Classification of Organic Reactions: Similarity of Reactions Based on Changes in the Electronic Features of Oxygen Atoms at the Reaction Sites[1]

Hiroko Satoh,*,† Oliver Sacher,‡ Tadashi Nakata,† Lingran Chen,‡,⊥ Johann Gasteiger,*,‡ and Kimito Funatsu*,§

Synthetic Organic Chemistry Laboratory, The Institute of Physical and Chemical Research (RIKEN),
2-1 Hirosawa, Wako, Saitama 351-01, Japan, Computer-Chemie-Centrum, Institute of Organic Chemistry,
University of Erlangen-Nürnberg, Nägelsbachstrasse 25, D-91052 Erlangen, Germany, and Department of
Knowledge-Based Information Engineering, Toyohashi University of Technology,
1-1 Tempaku, Toyohashi, Aichi 441, Japan

Received December 17, 1997

Organic reactions occur as a result of complicated interactions among many factors: structural and electronic features of reactants, reagents, catalysts, temperature, etc. In this study, organic reactions were automatically classified based on these factors. A dataset of 131 reactions was investigated focusing on the changes of electronic features on the oxygen atoms at the reaction sites by principal component analysis and self-organizing neural networks analyses. Good correlations were found between the similarities in the changes of the electronic features of oxygen atoms of the reaction sites and the similarities in the substructural transformations at the reaction sites as well as with the known reaction types. These results demonstrate that a classification based on changes of electronic features is closely related to the classifications which chemists have been establishing from various points of view. Furthermore, this indicates the possibility for the automatic and systematic classification of a large number of organic reactions.

1. INTRODUCTION

Synthetic chemists have many problems to solve: what starting materials and reaction conditions could give a molecule having a desired structure, will the desired reaction actually occur, will the molecule be produced as a major product, will side reactions occur, what is the reaction mechanism, etc.

These problems are difficult to solve because there are many factors to be considered and the degree of contributions and interactions varies. High-dimensional complexity which is inherent in reactions makes it difficult to solve these problems.

Experimental chemists are aware of the complexity because they actually handle reactions. In most cases they solve these problems using something like a *reaction map* in their brain. Such a reaction map organizes reactions on the basis of chemical knowledge they have learned from textbooks and their experiments. So each chemist has his/her own reaction map.

The reaction map holds a key to the solution. What is the contents of the reaction map? How is the reaction map constructed? Chemists have observed reactions, analyzed reactions in detail, accumulated several data obtained in the process of analysis, organized them systematically, and established many theories and rules. In a word, chemists have learned from experimental facts of reactions.

The number of observed reactions is extremely large and is increasing day by day. Part of them is published in articles and/or stored in databases on a computer. The number of the published and stored reactions goes into the millions.

A reaction map in a brain of a chemist is constructed from only a part of these reactions, because anyone cannot see all of these reactions.

How can this process of organizing the reactions be handled on a computer? The strong points of a computer are the even, speedy, and accurate treating of a large amount of data. This offers the possibility to produce a reaction map from a huge amount of computer readable reactions.

The following steps have been made to construct a reaction map utilizing the strong points of a computer. The first step is the classification of reactions based on factors controlling the course of reactions. The classifications will be performed focusing on several similarities. The next step is the systematic organization of the results of the classifications to construct a reaction map on a computer. The constructed reaction map will be employed for solving problems in computer-assisted reaction prediction and synthesis design, and this will be discussed in future articles.

It was shown previously that reaction classification can be performed in an unsupervised learning process.[2,3] Reactions were analyzed based on electronic features at the reaction sites of the reactants or of the products by a self-organizing neural network (Kohonen network). These studies focused on reactions having the same atoms and bonds participating in the bond rearrangement scheme. It was shown that the combination of physicochemical variables and the self-organizing features of a Kohonen network are a

* Corresponding author.
† The Institute of Physical and Chemical Research (RIKEN).
‡ University of Erlangen-Nürnberg.
§ Toyohashi University of Technology.
⊥ Present address: MDL Information Systems Inc., 14600 Catalina Street, San Leandro, CA 94577.

S0095-2338(97)00119-4 CCC: $15.00
Published on Web 02/28/1998

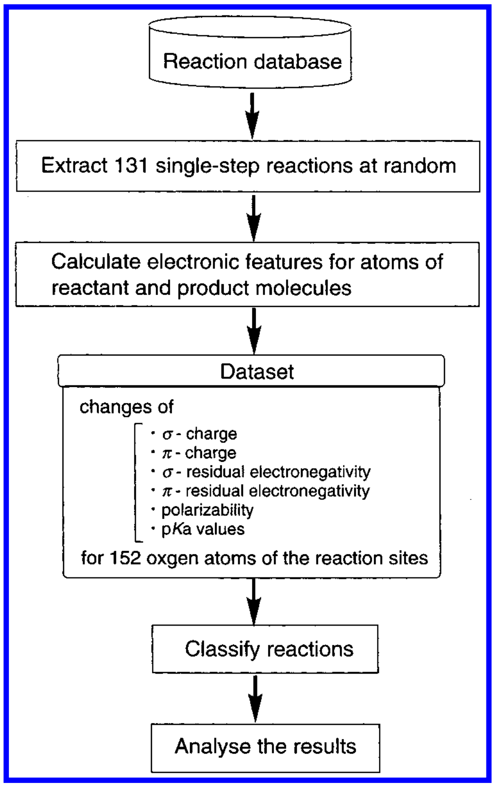

Reaction database
Extract 131 single-step reactions at random
Calculate electronic features for atoms of reactant and product molecules
Dataset
changes of
· σ - charge
· π - charge
· σ - residual electronegativity
· π - residual electronegativity
· polarizability
· pKa values
for 152 oxgen atoms of the reaction sites
Classify reactions
Analyse the results

Figure 1. Procedures for classification of reactions and its analysis.

powerful means for the identification of reaction types and their scope.

In this study, we have chosen reactions having oxygen atoms at the reaction site *at random* from a reaction database and classified the reactions based on *differences* of electronic features of the corresponding oxygen atoms of reaction sites between reactants and products by the same neural network technique. This work shows the possibility to automatically produce a classification of *arbitrary* reactions based on electronic features of reaction sites by the above mentioned combination of methods for the calculation of electronic features and a Kohonen neural network.

2. METHODS

2.1. Procedures for Classification and Analysis. Figure 1 shows an outline of the procedure for the classification of reactions and the analysis of the results.

First, 131 single-step reactions were extracted at random from a reaction database which is one of the modules of a computer program systems for synthesis design AIPHOS (Artificial Intelligence for Planning and Handling Organic Synthesis).[4] Resources of it are SYNLIB,[5] and reactions are directly collected from the literature. Part of the reactions are shown in Figure 2. All of the 131 reactions were checked and corrected by going back to the original references.

Then, electronic features for the atoms of reactant and product molecules were calculated by empirical methods.[6−8] The changes in six of these physicochemical parameters—σ-charge, π-charge, σ-residual electronegativity, π-residual electronegativity, polarizability, and pK_a values[9]—at the 152 oxygen atoms of reaction sites in going from the reactants to the products were taken as a characterization of the individual reactions and used for their classification (Figure 3).

The reactions of this dataset were classified by principal component analysis and Kohonen neural network methods. These methods for classification are explained in more detail in section 2.2.

The results of classification were then analyzed for similarities in the changes of electronic features at the oxygen atoms of the reaction sites, for similarities in the transformations of substructures at the reaction sites, and compared with reaction types intellectually assigned by chemists.

2.2. Methods for Classification. Both principal component analysis and the self-organizing neural network introduced by Kohonen[10] were used for classification. Both methods make it easy for the human eye to view and understand the distribution of data in multidimensional spaces by reducing the number of dimensions while preserving information on the distributions of the data in the initial space.

2.2.1. Principal Component Analysis. A principal component analysis (PCA) is a pattern recognition method used for analyzing the distribution of data in multidimensional spaces. PCA rotates the coordinate axes such that the dispersion (the variance) of data is largest in the first few axes. The new axes are generated by linear transformation of the initial axes.

2.2.2. Kohonen Network. The Kohonen network is a self-organizing neural network method introduced by Teuvo Kohonen for the generation of feature maps.[10] This method projects data from multidimensional spaces into a planar map preserving the topological relationships among the data in the initial space.

The Kohonen network is trained and data are projected into the neuron as follows (Figure 4): In the planar map each neuron's weight vector has the same dimension as that of the input data. Initial weight vectors in the planar map are random numbers. When one of the input data is entered into the network, a neuron which has the most similar weight vector with the input data is found, and weight vectors of neurons around the found neuron are corrected. The similarity is measured by values corresponding to Euclidean distance between the vectors:

$$d_{ij} = \sum_{i=1}^{m} (x_{is} - w_{ij})^2$$

here, m is the dimension of the input data, s is the number of input data, and j is the number of weight vectors. These processes are performed for every input data. In this manner, the planar map is trained to simulate the distribution of the m-dimensional input data. Finally, in the trained planar map, every input data is projected into the neuron which has the most similar weight vector. As a result, similar input data gather in the same or neighboring neurons in the planar map.

2.2.3. Comparison between PCA and Kohonen Network. The strong points of a PCA as compared with a Kohonen network are as follows: the recognition of boundaries of clusters usually is easy, and the analysis of the multidimensional space from various directions is possible. The weak points of a PCA as compared with a Kohonen network are that the analysis of the results in the plots of the various components can be complicated and difficult.

On the other hand, the strong point of a Kohonen network compared with the strong point of a PCA is that only one

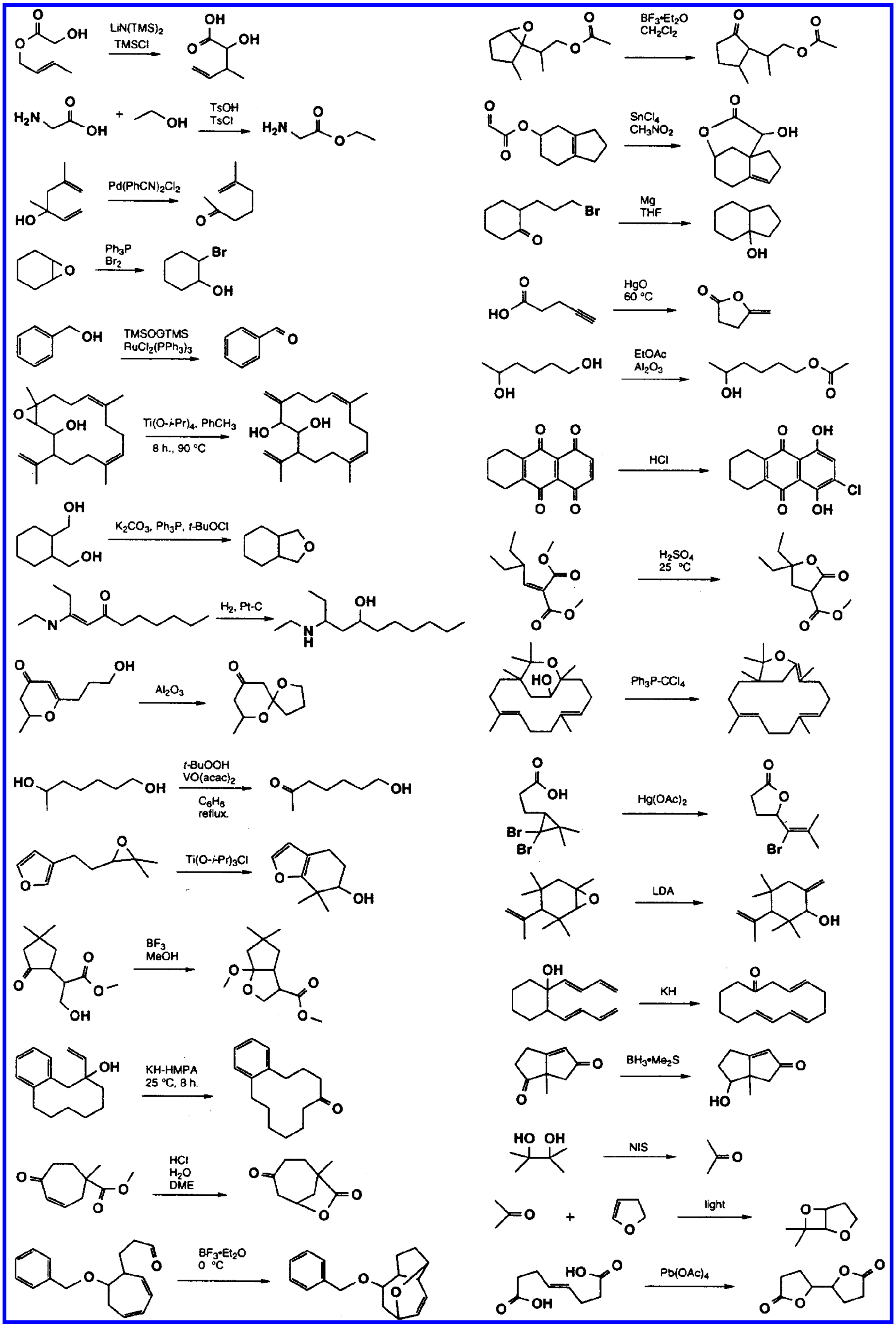


Figure 2. Part of the reactions used for the classification and analysis.

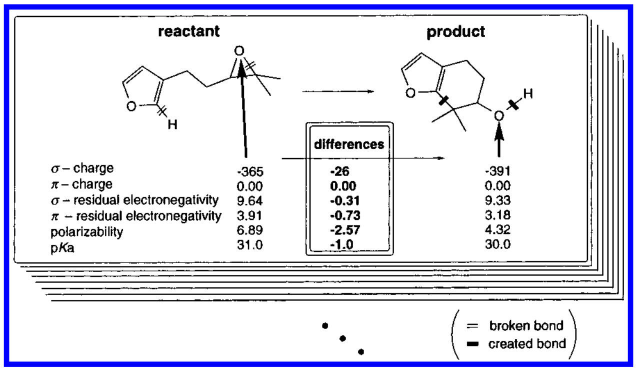


Figure 3. A data set for the analysis. The changes in σ-charge, π-charge, σ-residual electronegativity, π-residual electronegativity, polarizability, and pK_a values at the 154 oxygen atoms of the reaction sites in going from the reactants to the products are taken as a characterization of the individual reactions and are used for their classifications. In this example, differences in these values between an oxygen atom in the epoxide of the reactant to an oxygen atom in the hydroxy group of the product are calculated.

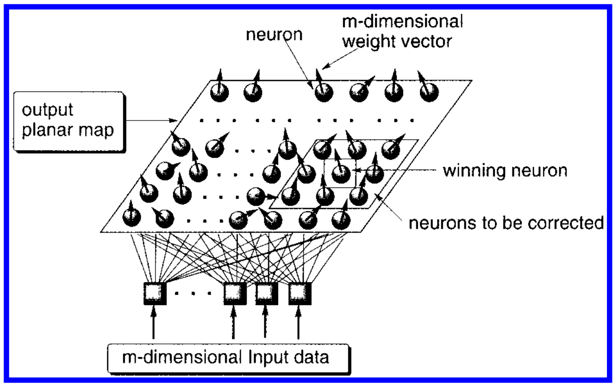


Figure 4. Kohonen network. Those neurons around the winning neuron whose weight vector is most similar to the entered input data are corrected. The processes are performed for every input data to train the Kohonen network. After the training, in the planar map, every input data is projected into the neuron which has the most similar weight vector.

planar map is output in one study. This makes the understanding of a data set easy.

It is generally assumed that the recognition of boundaries of clusters in a Kohonen network is difficult. However, it was recently shown that a careful analysis of the differences in the weight vectors in a Kohonen network allows a clear separation of clusters.[3]

The major difference between the methodology of a PCA and a Kohonen network is that a Kohonen network enforces the reduction of multidimensional space to two dimensions, whereas a PCA only performs a rotation of the coordinate axes.

Detailed explanations on PCA and a Kohonen networks are given in ref 11.

2.3. Analyses of the Results of Classification. Considering these strong and weak points of a PCA and a Kohonen network, the results of the classifications were analyzed in the following manner.

First, it was examined whether a Kohonen network successfully projected data that are close in the multidimensional space onto the same or close neurons. Namely, the results from mapping by a Kohonen network were compared with the results from a PCA. It was examined whether data belonging to the same cluster in the PCA gather on the same or close neurons on a two-dimensional map resulting from the Kohonen network.

Reactions in these studies by a PCA and a Kohonen network are considered similar on the basis of changes in the electronic features of oxygen atoms in the reaction sites. As an alternative, reactions in the planar map resulting from a Kohonen network were analyzed on the basis of the changes in substructures occurring during reactions. Furthermore, reaction types were assigned intellectually and compared with the clusters found in the PCA.

3. RESULTS AND DISCUSSION

3.1. Result from Kohonen Network. The Kohonen network program used in this study was obtained by modifying the SOM_PAK program (version 3.1) developed by Teuvo Kohonen et al.[12]

A Kohonen network was trained with properties of 152 oxygen atoms, which are in the reaction sites of the 131 reactions; each oxygen atom was represented by six physicochemical parameters, σ-charge, π-charge, σ-residual electronegativity, π-residual electronegativity, polarizability, and pK_a values, thus forming points in a six-dimensional space. In the first stage of training, the number of steps was 1000, the initial learning rate parameter was 0.05, and the initial radius of the training area was 10. In the second stage, the number of steps was 10 000, the initial learning rate

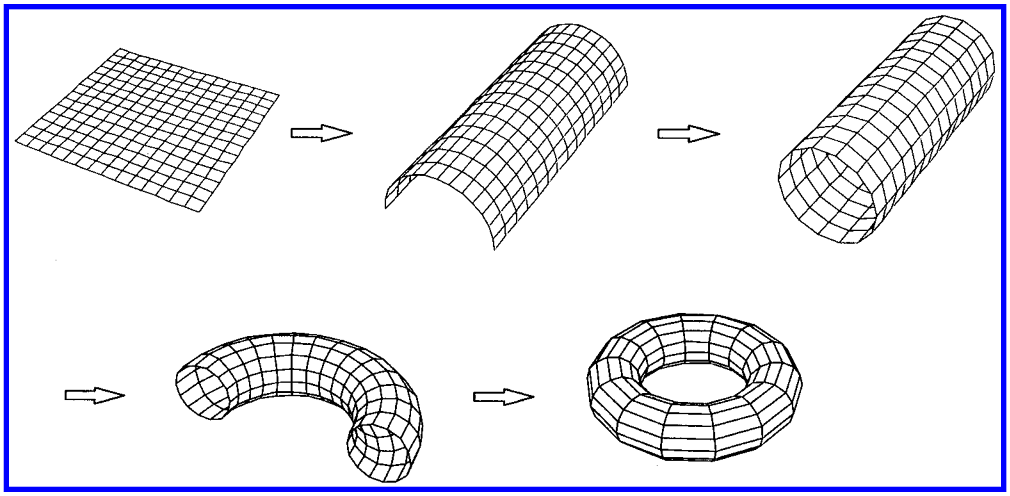
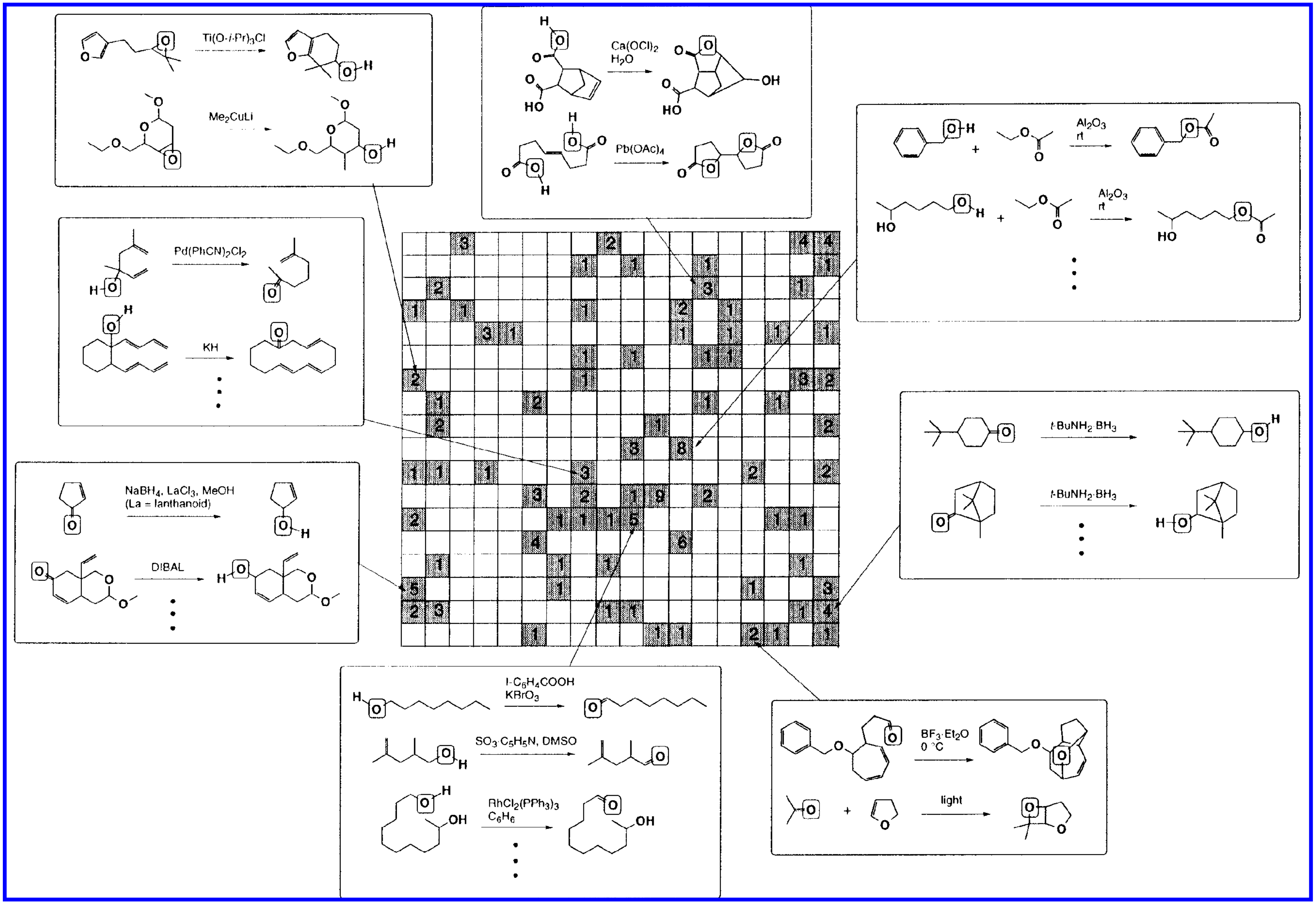


Figure 5. Toroidal map. A plane without beginning and end.



		3						2								4	4
							1		1			1					1
	2											3				1	
1		1					1				2		1				
			3	1							1		1		1		1
							1		1			1	1				
2							1									3	2
	1				2							1			1		
	2									1							2
									3		8						
1	1		1				3							2			2
					3		2		1	9		2					
2						1	1	1	5						1	1	
					4						6						
	1					1		1								1	
5						1								1			3
2	3							1	1							1	4
					1					1	1			2	1		1

Figure 6. Resulting map from the Kohonen network. Each of the squares on the map is a neuron. Neurons colored gray have fired by the 152 input data, i.e., these neurons obtained reactions. Some of the reactions are shown around the map. Oxygen atoms characterizing these reactions are encircled. The number in each neuron indicates the number of reactions which gathered on the corresponding neuron.

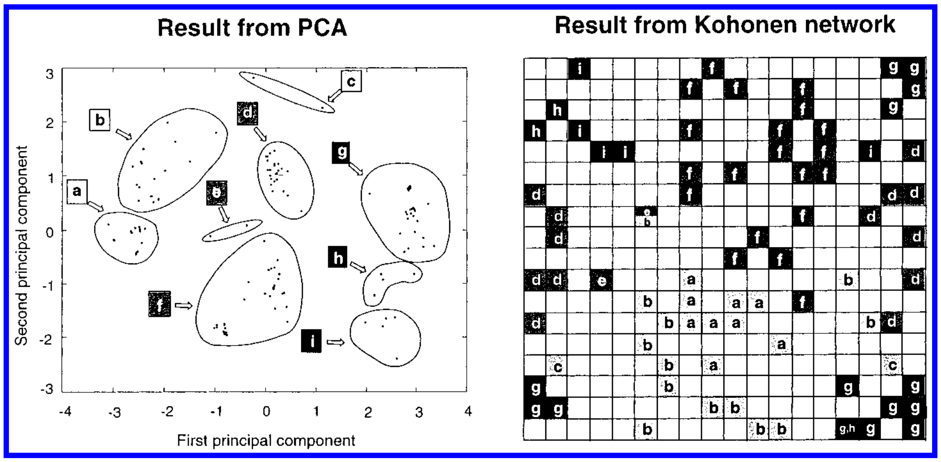


Figure 7. Comparison of the Kohonen map with the result from a PCA. The left side shows the results from a PCA in a plot of the first two components; each point in the plot represents a reaction. The right side shows the planar map resulting from the Kohonen neural network; squares belonging to the same cluster from the result of the PCA (cluster a−i) have been given the same notation.

parameter was 0.02, and the initial radius of the training was 3. A bubble type function was used for scaling corrections on neighbor weights. After training the 152 data were again sent through the network, and thus their location on a planar map was determined.

Actually the plane of projection was the surface of a torus. For visualization the torus is cut along two perpendicular lines and the surface spread into a plane.[11] Thus, the top of this map connects with the bottom, and the left edge connects with the right one; this is illustrated in Figure 5. The size of the map is 18 × 18 neurons. The planar map produced is shown in the center of Figure 6. Each of the squares on the map is a neuron. Neurons colored gray have fired by the 152 input data, i.e., these neurons obtained reactions; some of the reactions are shown around the map of Figure 6. In these reaction schemes, oxygen atoms characterizing these reactions are encircled. The number in each neuron indicates the number of reactions which gathered on the corresponding neuron. Other neurons, which are not colored gray, have not fired by the input data, i.e., they have not obtained any reactions.

3.2. Comparison with the Results from PCA. The left side of Figure 7 shows the results from a PCA in a plot of the first two components. Each point in the plot represents a reaction. The horizontal axis is the first principal component, and the vertical axis is the second principal component. The accumulated proportions of the first plus the second principal components are 97.13%, 70.09%, 98.63%, 97.04%, 76.67%, and 91.09% for the parameters, σ-charge, π-charge, σ-electronegativity, π-electronegativity, polarizability, and pK_a, respectively. In other words, the first and second principal components explain almost entirely the information contained in these six parameters. Nine clusters a−i are distinguished for comparing with the results from the Kohonen network as follows.

The planar map on the right side of Figure 7 is the result from the Kohonen network; squares belonging to the same cluster from the result of the PCA (cluster a−i) have been given the same notations (small letter).

The comparison of the plot from the PCA (left side of Figure 7) with the results from the Kohonen network (right side of Figure 7) shows that data close in the multidimensional space are successfully projected into the same or close neurons on the map from the Kohonen network (Remember that this plane is actually the surface of a torous, and neurons at the right margin find their neighbors at the left margin (cf. Figure 5).).

In this case, the PCA gave a good classification with reactions of different types clearly separated in the plot; however, in cases where the inherent dimensionality of the investigated information is higher than two, PCA has the drawback that several plots of different combinations of two coordinate axes have to be analyzed. This makes the analysis of the location of each point in a PCA more complicated than the analysis of the planar map from the Kohonen network. This is the reason why the result from the Kohonen network is used for a more detailed analysis of the classification in what follows.

3.3. Correlation with Substructural Transformations. In Figure 8, the classification by the Kohonen network is examined for substructural transformations at the reaction sites of the input reactions.

In Figure 8A, the planar map is labeled based on the similarity of substructures at the reaction sites of the reactants; Figure 8B shows the similarities of substructures at the reaction sites of the products. Finally, Figure 8C gives the labeling of the map based on similarities of changes in the substructures at the reaction sites in going from the reactants to the products. Thus, for example, neurons with reactions of carbonyl groups (having substructure R=O) are labeled a in Figure 8A (Be aware that the labels of small letters for the neurons in Figure 8−13 are changing from figure to figure and are also different from those in Figure 7.). Those carbonyl reactions that give products with hydroxy groups (having substructure R−OH) are labeled b in Figure 8B. These reactions starting from a carbonyl group and yielding a hydroxyl group are labeled a in Figure 8C.

The results shown in Figures 8A−C indicate that reactions having similar transformations in substructures are gathered in the same or close neurons on the map. In other words, there is a good correspondence between the changes in the substructures at the reaction sites and the changes in the electronic features of the oxygen atoms of the reaction sites and their mapping in the Kohonen network.

3.4. Correlation with Reaction Types. Figures 9−13 show comparisons between the reaction types intellectually assigned by chemists and the classification by the Kohonen

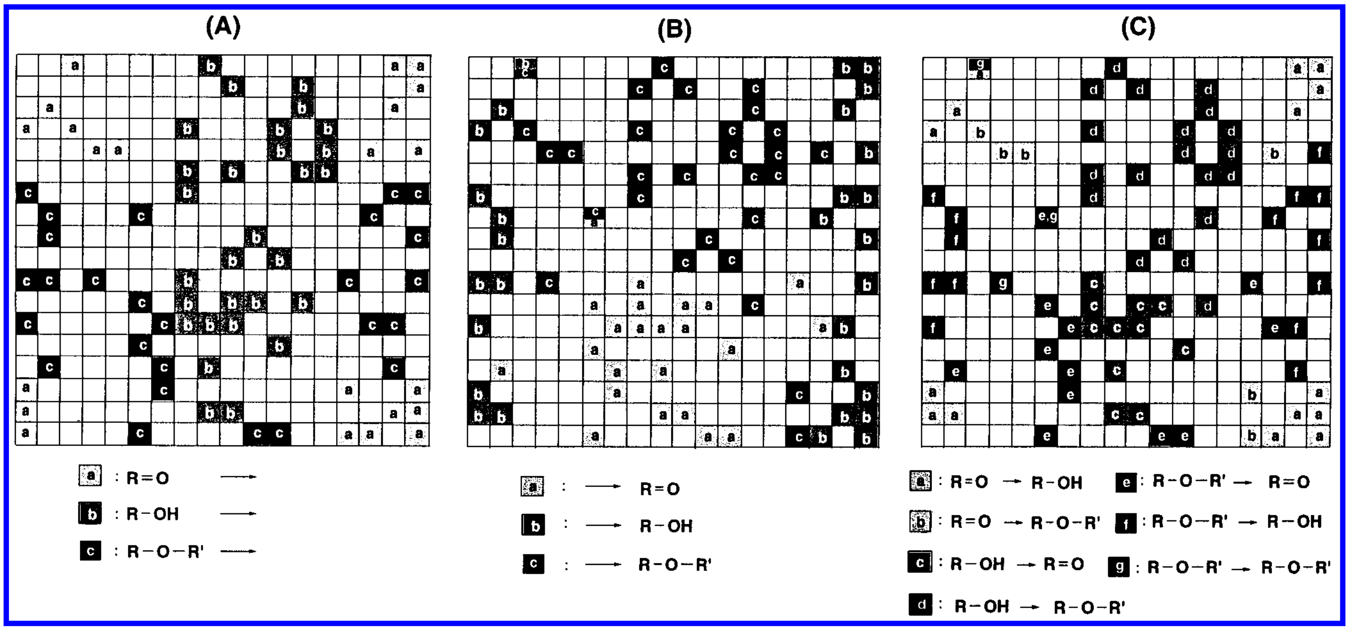


Figure 8. Correlation with substructural transformations. Map A is colored based on the similarity of substructures at the reaction sites of the reactants. Map B is colored based on the similarity of substructures at the reaction sites of products. Map C is colored based on changes in the substructures at the reaction sites in going from the reactants to the products.

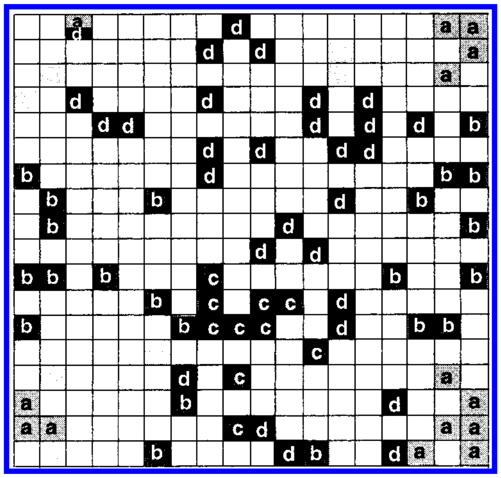


Figure 9. Correlation with reaction types (1). The analysis of all reactions. These are classified into four reaction types: reductions and alkylations (labeled a), cleavage of epoxides, ethers, lactones, and esters (labeled b), oxidation of alcohols (labeled c), and formation of epoxides, ethers, lactones, and esters (labeled d).

network. Such an analysis can be performed from different points of views; one of them is presented here.

Figure 9 shows the analysis of all reactions in the map. Neurons on the map are labeled based on similarity of reaction types. Four different reaction types are recognized, which are labeled a−d in Figure 9. Neurons that have no label have not obtained any reactions or belong to reaction types that are not further analyzed here. These results emphasize that reactions belonging to similar reaction types are on the same or neighboring neurons, and, furthermore, the distribution in the map of Figure 9 is close to that in the map of Figure 7, which corresponds to the clusters found in the PCA. These four reaction categories with the different labels in Figure 9 were analyzed in more detail as shown in Figures 10−13.

Figure 10 shows a more detailed analysis of reactions on neurons labeled a in Figure 9. All of these reactions have a transformation of substructures from R═O to R−OH in their reaction sites. These reactions are classified into three reaction types: reductions (a in Figure 10), reductive alkylations (b in Figure 10), and aldol reactions (c in Figure 10). These reaction types are illustrated on the side of the Kohonen map in Figure 10. Encircled oxygen atoms in the schemes were used for the characterization of the corresponding reactions. Each of these three types forms a separate area in the map. The area of reductive alkylations lies between the area for reductions and the aldol reaction area. This is a reasonable mapping according to similarities between them.

Figure 11 shows a more detailed analysis of reactions on neurons labeled b in Figure 9. The reaction types are illustrated around the Kohonen map in Figure 11. The reason why the b area in Figure 9 is so widely spread is that these reaction comprise a wide range of features in the reactants: cleavage of ethers, epoxides, lactones, and esters. A more detailed analysis recognizes 11 reaction types: cleavage of epoxides (a in Figure 11), rearrangement of epoxides (b in Figure 11), formation of diketones by Pd catalyzed reaction of α, *β*-epoxyketones (c in Figure 11), formation of furans by cleavage of epoxides (d in Figure 11), cleavage of cyclic ethers (e in Figure 11), cleavage of lactones and esters with halogens (f in Figure 11), hydrolysis of enol ethers (g in Figure 11), hydrolysis of enol esters (h in Figure 11), oxidation of aromatics (i in Figure 11), McMurry coupling (j in Figure 11), and Claisen rearrangement (k in Figure 11). Each of these types forms an area in the map. Differences among reactions corresponding to the cleavage of epoxides are distinguishable as four parts of a separate area (a−d). The reason why neurons having reactions belonging to e are somewhat separated is that reaction e contains two types of products as shown in the reaction schemes.

Figure 12 shows a more detailed analysis of reactions on neurons labeled c in Figure 9. These reaction types are illustrated around the Kohonen map in Figure 12. All reactions in this area are oxidations of alcohols. A more detailed analysis recognizes seven reaction types: oxidation

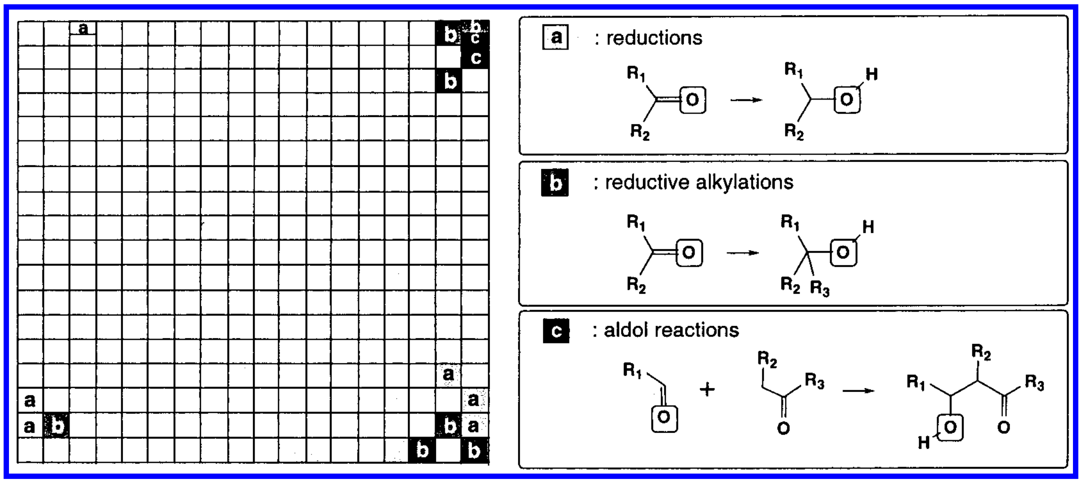


Figure 10. Correlation with reaction types (2). Detailed analysis of the a area of Figure 8.

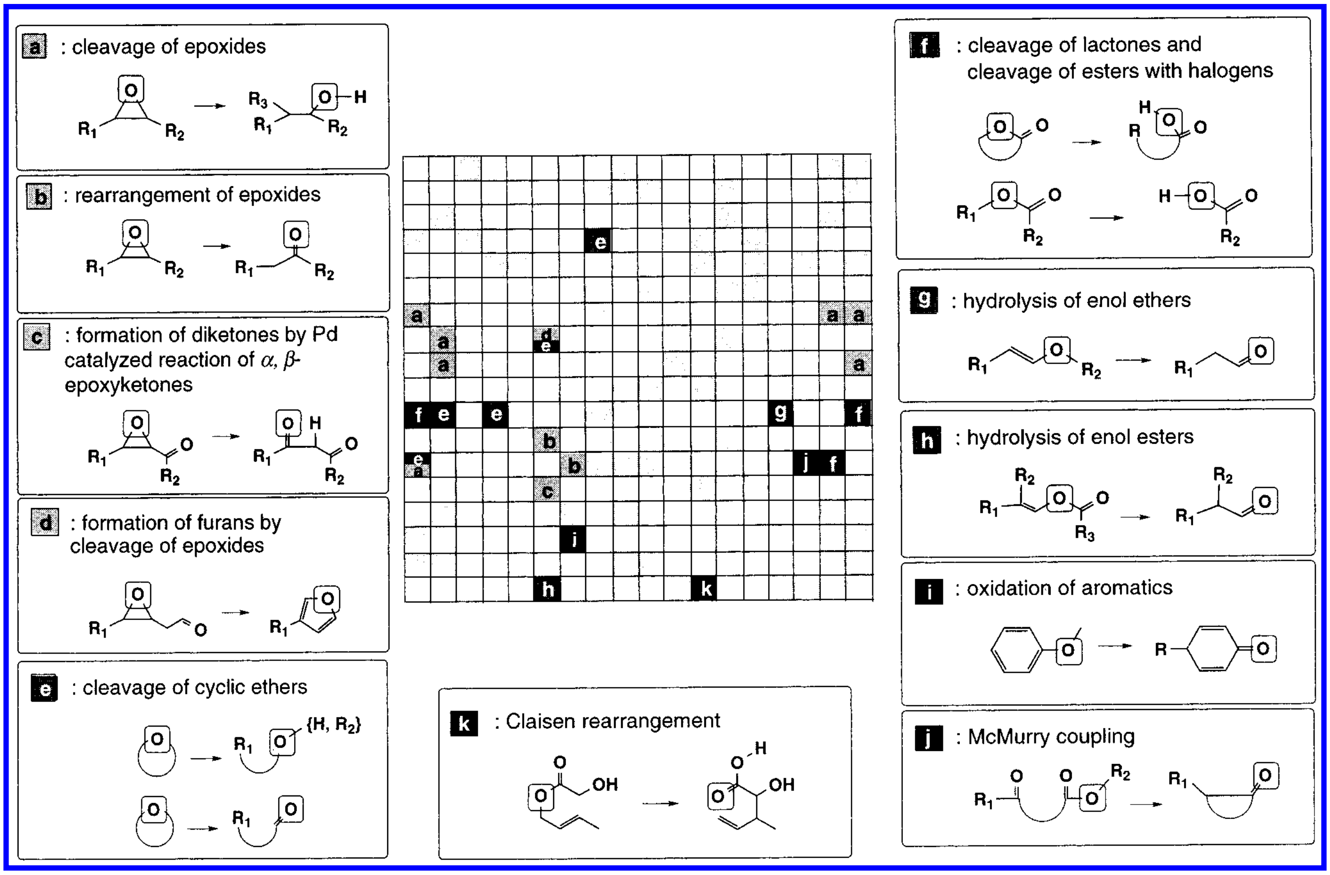

a : cleavage of epoxides
R1
R2
R3
O
H
b : rearrangement of epoxides
c : formation of diketones by Pd catalyzed reaction of α, β-epoxyketones
d : formation of furans by cleavage of epoxides
e : cleavage of cyclic ethers
{H, R2}
k : Claisen rearrangement
OH
f : cleavage of lactones and cleavage of esters with halogens
R
g : hydrolysis of enol ethers
h : hydrolysis of enol esters
i : oxidation of aromatics
j : McMurry coupling
a
b
c
d
e
f
g
h
i
j
k

Figure 11. Correlation with reaction types (3). Detailed analysis of the b area in Figure 8.

of primary alcohols (a in Figure 12), oxidation of acyclic secondary alcohols (b in Figure 12), oxidation of cyclic secondary alcohols (c in Figure 12), oxidation of benzyl alcohols and allyl alcohols (d in Figure 12), oxidation of α-diazo β-hydroxy ketones (e in Figure 12), oxidation of α-hydroxy ketones (f in Figure 12), and oxidation of phenols (g in Figure 12). Each type somehow forms an area in the map. However, differences among a−d are not perceived. All reactions belonging to a−e are on the same or neighboring neurons.

Figure 13 shows a more detailed analysis of reactions on neurons labeled d in Figure 9. The reaction types are illustrated around the Kohonen map in Figure 13. The d area in Figure 9 is also widely spread because these reactions comprise a wide range of features in the products: formation of ethers, epoxides, lactones, and esters. A more detailed analysis recognizes eight reaction types: cyclic etherification of diols (a in Figure 13), haloetherification reactions (b in Figure 13), halolactonization reactions (c in Figure 13), intramolecular acetalizations (d in Figure 13), oxidative lactonization of diols (e in Figure 13), acetylation of alcohols (f in Figure 13), formation of dialkyl carbonates from alcohols (g in Figure 13), and [4 + 2] photoreactions and [2 + 2] photoreactions (h in Figure 13). The notation i in Figure 13 collects all other reactions which were not assigned any reaction type. Each type forms an area of its own in the map. Similar reaction types such as b (haloetherifications) and c (halolactonizations) are on the same or adjacent neurons. Reaction types labeled e−g, which are similar to each other, are also on the same or adjacent neurons.

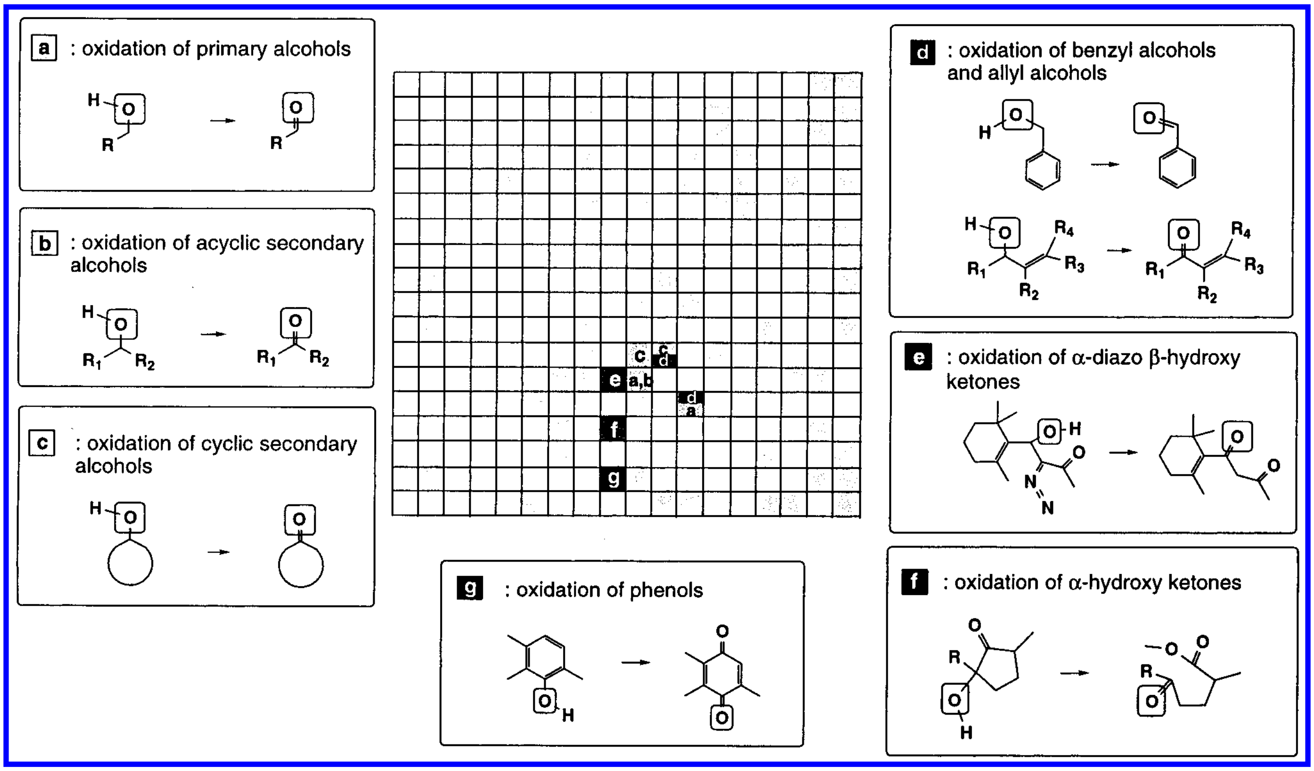

a : oxidation of primary alcohols
b : oxidation of acyclic secondary alcohols
c : oxidation of cyclic secondary alcohols
d : oxidation of benzyl alcohols and allyl alcohols
e : oxidation of α-diazo β-hydroxy ketones
f : oxidation of α-hydroxy ketones
g : oxidation of phenols
c
c d
e a,b
d a
f
g

Figure 12. Correlation with reaction types (4). Detailed analysis of the c area in Figure 8.

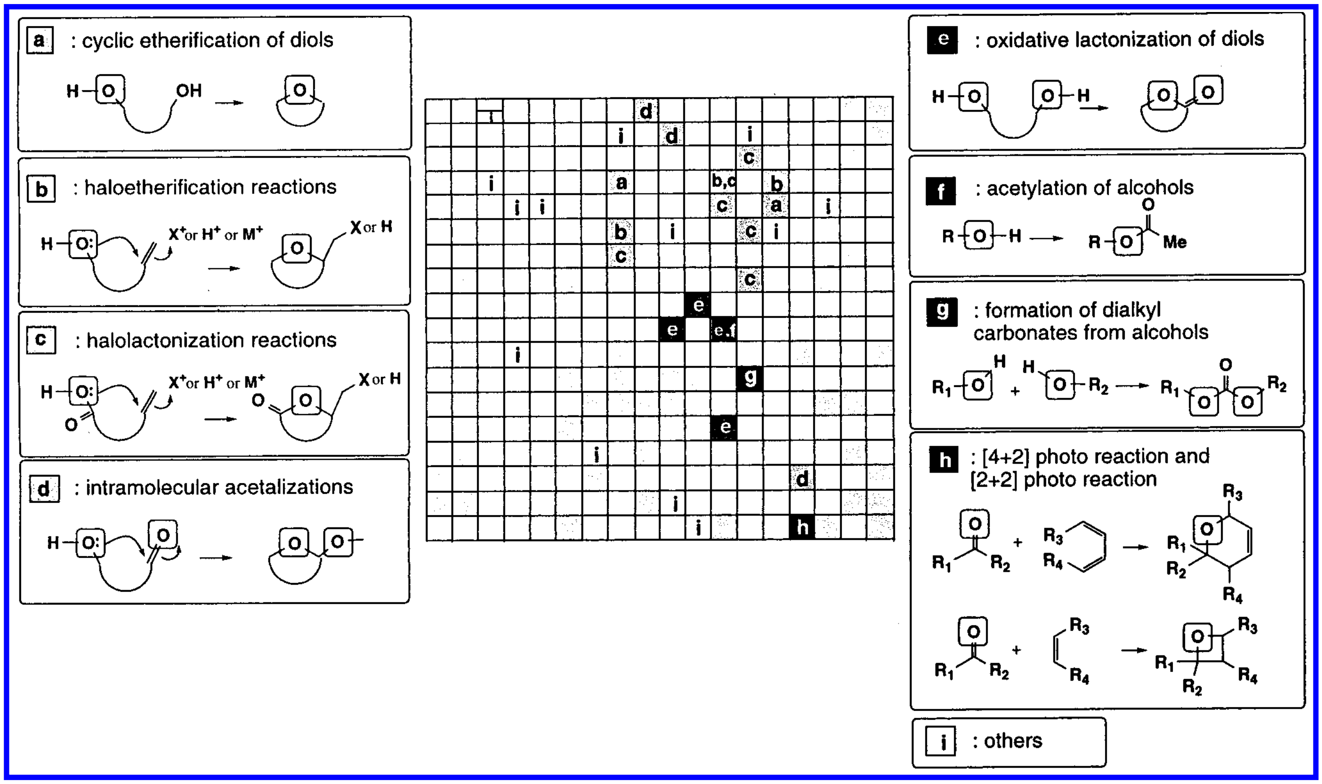

a : cyclic etherification of diols
b : haloetherification reactions
X⁺ or H⁺ or M⁺
X or H
c : halolactonization reactions
d : intramolecular acetalizations
e : oxidative lactonization of diols
f : acetylation of alcohols
g : formation of dialkyl carbonates from alcohols
h : [4+2] photo reaction and [2+2] photo reaction
i : others

Figure 13. Correlation with reaction types (5). Detailed analysis of the d area in Figure 8.

4. CONCLUSION

The results of Kohonen network analyses show that changes in the electronic features in the reaction sites show a good correspondence with changes in the substructures at the reaction sites and also with intellectually assigned reaction types and categories. This means that the representation of reactions by changes in the electronic features at the reaction sites is a valuable approach to the identification of reaction types and categories that have been established by chemists through analyses of reactions from several different points of view.

The changes in the electronic features of reactions are numerical data and thus they provide an easy way for treatment by a computer. The procedures for calculating these electronic factors work fast, and, thus, this approach provides an opportunity for rapidly and accurately processing large amounts of data.

Accordingly, these results indicate the possibility for the systematic classification and organization of large amounts of known organic reactions and construction of novel reaction maps.

The Kohonen neural network has shown its power for similarity perception proceeding through a nonlinear projection of a high-dimensional space into a two-dimensional map.

ACKNOWLEDGMENT

We thank Dr. Kazuo Nagasawa of Synthetic Organic Chemistry Laboratory in The Institute of Physical and Chemical Research (RIKEN) for his help in understanding reactions used for classification. We thank Mr. Daisuke Nishiguchi and Mr. Toshiro Kimura of Department of Knowledge-Based Information Engineering in Toyohashi University of Technology for their technical supports in principal component analysis and Kohonen network computations. We thank Dr. Wolf-Dietrich Ihlenfeldt of Computer-Chemie-Centrum in the Institute of Organic Chemistry at The University of Erlangen-Nürnberg for his assistance in computing physicochemical parameters.

Supporting Information Available: All reaction schemes with their references used for the classification and their positions in the Kohonen map (16 pages). See any current masthead page for ordering and Internet access instructions.

REFERENCES AND NOTES

(1) An original work of this was presented at the Fourth International Conference on Chemical Structures; Noordwijkerhout, June 1996. See: Satoh, H.; Kimura, T.; Funatsu, K. *In proceedings of the Fourth International Conference on Chemical Structures*; Noordwijkerhout, June 1996; p 25.

(2) Chen, L.; Gasteiger, J. *Angew. Chem.* **1996**, *108*, 844, *Angew. Chem., Int. Ed. Engl.* **1996**, *35*, 763.

(3) Chen, L.; Gasteiger, J. *J. Am. Chem. Soc.* **1997**, *119*, 4033.

(4) Funatsu, K.; Sasaki, S. *Tetrahedron Comput. Method.* **1988**, *1*, 27.

(5) Distributed Chemical Graphics, Inc. permitted us to use SYNLIB for our research works.

(6) Gasteiger, J.; Marsili, M. *Tetrahedron* **1980**, *36*, 3219.

(7) Saller, H.; Gasteiger, J. *Angew. Chem.* **1985**, *97*, 699, *Angew. Chem., Int. Ed. Engl.* **1985**, *24*, 687.

(8) Gasteiger, J.; Hutchings, M. G. *J. Chem. Soc., Perkin Trans. 2* **1984**, 559.

(9) Gushurst, A. J.; Jorgensen, W. L. *J. Org. Chem.* **1986**, *51*, 3513.

(10) Kohonen, T. *Biol. Cybern.* **1982**, *43*, 59.

(11) Zupan, J.; Gasteiger, J. *Neural Networks for Chemists. An Introduction*; VCH: Weinheim, 1993.

(12) SOM_PAK, the Self-Organizing Map Program Package was prepared by the SOM Programming team of the Helsinki University of Technology, Laboratory of Computer and Information Science, Pakentajanaukio 2 C, SF-02150 Espoo, Finland. This is available via Internet (cochlea.hut.fi or 130.233.168.48) by anonymous FTP connection.

CI9701190