Molecular Surface Analysis: A Computer Assisted Search for Structure-Odor Relationship

By Robert S. Marmor, Lorillard Research Center, Greensboro, NC.

The techniques used in isolation and identification of the components of flavor and fragrance materials have so improved over the years that we can produce almost any desired reconstituted facsimile, often of excellent quality. And yet we know very little about what is going on at the molecular level, of how molecules interact with olfactory receptors, or what governs the specificity and intensity of the response.

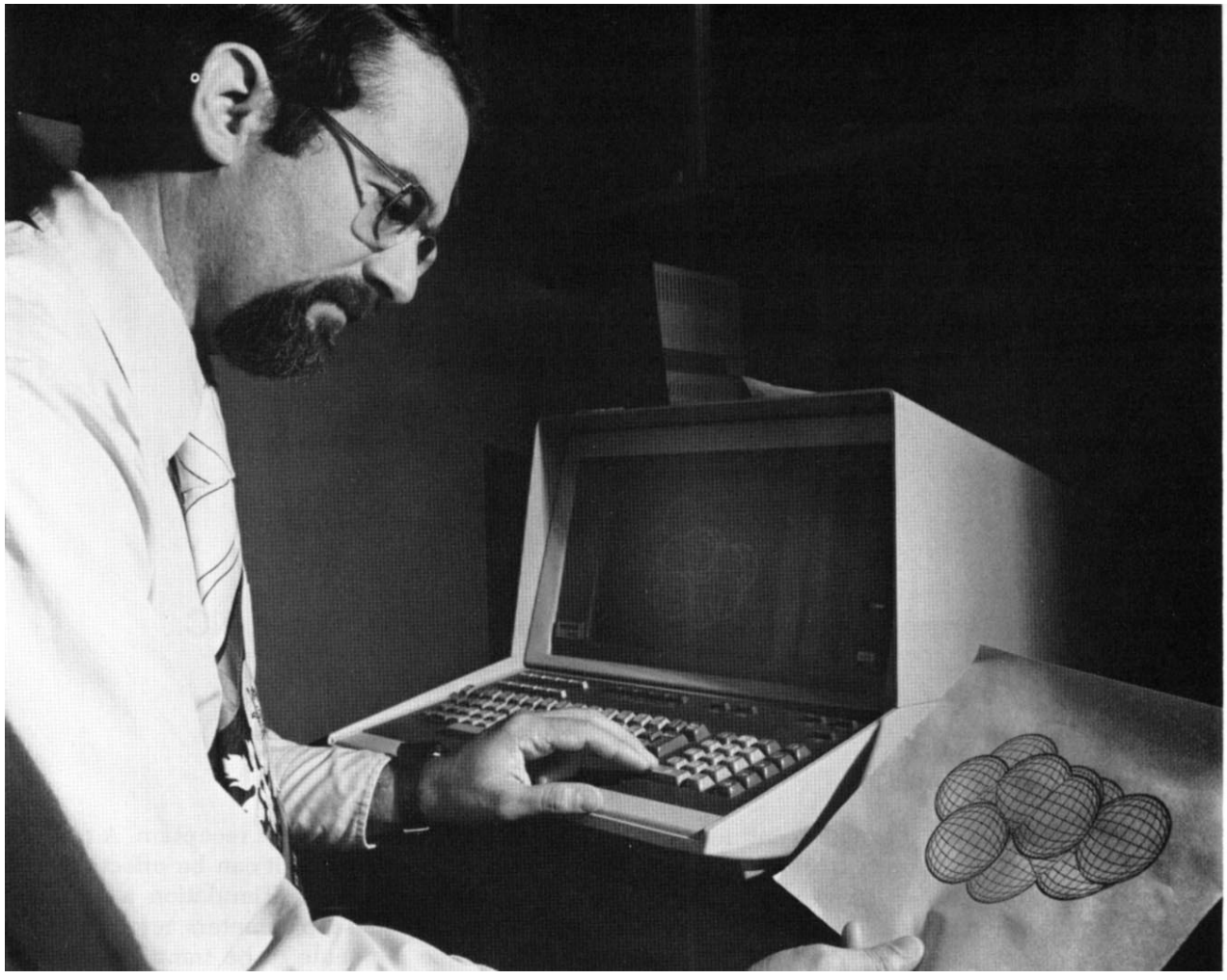
If we were to gain greater understanding of the basic mechanism of olfaction it would be a tremendous aid in our efforts to develop new and improved flavors and fragrances. There are two main reasons why this is so. First, we would then be able to quantitatively correlate olfactory response with physicochemical parameters, i.e., we could predict the type of odor or flavor and its strength from measurable data. This would allow mathematical description of olfactory nuances and permit optimization of recipes. Secondly, we would then be able to design new molecules with predicted organoleptic properties. Such a rational approach would likely result in a higher success rate in the synthesis of new and useful flavors and fragrances. Consider that at present the usual operation is guided by the philosophy "let's make and evaluate every synthetic analog we can think of."

The most prevalent olfactory theory invokes the "key in lock" model of reception. A receptor site is stimulated when it can be effectively occupied by a molecule. Stimulation would then correlate with those parameters relating to the ability of the molecule to be transported to the receptor site, to its chemical affinity for the site, and to its physical dimensions, or shape, which would regulate its ability to enter the site. Since physiological experiments have been unable to show high specificity of response in olfactory receptors, the interaction mechanism is probably more complex, perhaps involving the fit of a molecule into several different receptors.

While most investigators agree that molecular shape is likely to be an important parameter in structure-odor relationships, there have been few studies in which the shapes of molecules have been adequately represented or evaluated in a mathematically rigorous fashion. This is in part due to the complex molecular shape, and it is not meaningfully represented by a single number as would be, for example, the partition coefficient. Furthermore chemists are accustomed to constructing hand held molecular models and making predictions based on visual pattern recognition. We believe there are better ways of doing this.

The purpose of this article is threefold. First, we shall explain how we have adapted the small computer to carry out comparisons between the

 0272-2666/83/0004-3701$04.00/00



Entering molecular data into the computer.

surface regions of molecules and to identify regions of significant spatial similarity. Secondly, we shall demonstrate how this method was applied to the development of new synthetic cedarwood tobacco flavorants. We will show how we were able to predict the shape of the putative receptor site, but we will not be discussing specific results of our synthesis program. The utility of this approach is currently being evaluated in our laboratories and will be the subject of future publications. Finally we would like to point out that this approach could be implemented in most R&D settings at minimal training and equipment costs.

Molecular Surface Analysis

The first step is to enter the data into the computer. Programs have been written by us and others which convert basic information about the molecules into stored structural information. Required data include which atoms are bonded, what are the ideal bond lengths and angles (as entered from tables) and what are the Newman angles (estimated from protractor measurements on hand held molecular models). This information is transformed into a ball and stick model visible on the computer screen *(see photo).*

Another program, written by us but also available from others, performs molecular mechanics structural optimization using force field techniques. This procedure moves the atoms around to their minimal energy positions, converting the original approximate entry into one which is usually quite close to the actual dimensions of the molecule.

The molecular surface analysis procedure does require that the molecules be structurally rigid. If they are not, you must select the particular conformation you wish to examine. This

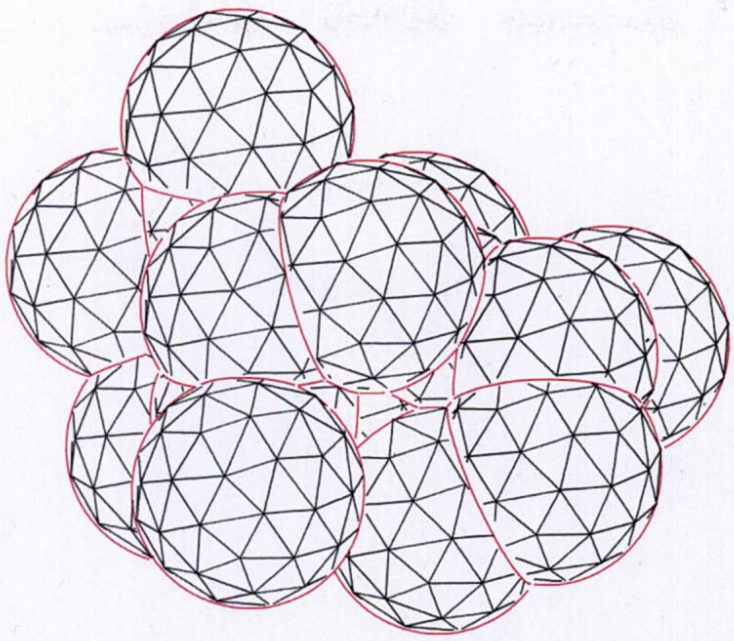


Figure 1. Equidistant dot surface model. Mathematical representation of the surface of Cedrol.

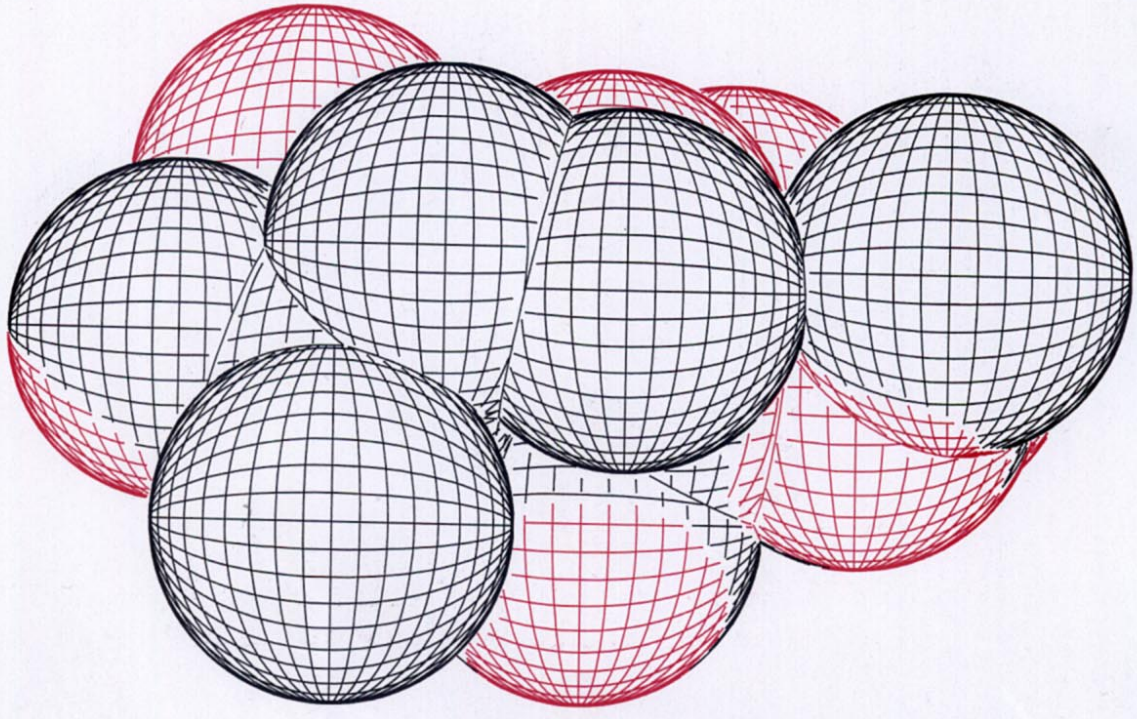


Figure 2. Portion of surface of Cedrol shown in black is coincident with that of Sclareol Oxide.

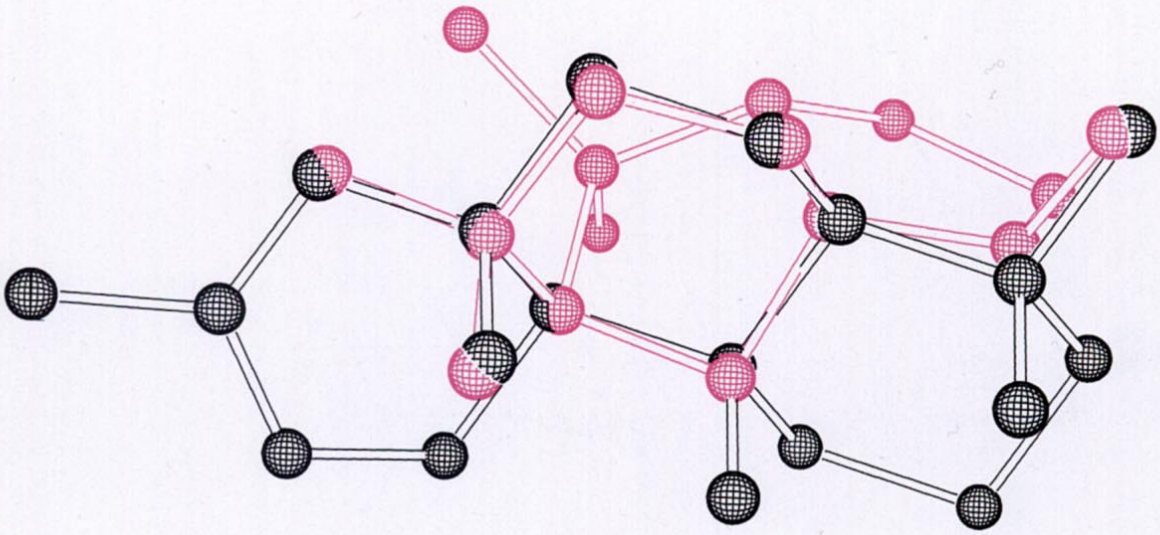


Figure 3. Cedrol (red) vs. Sclareol Oxide (black).

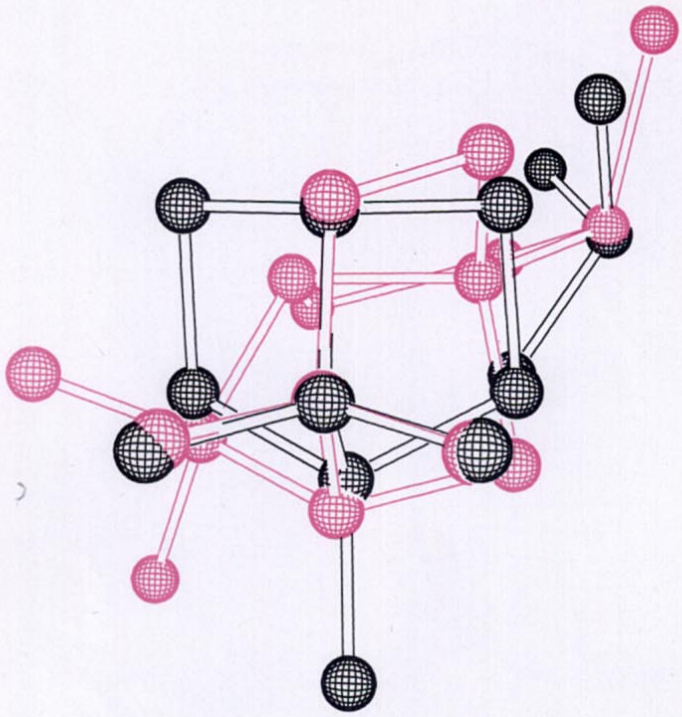


Figure 4. Cedrol (red vs. Isopropoxycamphane (black).

requirement of structural rigidity is not necessarily as much of a limitation as one would think, since it has often been the case in pharmaceutical development that valid structure-activity relationship predictions have been made from floppy drug molecules when they are examined in their minimum energy conformations. The computer then calculates the location of the surface of the molecule from the Van der Waals radii of the atoms, and expresses this mathematically as the Cartesian coordinates of hundreds of dots spread uniformly across the molecular surface (figure 1).

When molecule A is compared by the computer to molecule B, all the ways in which molecule A can be rotated and translated relative to B are examined. All regions are found where there exists significant similarity of shape of the molecular surfaces. For example, the computer might determine that the best fit was found when molecule A was rotated 23°, 47° and 182° and then translated 1.2, −2.1 and 0.6 Angstroms about the x,y and z axes respectively, relative to a stationary molecule B. It might report that at this configuration, 42% of the surface of molecule A was found to be coincident with that of molecule B. Also printed would be the second best fit, the third best, and so on. The computer would then plot pictures of these selected configurations.

If a series of compounds all possessing a similar flavor or fragrance note are examined by this molecular surface analysis procedure, it will become evident if certain regions are repeatedly involved in favorable surface match-ups. Such was the case in our examination of cedarwood tobacco flavorants.

Cedarwood Tobacco Flavorants

Eight molecules were selected for comparison, all reported to possess the odor or taste of cedarwood or a related aromatic woody oriental tobacco note:

Cedrol
Sclareol Oxide
6-Hydroxythiaspirane
6-Acetoxythiaspirane
Isopropoxycamphane
trans-1,7,7,-Trimethylbicyclo[4.4.0]decan-3-one
Caryophyllene
β-4,8,13-Duvatriene-1,3-diol

Cedrol is a rigid molecule, flexible only to the extent of allowing boat and chair conformations of its six-membered ring. The chair form was used for molecular surface analysis. It was compared with each of the molecules listed above.

Cedrol vs. Sclareol Oxide

Sclareol oxide is a rigid molecule capable only of boat-chair conformational changes. The all-chair conformation was used for molecular surface analysis with cedrol. The best fit resulted in a configuration in which 49.6% of the surface area of cedrol was coincident with that of sclareol oxide (fig. 2). In this configuration a ridge of eight atoms along the edge of the cedrol molecule was in close proximity to atoms of sclareol oxide (fig. 3). Interestingly, the oxygen atom of cedrol was one of the eight atoms and had been paired with the oxygen atom of sclareol oxide.

Cedrol vs. 6-Hydroxythiaspirane

6-Hydroxythiaspirane is also a fairly rigid molecule except for boat-chair conformational possibilities. Molecular surface analysis on the chair form resulted in a best fit configuration in which 49.1% of the surface area of cedrol was coincident with that of 6-hydroxythiaspirane. In this configuration all but one atom found in the "cedrol ridge" identified in the cedrol-sclareol oxide comparison was again seen to be in close proximity to atoms in 6-hydroxythiaspirane. Moreover cedrol's oxygen atom was paired with one of the oxygen atoms of 6-hydroxythiaspirane.

In the second best fit accounting for 47.7% of cedrol's surface area, this same ridge is in modest alignment with atoms of 6-hydroxythiaspirane including the one cedrol ridge atom not involved in the first fit. However in this second best fit cedrol's oxygen atom is not paired with another oxygen atom.

Cedrol vs. 6-Acetoxythiaspirane

6-Acetoxythiaspirane has boat-chair ring conformational possibilities as well as an infinite number of rotamers for the acetate group. A rotamer was chosen which places the acetate group away from the five-membered ring. After molecular surface analysis the best fit of this rotamer with cedrol accounted for 46.4% of the surface area of cedrol and involved the methylcarbinol group of cedrol and the acetate group of 6-acetoxythiaspirane. This fit was entirely dependent on the position of the acetate group.

The second best fit accounted for 46.2% of the surface area of cedrol. Here the same "cedrol ridge" seen previously was in close proximity to atoms of 6-acetoxythiaspirane. In this fit the rotational positions of the acetate group was unimportant.

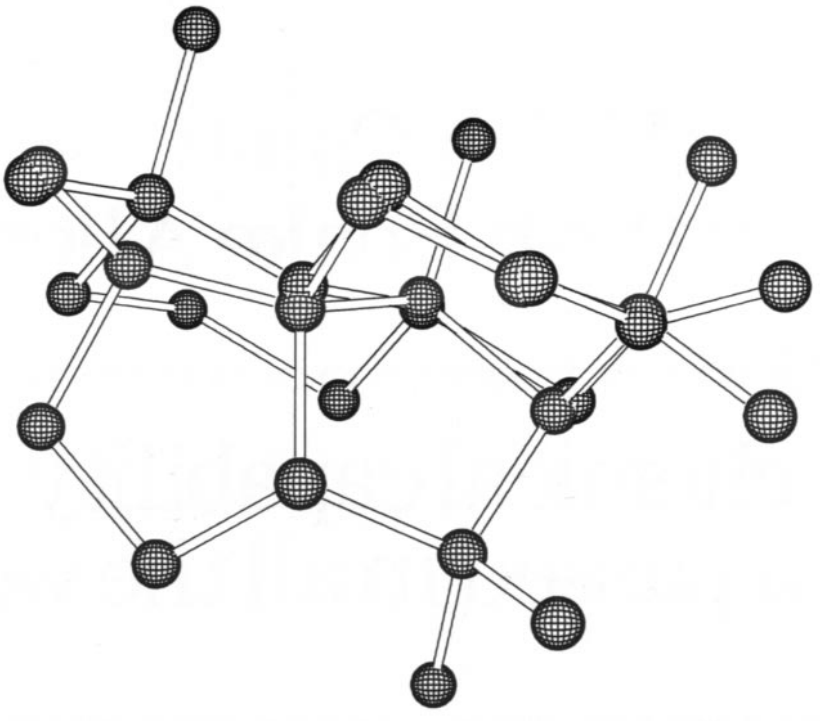


Figure 5. Cedrol (red) vs. trans-1,7,7-Trimethylbicyclo[4.4.0]decan-3-one.

Cedrol vs. Isopropoxycamphane

Isopropoxycamphane, also known as bornyl isopropyl ether, has an inflexible ring system and a free rotating isopropyl group. This group was held in a minimum energy conformation of zero Newman angle between the bornyl carbon and the lone hydrogen on the center carbon of the isopropyl group. Molecular surface analysis with cedrol revealed in the best fit that 46.8% of the surface area of the smaller isopropoxycamphane molecule was coincident with the surface area of cedrol. This fit involved the methylcarbinol group of cedrol.

There was little interest in this fit or the others below it until we examined the eighth fit, involving 40.8% of the surface area of isopropoxycamphane. In the eighth fit, most of the "cedrol ridge" was identified in the coincident area region. Cedrol's oxygen was not paired with isopropoxycamphane's oxygen. It was interesting to study the 3-D graphs in this case because they made clear that the isopropyl group was capable of being rotated into a more favorable position to overlap with the cedrol ridge.

A second run was made with the isopropoxycamphane rotamer suggested by the previous results. In this case, the best fit involved 52.6% of the surface area of isopropoxycamphane and was clearly a good matchup of the cedrol ridge, but did not pair cedrol's oxygen with another oxygen (fig. 4).

Cedrol vs. trans-1,7,7-Trimethylbicyclo [4.4.0]decan-3-one

Except for the possibility of boat and chair conformations, trans-1,7,7-trimethylbicyclo [4.4.0]decan-3-one is a completely rigid molecule. Molecular surface analysis of the chair form with cedrol gave in the best fit a 54.2% coincident area of the ketone. The methylcarbinol region of cedrol was involved. In the third and sixth best fits (49.3% and 44.6% coincident

area respectively) more of the "cedrol ridge" region was involved. The seventh best fit (44.1% coincident area) was a very good fit of the entire cedrol ridge. The ketonic oxygen in this configuration was seen to be positioned between the methyl and hydroxyl groups of cedrol (fig. 5).

Molecular surface analysis of the boat form of the ketone gave similar results but not as striking. The twelfth best case represented 44.1% coincidence of the surface area of the ketone in the boat form and gave a modest fit of the cedrol ridge region.

Cedrol vs. Caryophyllene

The caryophyllene molecule is very flexible and capable of many low energy rotamers. One configuration chosen at random was analyzed. The best fit involved pairings of the gem-dimethyl groups of the two molecules. None of the top twelve fits were reminiscent of previous alignments.

A second conformation was estimated based on a forced fit with the previously identified "cedrol ridge. This conformation turned out to be relatively unhindered. Molecular surface analysis on this conformation gave a best fit with 50.3% coincident surface area of caryophyllene but an uninteresting configuration. The fifth best fit with 42.9% coincident surface area gave a good fit of the cedrol ridge region and a fair proximity of cedrol's oxygen to an electron rich double bond of caryophyllene.

Cedrol vs. β-4,8,13-Duvatriene-1,3-diol

The duvatrienediol molecule is extremely floppy. It is not possible to predict regions from this molecule because there is no reason to prefer one conformation over another. However, given the previous observations of a "cedrol ridge," it is possible as in the carophyllene case to test the results so far.

A rotamer model was chosen in which the methylcarbinol groups of cedrol and the duvatrienediol were superimposed, and the rest of the molecule visually aligned so as to be in close proximity to the cedrol ridge atoms. Having "stacked the deck," it was comforting to find that molecular surface analysis could give as the best fit a very good alignment with the cedrol ridge region, where 56.3% of the surface area of cedrol was coincident with that of the surface area of the duvatrienediol.

Discussion of Results

A correlative trend has been clearly demonstrated between the shape of a sizeable region of the surface of the cedrol molecule and that of a number of cedarwood flavorants of diverse molecular structure. There is little likelihood one would have foreseen this relationship and certainly no resonable way one could have quantified the relationship without the help of the computer.

This cedrol region may relate to the odor and taste of cedarwood. If this is in fact true, then the dimensions of the putative olfactory receptor are likely to be those of a concave surface fitted by this region. We state this cautiously for several reasons. This observed relationship may be due to interaction with something other than an olfactory receptor. From a practical point of view it does not matter so long as the observed relationship results in accurate prediction, but from a theoretical point of view we cannot rule out the possibility of interaction with a mediator.

Another possibility is that we have discovered only one of the necessary interactions for perceptions of the cedarwood note. There may be other receptors that require activation. And finally, we have only assumed that it is the cedarwood quality that is associated with a certain molecular shape. There is no reason to believe that these eight molecules we have examined do not have other physiological properties in common, any one of which may be associated with the observed spatial relationship.

In spite of these disclaimers, the molecular surface analysis procedure is useful in both applied and theoretical application. In our search for new cedarwood flavorants we now have a rationale for proceeding in an exploratory synthesis program.

We are able to adapt this procedure to other problems. Our technical staff was eager to gain hands-on experience with the computer, and quickly learned to write useful research programs. We are interested in determining quantitative structure-activity relationships using the "percent of coincident surface" for a molecule versus a proposed receptor site as a steric scalar parameter. Also of interest are qualitative classification or discrimination techniques which might be generally useful for olfactory prediction. Although tobacco flavor development is our main interest, molecular surface analysis could be useful in other fields where structure-activity relationships are studied.

Acknowledgement

This article is based on papers presented at the 35th Tobacco Chemists' Research Conference, Winston-Salem, North Carolina, October 1981 and the 184th American Chemical Society National Meeting, Kansas City, Missouri, September 1982.