Correlation of moth sex pheromone activities with molecular characteristics involved in conformers of bombykol and its derivatives

(olfaction/conformational analysis)

TOSHIHIDE KIKUCHI

Department of Biological Science, Tohoku University, Kawauchi, Sendai 980, Japan

ABSTRACT   Molecular characteristics of bombykol and its 11 derivatives, which reveal significant correlations with biological activities for single sex pheromone receptor cells of four moth species, Bombyx mori, Aglia tau, Endromis versicolora, and Deilephila euphorbiae, were examined on the assumption of the "bifunctional unit model." Probabilities of bifunctional unit formations of those 12 compounds were assessed with frequency distribution patterns of distances between the proton acceptor, the proton donor, and the methyl group involved in a total of 1,200 conformers. A highly significant correlation exists between biological activity for each species and the probability of a particular bifunctional unit formation: a proton acceptor (A)–a methyl group (Me) unit $(A–Me$ distances: about 6 Å) for Deilephila $(r = 0.94)$; a proton acceptor (A)–a proton donor (D) $(A–D$: about 11 Å) for Aglia $(r = 0.83)$; two antagonistic proton donor–methyl units $(D–Me$: about 14 and 5 Å for favorable and adverse unit, respectively) for Bombyx $(r = 0.94)$ and Endromis $(r = 0.92)$.

There has been increasing interest in demonstrating molecular characteristics of insect pheromones in governing their biological activities. For alarm pheromone in ants (1, 2) and trail-following pheromonein termites (3), the molecular shape or the vibrational pattern of these pheromone molecules is presumed to play an important role in governing biological activities, but the structure-activity relationships in a wide variety of other pheromones still remain undefined. For the sex pheromone of the silkworm moth, I now present evidence that the pheromone activity may be produced by a pair of functional groups occupying a particular interval ("bifunctional unit"), which is involved in certain conformers of this pheromone. Attention is, therefore, focused on the idea that "bifunctional unit" may be the basis of pheromone perceptions in insects and on the basic idea in the conformational analysis of pheromone molecules. The new idea, "bifunctional unit model," has emerged from the studies of an induced olfactory mutant of Drosophila with abnormal responses to 18 specific chemicals, the molecular feature common to these specific chemicals being a "bifunctional unit," consisting of a proton acceptor (A) and a proton donor (D) with an average $A–D$ distance of about 3 Å (5, 6). From this idea it became, therefore, of particular interest to see whether or not biological activities of other insect pheromones can be equally explained with reference to the "bifunctional units." Bombykol, the sex pheromone of the silkworm moth (Bombyx mori), and its 11 derivatives may well serve our purpose when used in examining the structure-activity relationships, because biological activities of these compounds have been definitely assayed with single cell recordings (6). To define probable "bifunctional units" involved with those compounds, precise information is required about the conformational states of these compounds in gas phase, since the

distance between functional groups involved in these compounds depends upon their conformations as well as their configurations. With considerations of frequency distribution patterns of the distances between a pair of methyl and hydroxyl groups involved with conformers of bombykol and its derivatives, the attempts described below could provide demonstrations on the relationships between pheromone activities and their molecular characteristics.

MATERIALS AND METHODS

Biological Activities. In the present analysis, biological activities of bombykol and its 11 derivatives on four lepidopterous species, Bombyx mori (Bombycidae), Aglia tau (Saturnidae), Endromis versicolora (Endromidae), and Deilephila euphorbiae (Sphingidae), are defined as the threshold values represented as minimal amounts of odor sources required to elicit potential changes on single sensory hairs or to evoke nerve impulses as were observed in single cell recordings (columns 2 to 5 of Table 1) (6).

Conformational Analysis of Pheromone Molecules. (a) "Functional regions." Both functional groups and π bonds involved in molecules, which might play a role in producing pheromone actions, are defined as "functional regions," and the combination of two "functional regions" is termed a "functional system." The "functional system" that has a particular interval of "functional regions" and plays a significant role in governing a biological activity is defined as a "bifunctional unit." Thus, the three "functional regions" that can be expected to become constituents of "bifunctional units" were considered (Fig. 1): (i) a proton donor (D) (—OH on C_1); (ii) one or two proton acceptors (A) (C_9 for compound 7, 10, or 12, both C_9 and C_{12} for compound 9, C_{10} for compound 5, C_{11} for compound 11, C_{10} double and C_{12} triple bond for compound 8, and a conjugated system of C_{10} and C_{12} for compound 1, 2, 3, 4, or 6 as shown in the first column of Table 1; and (iii) a C_{16}, C_{17}, or C_{18} methyl group (Me). From the combinations of those three regions, three possible types of functional systems can be assumed: (i) a proton acceptor (A)–proton donor (D) $(A–D$ system); (ii) a proton acceptor (A)–methyl group (Me) $(A–Me$ system); and (iii) a proton donor (D)–methyl group (Me) $(D–Me$ system). As for compounds 8 and 9, possessing two proton acceptors of the C_{10} double and C_{12} triple bond (compound 8, Fig. 1) and two of the C_9 and C_{12} double bond (compound 9), respectively, two $D–A$ and two $A–Me$ systems were considered.

(b) "Station" for measurements of distances between functional regions. To take accurate measurements of distances between "functional regions," —OH proton, C_{16},

C_{17}, or C_{18} methyl C atom, and the center of π bonds (or the center of a conjugated system) were used as the "stations."

(c) *Basic idea in conformational analysis.* The distances between a pair of "functional regions" involved in bombykol and its derivatives depend upon both their configurations and conformations derived by rotations about each single bond, but at present little is known about the conformations of these alcohols. In the present conformational analysis, the application of the concept obtained from an analysis of simple acyclic hydrocarbons such as *n*-butane and *n*-pentane to that of these alcohols would be, therefore, of great utility in considering their conformational states and the distances between a pair of the "functional regions." For *n*-butane, there are three conformations to be considered, one in which the methyl groups are as far apart as possible in a zigzag *anti* plane structure and two equivalent conformations derived by a rotation of approximately 120° about the central single bond in either direction. The latter two conformations, called *gauche*, are of somewhat higher energy level than the former, *anti* form. The difference between free energy in a given conformer and that in the conformer that gives the minimal free energy level has been defined as "conformational free energy" ($-\Delta G°$) (7). From this definition, the "conformational free energy" ($-\Delta G°_{25}$) of the *gauche* form in *n*-butane is estimated to be about 0.4 kcal; the equilibrium constant K [(*anti*)/(*gauche*)] being about 2.0 (7). For the application of such a concept to *n*-pentane, this compound can be divided conveniently into two "*n*-butane systems" or "three-bond systems": C_1—C_2—C_3—C_4 and C_2—C_3—C_4—C_5. Thus, there are nine conformers to consider ($3^2 = 9$); one *anti-anti*, two *anti-gauche*, two *gauche-anti*, and four *gauche-gauche* forms.
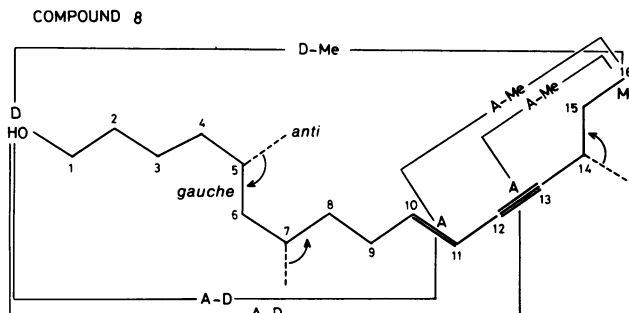

FIG. 1.   A conformer of hexadeca-10-*trans*-en-12-yn-1-ol (compound 8, Table 1) with two restricted rotations about C_4—C_5 and C_6—C_7 axes and a free rotation about C_{13}—C_{14} axis.

For further applications of such procedures to bombykol and its derivatives, they were divided conveniently into certain "three-bond systems": H—O—C_1—C_2, O—C_1—C_2—C_3, C_1—C_2—C_3—C_4, ... Thus, the number of possible conformers involved in each compound is given according to the following formula:

$$N = 3^{s \cdot c \cdot 2} \quad \text{or} \quad 3^{n \cdot 2} \qquad [1]$$

where N is the number of possible conformers, s is the number of C—C, C—O, and O—H single bonds, c is the number of the C—C single bond involved in the conjugated system, $n(= s - c)$ is the number of single bonds to rotate in either of three restricted directions, and $n - 2$ is the number of "three-bond systems." With this formula the number of conformers involved in the 12 compounds were estimated as listed in column 5 of Table 1. Since compound 8 possesses

Table 1.   Sex pheromone receptor thresholds for four moth species and probabilities of bifunctional unit formations*

| Compounds | Threshold values (μg) | | | | Probabilities ($\times 10^{-2}$) | | | | | |
	Deilephila	Aglia	Bombyx	Endromis	P_{f-1}	P_{f-2}	P_{f-3}	P_a	P_o $(P_{f-3}- 2P_a)$	P'_o $(P_{f-3}- 3P_a)$
1. Hexadeca-10-*trans*, 12 *cis*-dien-1-ol	10^{-1}	10^{-1}	10^{-3}	10^{-2}	60	37	40	1	38	37
2. Hexadeca-10-*cis*, 12-*trans*-dien-1-ol	10^{-1}	10^{-1}	10^{-1}	10^{-1}	65	33	37	3	31	28
3. Hexadeca-10-*cis*, 12-*cis*-dien-1-ol	10^{0}	10^{0}	10^{0}	10^{0}	58	38	34	2	30	28
4. Hexadeca-10-*trans*, 12-*trans*-dien-1-ol	10^{-1}	10^{-2}	10^{0}	10^{-1}	70	39	32	0	32	32
5. Hexadeca-10-*trans*-en-1-ol	10^{-1}	10^{-1}	10^{0}	10^{-1}	49	37	34	0	34	34
6. Heptadeca-10-*cis*, 12-*cis*-dien-1-ol	10^{1}	10^{-1}	10^{1}	10^{1}	42	36	34	2	30	28
7. Hexadeca-9-*cis*, en-1-ol	10^{1}	10^{1}	10^{2}	10^{3}	23	27	36	5	26	21
8. Hexadeca-10-*trans*-en-12-yn-1-ol	–	10^{-2}	10^{2}	10^{2}	23†,0‡	36†	33	4	25	21
9. Octadeca-9-*cis*, 12-*cis*-dien-1-ol	10^{2}	10^{0}	10^{3}	$>10^{3}$	9§	30¶	38	8	22	14
10. Octadeca-9-*cis*-en-1-ol	10^{2}	10^{0}	10^{3}	$>10^{3}$	2	28	28	4	20	16
11. Octadeca-11-*trans*-en-1-ol	10^{2}	10^{-1}	$>10^{3}$	$>10^{3}$	21	40	24	6	12	6
12. Octadeca-9-*trans*-en-1-ol	10^{3}	10^{2}	$>10^{3}$	$>10^{3}$	3	23	21	7	7	0

* P_{f-1} for unit F-1 in Fig. 2a, P_{f-2} for unit F-2 in Fig. 2b, P_{f-3} for unit F-3 and P_a for unit A in Fig. 2c.
† For C_{10}=;   ‡ C_{12}= ;   § C_9= ;   ¶ C_{12}=.

Table 2. Molecular characteristics of 12 compounds and their conformers to be sampled

Compounds*	s	c	n	N	Conformers containing r gauche forms $(Nr/N, \%)$†											
					$r = 0$	1	2	3	4	5	6	7	8	9	10	$11\leq$
8	15	0	$13^{\ddagger}$	3^{11}	2	9	20	26	22	13	6	2	0	0	0	0
1, 2, 3, 4	15	1	14	3^{12}	1	7	17	24	23	16	8	3	1	0	0	0
6	16	1	15	3^{13}	1	5	14	22	24	18	11	4	1	0	0	0
5, 7	16	0	16	3^{14}	1	4	11	19	23	20	13	6	2	1	0	0
9	17	0	17	3^{15}	0	3	9	17	22	21	15	8	4	1	0	0
10, 11, 12	18	0	18	3^{16}	0	2	7	15	21	21	16	10	5	2	1	0

s, number of C—C, C—O, and O—H bonds; c, number of C—C in conjugated system; n, number of single bonds to rotate $(n = s - c)$; N, number of possible conformers $(N = 3^{n-2})$.

* Same as in Table 1. † Calculated from formulas [1] and [2]. ‡ Deducted 2 single bonds, C_{11}-C_{12} and C_{13}-C_{14}, from 15 ones (s) (see explanations in Fig. 1 and Materials and Methods).

the C_{12} triple bond, the direction of the C_{14}—C_{15} axis is presumed to be at random, being due to a free rotation about the C_{13}—C_{15} axis.

(d) Procedures to build up models of conformers. Conformers of each compound were built up with a chemical model, W. Büchi Glasapparatefabrik, from a hydroxyl group on C_1 to C_{16}, C_{17} or C_{18} methyl group with the following assumptions: (i) The conformer with a zigzag anti-anti-anti- ... plane structure is defined as the original form; (ii) the rotation about each of the C—C, C—O, and O—H single bonds is independent of each other, giving restricted rotations of approximately 120° about each axis; (iii) the probabilities of three restricted rotations about each axis can be regarded as about 0.7, 0.15, and 0.15 for anti and two of gauche, respectively, using the equilibrium constant K of about 2.0; (iv) for compound 8, possessing the C_{12} triple bond (Fig. 1), there exist complete freedoms of rotation about the C_{13}—C_{14} axis; (v) the rotations about the C_{11}—C_{12} single bond involved in the conjugated systems in compounds 1, 2, 3, 4, and 6 are completely restricted; and (vi) in the cases where the interatom repulsions should be considered, the conformers are rearranged with van der Waals or kinetic theory radii so as to give a state of potential minimum.

(e) Sampling of conformers. For a large number of possible conformers involved in each compound, 100 conformers of each compound were sampled according to the following formula:

$$Nr = {}_{n-2}Cr \ 0.7^{n-2-r} \cdot 0.3^{r} \qquad [2]$$

where n is the number of the C—C, C—O, and O—H single bonds involved in each compound, r is the number of gauche forms involved in each conformer, and Nr is the number of conformers containing r gauche forms. With this formula percent values of Nr ($= Nr/N \times 100$) were estimated as listed in columns 6 to 17 of Table 2. For the conformers thus sampled, the positions of r single bonds relating to the formations of gauche forms and the directions of the rotation about each of r single bonds were computed with a random table. For compound 8, possessing the C_{12} triple bond (see Fig. 1), I also randomized the direction of the C_{14}—C_{15} axis, assuming a free rotation about the C_{13}—C_{14} axis at 10-degree intervals.

(f) Procedures to define "bifunctional units." A total of 1200 conformers (100 conformers from each of the twelve compounds) was built up and 4000 distance values between "functional regions" were measured: 1200, 1400, and 1400 for D-Me, D-A, and A-Me systems, respectively. Thus, 40 frequency distribution patterns of distances (12, 14, and 14

for D-Me, D-A, and A-Me, respectively) were compared with each other to define any significant systems with a particular distance ("bifunctional unit") that may occur, the probability of which is most closely correlated with the biological activity.

RESULTS

Frequency Distributions of Distances among a Proton Acceptor, a Proton Donor, and a Methyl Group. Fig. 2 shows frequency distributions of the distances between a proton acceptor (A) and a methyl group (Me) (Fig. 2a), between a proton acceptor (A) and a proton donor (D) (Fig. 2b), and between a proton donor (D) and a methyl group (Me) (Fig. 2c) involved in conformers of compounds 1, 6, 9, and 12.

Correlation of Biological Activity with the Probability of "Bifunctional Unit" Formation. (a) Deilephila euphorbiae. As listed in column 2 of Table 1, threshold values of bombykol and its 11 derivatives for Deilephila euphorbiae range from 10^{-1} to 10^3 μg. Comparison between distance frequency distribution patterns of the three systems involved in these compounds has indicated that the biological activities for Deilephila can be explained by the presence of a particular proton acceptor–methyl system, the distances of which were about 5–6.2 Å (bifunctional unit F-1 in Fig. 2a), correlation coefficient between the probability of the bifunctional unit F-1 formation (P_{f-1} in column 6 of Table 1), and the biological activity (column 2) being about 0.94.

(b) Aglia tau. Columns 3 and 7 of Table 1 show the threshold values for Aglia tau and the probabilities (P_{f-2}) of a particular proton acceptor (A)–proton donor (D) system with A-D distances of 10.1–11.9 Å (bifunctional unit F-2 in Fig. 2b). There exists a significant correlation between the biological activity and the probability (P_{f-2}) ($r = 0.83$).

(c) Bombyx mori. A marked difference of pheromone activity among bombykol and its 11 derivatives for the silkworm moth, Bombyx mori, as listed in column 4 of Table 1, can be explained if we assume that, whereas one of the proton donor (D)–methyl (Me) systems (D-Me distance: 13–15.7 Å, bifunctional unit F-3 in Fig. 2c) plays a favorable role in producing pheromone action, another D-Me system involved in certain conformers (D-Me distance: 3.2–6.0 Å, bifunctional unit A in Fig. 2c) reveals an adverse action against the favorable unit F-3. The probabilities for the favorable unit F-3 (P_{f-3}) and the adverse unit A (P_a) were listed in columns 8 and 9 of Table 1. The pheromone action of each compound was represented as the "net probability" P_o between two antagonistic units F-3 and A in the following formula: $P_o = P_{f-3} + kP_a$, where k is a constant for an ad-
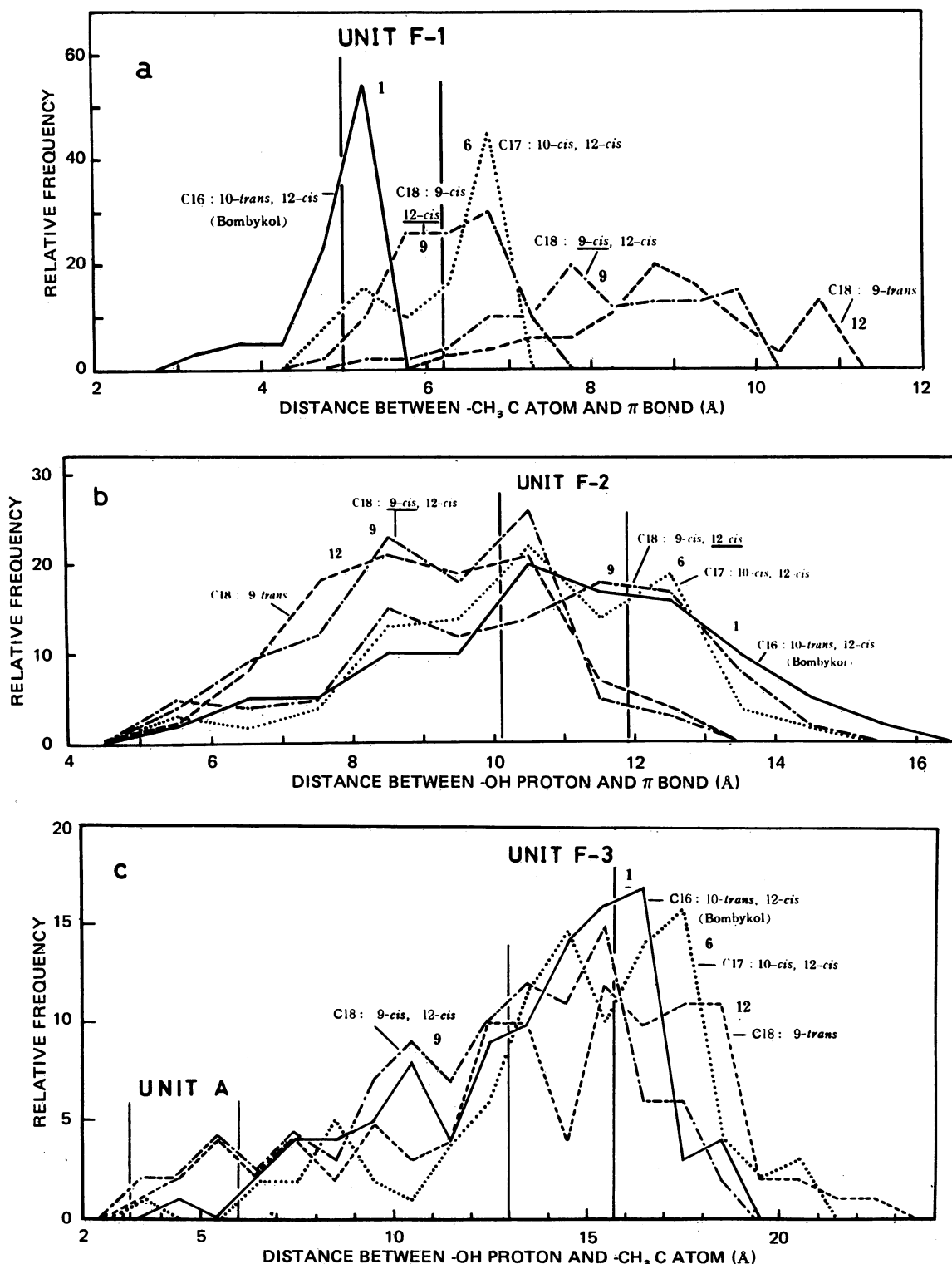
FIG. 2.   Frequency distributions of distances between a proton acceptor and a methyl (a), a proton acceptor and a proton donor (b), and a proton donor and a methyl system (c), obtained from 100 distance values involved in 100 conformers of each compound. Each numeral is the same as in column 1 of Table 1.

verse contribution of the unit A on the biological action. P_o values for the 12 compounds are listed in column 10 of Table 1 with a k value of −2.0 for *Bombyx* (the k value, −0.2, was estimated so as to give the highest correlation coefficient between biological activity and the P_o value). It is obvious from this table that the biological activity of 10 active compounds (compounds 1 to 10) has a striking corre-

lation with the calculated value (P_o), the correlation coefficient being about 0.94. The other two inactive compounds at 10^3 μg level (compounds 11 and 12) gave the smallest values among the 12 compounds examined, indicating that two antagonistic units perhaps play an important role in governing biological activities.

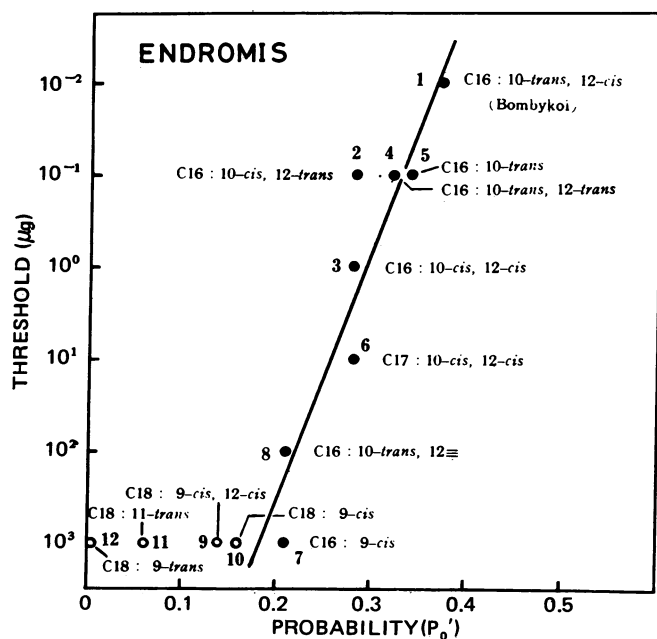(d)   *Endromis versicolora*. Biological activities of the 12

Biochemistry: Kikuchi

Proc. Nat. Acad. Sci. USA 72 (1975)   3341

FIG. 3. Logarithmic values of the threshold of 8 compounds for *Endromis versicolora* (column 5 of Table 1) plotted against net probabilities (P_o') for two antagonistic units of F-3 and A listed in the right column of Table 1. $r = 0.92$. Each numeral is the same as in Table 1. Notice four biologically inactive compounds at the level of 10^3 μg (*open circles*) gave insufficient P_o' values, as compared with those of eight active ones (*solid circles*).

compounds for *Endromis* can be explained from the probabilities of two antagonistic *D-Me* units: the favorable unit F-3 and adverse unit A, used for the silkworm moth, *Bombyx mori*. The "net probabilities" for *Endromis* (P_o') are listed in the right column of Table 1, using the same formula: $P_o' = P_{f-3} + klP_a$ as in the silkworm moth, where a k value of -3.0 is given for *Endromis* instead of -2.0 for *Bombyx* (the k value of -3.0 was obtained in the same manner as in *Bombyx*). Fig. 3 shows the logarithmic threshold values of eight active compounds for *Endromis*, compounds 1 to 8, plotted against the net probabilities (P_o') listed in Table 1. It is apparent from this figure that biological activities of these compounds are strikingly correlated with the calculated values (P_o') $(r = 0.92)$, and that all of the four inactive compounds at 10^3 μg level (compounds 9, 10, 11, and 12 in Table 1) gave insufficient P_o' values (0.16 to 0.00), as compared with those of active ones.

DISCUSSION

Although the excitation of sex pheromone receptors of lepidopterous insects examined may involve complex prossesses, highly significant correlations between biological activity and probability indicate that the "bifunctional units" involved in conformers play a key role in producing pheromone actions. If this is so, the role of the unit for *Aglia* (unit F-1) or *Deilephila* (unit F-2) might be explained if it is ascertained that there exists a receptor site $(f-1$ or $f-2)$ on the cell membrane which interacts with the corresponding unit involved in odor molecules. For *Bombyx* and *Endromis*, bio-

logical activities of odor molecules might be explained if it is ascertained that there indeed exist two receptor sites $(f-3$ and $a)$ which interact specifically with the corresponding unit F-3 and A, and that the excitation of a receptor cell can be evoked by interactions of unit F-3 with site $f-3$, whereas those of unit A with site a result in conformational changes in the receptor cell membrane, whereupon site $f-3$ might be modified or eliminated. Besides these assumptions, the marked differences of biological activities among bombykol and its derivatives might be explained if the excitation of a receptor cell is initiated by the simultaneous interaction of molecules with receptor sites. From the gradient of regression line shown in Fig. 3 and those derived from Table 1, the minimal number of pheromone molecules to evoke the excitation on a receptor cell can be regarded as about 17 molecules for *Endromis* and about 20, 15, and 4 for *Bombyx*, *Aglia*, and *Deilephila*, respectively. The minimal number of tritiated bombykol molecules to elicit the excitation of a receptor cell of the silkworm moth has been estimated to be approximately seven molecules per sec of stimulus duration (8). Thus, about three conformers involved in these molecules are presumed to be biologically active (7 molecules × 0.4 for P_{f-3} in Table 1). Such estimations may involve a certain degree of error, so that these different estimations derived from entirely different approaches may not be necessarily incompatible with each other at the present stage. For further confirmation of the role of the bifunctional units, it is important to see whether or not the biological activities of additional compounds can be predicted from calculated values. The threshold value of compound 8 for *Deilephila*, which has not been examined, is presumed to be 10^{-2} μg or so from a comparison of an average probability (P_{f-1}) of about 0.12 for unit F-1 with the biological activities shown in column 6 of Table 1 $(P_{f-1} = 0.23$ and 0.00 for the C_{12} triple and C_{10} double bond, respectively). In addition to this example, a mixture of bombykol with either *n*-propanol or *n*-pentanol, which presumably possesses only an adverse unit A, as shown in Fig. 2c, may result in masking or blocking the bombykol action both for *Bombyx* and *Endromis*. Information on biological actions of those predicted compounds should be of help in analyzing complex processes involved in pheromone perception in insects.

1. Amoore, J. E., Palmieri, G., Wanke, E. & Blum, M. S. (1969) *Science* 165, 1266-1269.
2. Tai, A., Matsumura, F. & Coppel, H. C. (1971) *J. Insect Physiol.* 17, 181-188.
3. Wright, R. H. & Brand, J. M. (1972) *Nature* 239, 255-256.
4. Kikuchi, T. (1973a) *Nature* 243, 36-38.
5. Kikuchi, T. (1973b) *Jap. J. Genet.* 48, 105-118.
6. Priesner, E. (1969) in *Olfaction and Taste*, ed. Pfaffmann, C. (Rockefeller Univ. Press, New York), Vol. 3, pp. 235-240.
7. Eliel, E. L., Allinger, N. L., Angyal, S. J. & Morrison, G. A. (1965) *Conformational Analysis* (John Wiley and Son, New York), p. 11.
8. Kaissling, K.-E. & Priesner, E. (1970) *Naturwissenschaften* 57, 23-28.