brake, *J. Mol. Struct.*, 1984, **118**, 63.
25 M. A. Winnik, *Chem. Rev.*, 1981, **81**, 491.
26 C. Tanford, 'The Hydrophobic Effect. Formation of Micelles and Biological Membranes,' Wiley, New York, 1980.

Received 11*th March* 1985; *Paper* 5/410