Simple Method for Identification of Skeletons of Aporphine Alkaloids from ^{13}C NMR Data Using Artificial Neural Networks

Alessandra R. Rufino,[†] Antônio J. C. Brant,[†] João B. O. Santos,[‡] Marcelo J. P. Ferreira,[†] and Vicente P. Emerenciano*,[†]

Instituto de Química, Universidade de São Paulo, Caixa Postal 26077, 05513-970, São Paulo, SP, Brazil, and Faculdade de Engenharia Química, Centro de Engenharias e Ciências Exatas, Universidade Estadual do Oeste do Paraná, Toledo, Paraná, Brazil

Received May 12, 2004

This paper describes the use of artificial neural networks as a theoretical tool in the structural determination of alkaloids from ^{13}C NMR chemical shift data, aiming to identify skeletal types of those compounds. For that, 162 aporphine alkaloids belonging to 12 different skeletons were codified with their respective ^{13}C NMR chemical shifts. Each skeleton pertaining to aporphine alkaloid type was used as output, and the ^{13}C NMR chemical shifts were used as input data of the net. Analyzing the obtained results, one can then affirm the skeleton to which each one of these compounds belongs with high degree of confidence (over 97%). The relation between the correlation coefficient and the number of epochs and the architecture of net (3-layer MLP or 4-layer MLP) were analyzed, too. The analysis showed that the results predicted by the 3-layer MLP networks trained with a number of the epochs higher than 900 epochs are the best ones. The artificial neural nets were shown to be a simple and efficient tool to solve structural elucidation problems making use of ^{13}C NMR chemical shift data, even when a similarity between the searched skeletons occurs, offering fast and accurate results to identification of skeletons of organic compounds.

1. INTRODUCTION

From 1975 up to nowadays, significant progress has been made in the research of aporphine alkaloids, leading to an increasing interest in their isolation from natural sources, mainly in plant families such as Papaveraceae, Annonaceae, Apocynaceae, Ranunculaceae, and the identification of those secondary metabolites afterward. So far, this class of natural products already possesses many compounds records of different chemical structures described in the literature,[1] and many of them have already been synthesized.[2] The greatest interest in studying such alkaloids is directly related to their varied biological activity, encompassing, for instance, anticancerous and anesthetic properties, platelet aggregation inhibitory action, insecticides, among many others.

In alkaloid research, like in all natural products, the interpretation and the structural elucidation of their ^{1}H NMR and ^{13}C NMR spectra both are almost always difficult and arduous tasks for the researchers engaged in this scientific field. It is noticeable that numerous computer-aided expert systems built so far[3] can work databases containing spectroscopic data and get reliable results helpful to the researchers in interpretation and prediction of chemical structures. Besides those expert systems, other computational techniques utilizing statistical methods can also be employed to do such tasks, and artificial neural networks (ANNs) are an excellent option.

ANNs, although they are not a so recent computational technique,[4,5] have reached an explosion of interest worldwide, especially from the past decade to present. They are being successfully applied across an extraordinary range of problem domains, in the most different areas of human knowledge, such as finance,[6−8] medicine,[9−11] engineering,[12,13] chemistry,[14−16] physics,[17,18] mainly in pattern recognition.[19−22] Indeed, anywhere that there are problems of prediction, classification or control, neural networks can be introduced.[8,14,23]

ANNs are robust with respect to small variances in data, such as noise, and may avoid some of the problems encountered with the similarity index. Another disadvantage of the classical sequential comparison of an unknown target spectrum with a set of library spectra (techniques of pattern recognition) is that the search and matching procedure is time-consuming for large libraries. A neural network could speed up the process due to its parallel structure.[23]

As ANNs are becoming each time a more common tool appliable in all areas of human knowledge, chemistry can therefore benefit a lot from this computational methodology from spectroscopic data from IR, NMR (^{1}H and ^{13}C), and MS, among others. Thus, by utilizing ANNs, one can get, for instance, results with very high levels of confidence in the classification of natural products' spectra,[24] structural determination of organic compounds,[25] and chemical shift predictions.[26,27]

This article shows how to use the artificial neural networks as a theoretical tool in structural determination of alkaloids from ^{13}C NMR chemical shift data, aiming to skeleton prediction types of those compounds.

* Corresponding author fax: +55-11-38155579; e-mail: vdpemere@iq.usp.br.
[†] Universidade de São Paulo.
[‡] Universidade Estadual do Oeste do Paraná.

10.1021/ci0498416 CCC: $30.25
Published on Web 04/16/2005

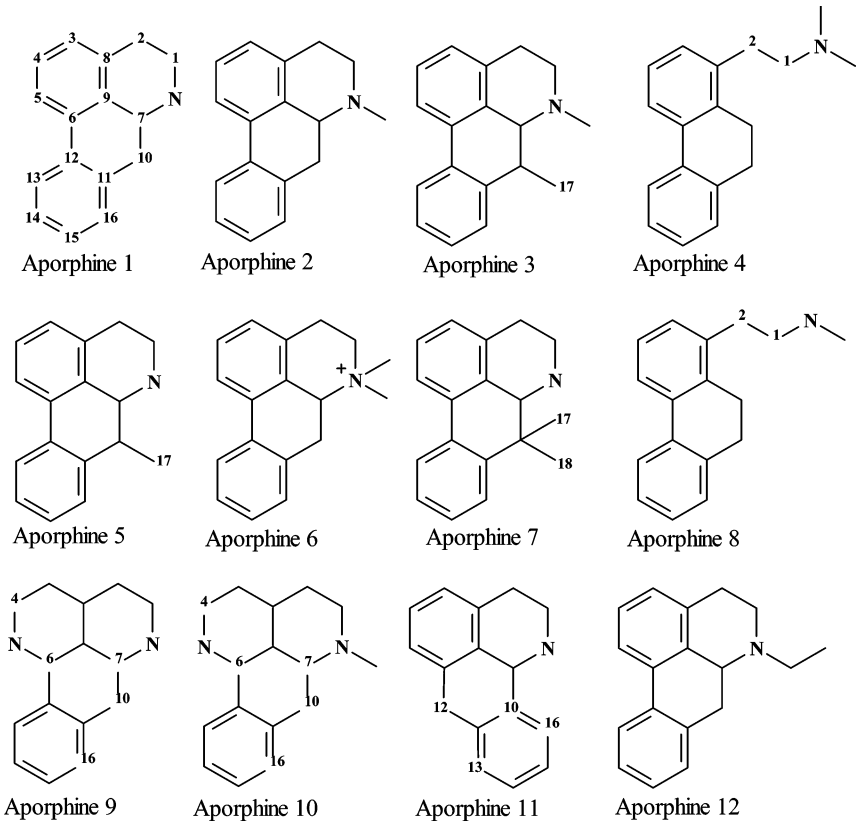

Figure 1. Skeletons of aporphine alkaloids.

2. METHODOLOGY

As the ANNs work through the learning method, their training must be done with the use of well detailed and correct data to avoid an erroneous learning way. So, an Excel worksheet was constructed (163 lines × 35 columns) with the codification of 162 aporphine alkaloids belonging to 12 different skeletons (Figure 1) with their respective chemical shifts of ^{13}C NMR. The construction of this worksheet is the main step for the attainment of good results.

In the first column of the worksheet, the alkaloids present in the database were placed under the form of codes (APF001-APF162). In the 12 following columns, were identified (0 or 1) the skeletons of each one of the 162 structures. In 18 of the last 22 columns are the ^{13}C NMR chemical shifts of each carbon atom of the alkaloid in question. The numeration of the substances proceeds according to a biogenetic numeration commonly used by the chemists of the area, and it was also used in this worksheet encompassing from C-1 to C-18. The 4 remaining columns differentiate which and how many substituents are at the nitrogen of aporphine alkaloids. The codification of the alkaloid 7-methyldehydroglaucine[31] (Figure 2), pertaining to the aporphine skeleton 05 better explains the construction of the worksheet.

In the worksheet containing chemical shifts, the C-1 carbon can appear in five different forms depending on its origin from the chemical moiety of the molecule: from a carbonyl group, a methylene group, an aromatic carbon, a methine group bonded to an O−R group and a methine group bonded to a positive nitrogen atom as well. Independent of the nature of the C-1 carbon atom, all the chemical shifts are placed in the same column. Therefore, the net gets to identify the

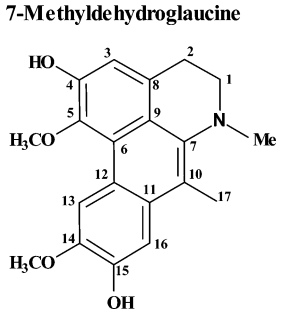

Figure 2. Aporphine alkaloid 7-methyldehydroglaucine.

nature of this carbon through the chemical shifts of C-1 and neighboring atoms. In the same way, the other chemical shifts corresponding to C-2 up to C-18 are placed in their corresponding columns. The total number of carbon atoms may vary in different skeletons, for example, the aporphine ones: 1, 6, 7, 12 (Figure 1). However, when the respective skeleton does not possess a determined carbon atom, for example C-17, C-18, C of methyl group or ethyl group both bonded to the nitrogen atom, the chemical shift value is placed as an interrogation signal (?).

The interrogation signals (?) are treated by the program as missing values and represent the absence of data, once the latter were not recorded in the database. These signals were not attributed because of the difficulty of interpretation of the spectrum or simply such data are not present in the molecule. The codification representing the compound 7-methyldehydroglaucine (Figure 2) is shown in Table 1.

Finished with the construction of the worksheet, the next step consists of carrying it to the program of artificial neural nets (for this work was used the program Statistica Neural Networks Version 4 − StatSoft[32]) and defining the learning

Table 1. Representation of the Compound 7-Methyldehydroglaucine on the Worksheet

Identification of the Skeletons					
Skel 1	Skel 2	Skel 3	Skel 4	Skel 5	Skel 6
0	0	0	0	1	0
Skel 7	Skel 8	Skel 9	Skel 10	Skel 11	Skel 12
0	0	0	0	0	0
^{13}C NMR Chemical Shifts					
C-1	C-2	C-3	C-4	C-5	C-6
49.6	24.9	111.8	149.8	144.6	124.1
C-7	C-8	C-9	C-10	C-11	C-12
140.7	129.7	121.7	120.8	129.6	121.1
C-13	C-14	C-15	C-16	C-17	C-18
109.1	147.1	148.8	104.6	14.1	?
N-Me1[a]	N-Me2[a]	N-Et(CH_2)[a]	N-Et(CH_3)[a]		
42.5	?	?	?		

[a] It is important to recall that the methyl and ethyl groups at the nitrogen are "disfunctionalized" in their representations, which does not mean that the same occurs in all molecules of the worksheet; N-Me2 represents the second methyl bonded to nitrogen. This can be in ionic form (Skeleton 6) or not (Skeleton 4).

conditions. Each skeleton pertaining to aporphine alkaloid type was used as an output of the program for each series of tests.

The ^{13}C NMR chemical shifts were used as input data of the net. In the program Statistica Neural Network − StatSoft, the division of the cases for the first learning of the net is carried out randomly through the ratios of 2:1:1 for training, verification and test of the net, respectively. After the training, depending on the results presented by the net, one has the option to fix these data (if the results are good) or to select a new distribution randomly until the net "understands" that there are examples of well distributed cases that will provide safeness along the execution of trainings, verifications and actual tests. The types of nets selected in this work were multilayer perceptrons with 3 and 4 intermediate layers of neurons (Figure 3).

The complexity of the network, i.e., the number of hidden units in the network, was determined automatically by Statistica Neural Networks Version 4 − StatSoft using Intelligent Problem Solver (IPS). In general, the increase of the number of hidden units leads to an increase of the modeling power of the neural network (it can model a more convoluted, complex underlying function) but also makes it larger, more difficult to train, slower to operate, and more prone to overfitting (modeling noise instead of the underlying function). The decrease of the number of hidden units has the opposite effect.

According to the theory of ANNs, if the data are from a fairly simple function or are very noisy, or if someone is dealing with too few cases, a network with relatively few hidden units is preferable.[23] In an experiment with different numbers of hidden units where larger networks have a better training performance but worse verification performance, probably there has been overfitting, and one should then revert to smaller networks.

In the IPS, the duration of the design process of the chosen medium conducted a fast search for an optimal network. In addition, the selection of networks was chosen to keep the nets with the best performance and to save a maximum number of 10 networks.

To research the probable skeleton of an unknown compound, it was only necessary to add in the worksheet a line

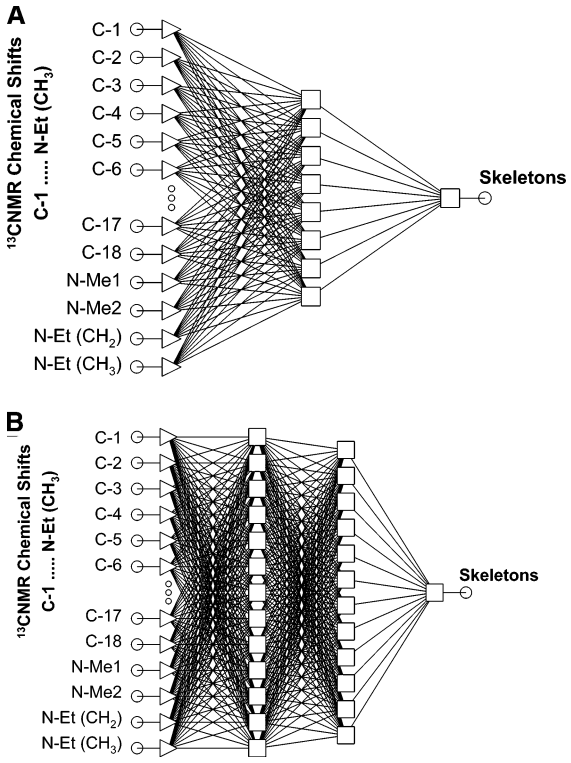

Figure 3. A. Training of the net multilayer perceptron with 3 and 4 intermediate layers of neurons: A) MLP of 3 layers and B) MLP of 4 layers. B) ^{13}C NMR chemical shifts of aporphine alkaloids, inputs; aporphine alkaloid skeletons predicted, outputs.

with the ^{13}C NMR chemical shifts of that compound and to use this line as the test. These chemical shifts were inserted in a random order, once that, if the skeleton is unknown, an assignment cannot be given.

When the training of the net is over, the coherence of the results and the possibility of a certain compound to belong or not to the researched skeleton can be evaluated through RMS of the mistakes shown by the compound in the test, which are observed and supplied by the program on a case by case basis.

For the test cases was reserved $^1/_4$ of the worksheet. To this were added the ^{13}C NMR chemical shift data of 12 compounds (Table 2) belonging to the skeletons of aporphine alkaloids (Figure 4) totally unknown by the net. These compounds were chosen in agreement with the representativeness of their respective skeletons in the worksheet, in other words, the skeletons present in the worksheet which are represented by less than 4 compounds were not used for the test, for the sampling was not appropriate for an efficient learning of the net. Thus, since the ANNs learn only through examples, it should be necessary to supply them various examples for their learning.

To identify the skeletons of the alkaloids in the test by simulating a real situation where these skeletons are unknown, were put the "missing values" (?) in the columns of their identification. Our objective with this procedure was to determine whether the net would be able to identify a correct skeleton of the furnished compound from the ^{13}C NMR data.

Individual trainings were carried out using each one of the 12 present skeletons in the worksheet as outputs (Skeleton 01 − Skeleton 12). The probability results were obtained

Table 2. ^{13}C NMR Chemical Shifts of Aporphine Alkaloids Which Will Be Identified by the Net

carbon	C-1	C-2	C-3	C-4	C-5	C-6	C-7	C-8	C-9	C-10	C-11	C-12	C-13	C-14	C-15	C-16	C-17	N--Me1	N--Me2	skeletons[a]
test A	42.8	29.2	108.3	146.9	142.6	112.3	54.0	129.4	126.4	37.4	129.1	122.8	148.5	145.6	114.5	123.4	?	?	?	SKEL 01
test B	156.6	176.1	112.8	153.3	155.0	124.7	132.7	124.9	120.0	114.3	133.3	127.2	128.1	127.7	128.4	129.6	?	33.5	?	SKEL 02
test C	52.8	23.5	111.0	149.1	149.2	130.5	62.8	123.5	129.9	33.7	123.0	122.9	114.1	144.2	144.8	145.5	?	40.3	?	SKEL 02
test D	55.3	36.0	116.5	145.8	144.3	126.3	140.0	132.8	127.3	106.3	133.1	126.3	126.9	127.0	127.1	131.9	47.8	152.1	?	SKEL 03
test E	60.1	31.1	106.6	146.6	155.6	124.3	120.9	131.5	125.6	121.5	126.6	131.7	115.7	119.6	126.9	151.3	?	44.6	44.6	SKEL 04
test F	45.6	19.8	147.1	143.4	145.2	116.5	172.0	122.7	117.4	73.1	145.2	121.1	129.5	113.6	159.5	109.5	33.7	?	?	SKEL 05
test G	45.8	26.4	105.6	151.2	142.4	115.6	169.7	132.8	113.6	72.5	145.2	106.6	157.6	98.4	161.3	101.1	32.4	?	?	SKEL 05
test H	61.5	23.5	108.0	148.0	142.2	118.2	69.7	119.8	118.3	28.6	123.7	122.5	112.1	146.2	145.5	114.3	?	42.6	53.4	SKEL 06
test I	60.3	23.8	110.6	152.9	142.2	126.0	69.1	125.2	118.4	28.9	124.3	120.2	143.5	149.7	111.5	119.6	?	42.9	53.5	SKEL 06
test J	59.4	32.5	116.0	150.5	145.5	124.2	127.3	129.3	125.4	128.2	129.3	132.4	121.6	126.7	126.7	125.7	?	49.1	?	SKEL 08
test K	148.1	123.2	118.9	147.1	?	150.7	147.5	138.3	119.3	181.5	132.0	135.0	125.1	134.4	131.1	128.1	?	?	?	SKEL 09
test L	42.7	24.2	140.9	134.8	144.0	108.1	54.1	119.1	130.8	37.6	128.9	122.5	148.3	144.5	113.6	123.4	?	?	?	SKEL 01

[a] Each one of the tested alkaloids is identified according to the skeleton it belongs to. In the other columns are their corresponding ^{13}C NMR chemical shifts.

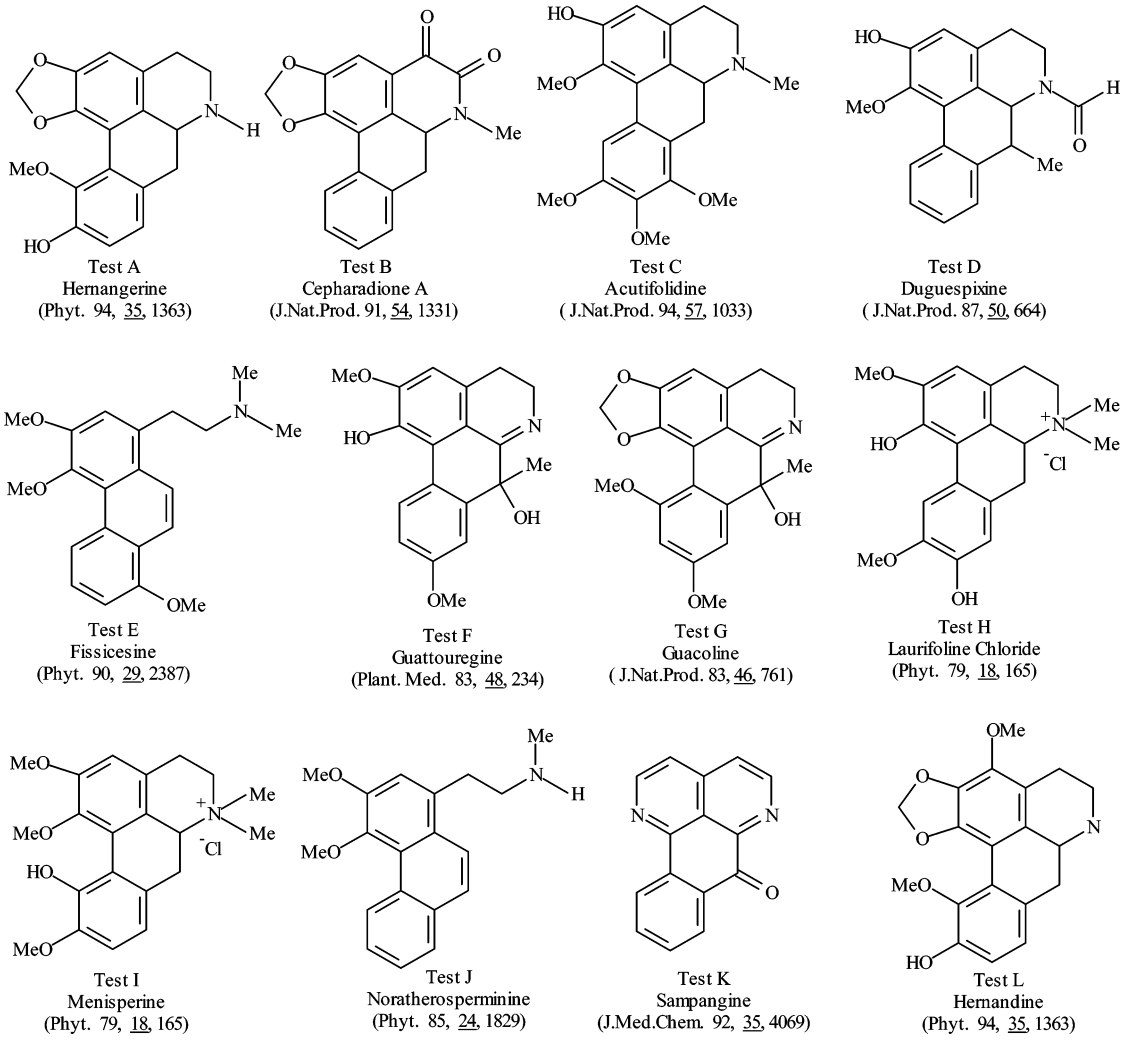

Figure 4. The 12 substances belonging to the skeletons of the aporphine alkaloids used for the test.

for the compounds used to test the prediction ability of the net. Thus, the net should furnish a result that varies from 0 to 1 which allows the user to define if the compound belongs to the respective skeleton as well as the correlation of the trainings and tests in the MLR (multiple linear regression) form with the real data versus observed data.

MLP of 3 and 4 layers were used, and the efficiency of both types was compared. The complexity of the net was determined automatically, and the 10 better nets were selected.

3. RESULTS AND DISCUSSION

After training the net with all of the skeletons, the obtained results are described in Table 3, which allowed us to identify the skeleton type of each one of the compounds tested by the net.

For the net training using Skeleton 01 as output was observed that the compounds A and L have respectively 98.78 and 99.05% of the probability to pertaining to this skeleton. It was also observed that none of the compounds

Table 3. Probability of the Test Compound To Belong to the Skeletons Researched

tested skeletons	test compound (%)												correlations		RMS error	
	A	B	C	D	E	F	G	H	I	J	K	L	training	test	training	test
SKEL 01	98.78	14.31	1.12	1.71	1.10	0.14	1.68	1.39	1.93	10.95	0.17	99.05	0.99178	0.86230	0.06158	0.25040
SKEL 02	3.69	99.69	97.84	5.78	17.4	3.76	5.44	11.95	3.69	8.79	1.91	24.45	0.99987	0.87208	0.00659	0.21450
SKEL 03	0.53	0.50	0.15	99.72	7.82	1.10	0.23	1.46	0.35	2.04	1.02	2.70	0.99873	0.94651	0.00861	0.13010
SKEL 04	0.06	0.16	0.59	0.23	99.45	1.25	1.85	2.42	1.93	1.87	1.22	0.48	0.99212	0.78715	0.01529	0.13650
SKEL 05	0.22	0.31	0.33	0.57	0.17	99.75	99.40	0.04	0.02	0.52	0.61	0.07	0.99421	0.78789	0.02343	0.17350
SKEL 06	0.87	2.38	6.71	1.66	1.88	0.36	1.43	98.79	99.96	0.34	0.59	0.49	0.98983	0.79407	0.04327	0.19130
SKEL 07	0.00	0.03	0.01	0.02	0.02	0.01	0.00	0.02	0.01	0.00	0.00	0.00	0.99999	0.99990	0.00041	0.01535
SKEL 08	4.92	4.35	4.29	5.29	4.57	4.22	4.66	4.19	4.56	98.65	4.67	5.54	0.99068	0.74252	0.04575	0.12870
SKEL 09	0.04	1.45	0.47	1.63	0.04	0.78	0.23	0.43	1.32	0.85	98.17	0.79	0.99885		0.00720	0.12900
SKEL 10	2.78	2.72	0.89	1.95	1.25	2.49	2.53	1.83	1.13	0.90	2.65	0.87	*a*	*a*	0.02478	0.07648
SKEL 11	0.01	0.01	0.01	0	0.00	0.01	0.01	0.01	0	0.01	0.00	0.02	*a*	*a*	0.13360	0.00780
SKEL 12	1.45	1.85	1.92	1.98	0.34	0.05	0.56	1.38	2.27	0.69	1.32	1.95	*a*	*a*	0.01035	0.01294

[a] Skeletons not possessing a minimal number of compounds that can supply a correlation.

Table 4. Identification of the Skeleton

test compound	skeleton	probability%
test A	SKEL 01	98.78
test B	SKEL 02	99.69
test C	SKEL 02	97.84
test D	SKEL 03	99.72
test E	SKEL 04	99.45
test F	SKEL 05	99.75
test G	SKEL 05	99.40
test H	SKEL 06	98.89
test I	SKEL 06	99.96
test J	SKEL 08	98.65
test K	SKEL 09	98.17
test L	SKEL 01	99.05

used for the test pertain to skeletons 10, 11 and 12. The results also indicate that there are no possibilities of the same compound belonging to more than one skeleton. In the same way all the other skeletons were trained, and one can conclude that the net identifies a skeleton positively and does not supply a false positive result. This fact can be evidenced when one observes, for example, the probability of test A belonging to Skeleton 03 (0.53%), whose value is practically insignificant. The results are summarily described in Table 4.

Analyzing Table 4, one can affirm, with a high degree of confidence, the skeleton to which each one of these compounds belongs.

From the proposals of skeletons shown in Table 4, the missing values, previously attributed to the tested compounds, were replaced by these proposals of skeletons. So, the net was retrained using a new worksheet where all the compounds were duly identified. Thus, some parameters such as, for example, type of used algorithm, number of epochs and number of layers of the net, were varied in order to obtain the highest performance of the net. Through this procedure, the net efficiency was increased, and the results were obtained with greater reliability.

The ANNs explored in this work identified all of the skeletons of the compounds submitted to the tests, showing high confidence probability results (over 97%), even in the cases where the substances belonging to the same skeleton had a great difference in the ^{13}C NMR chemical shift values, as occurred in the cases of the chemical shifts of carbon atoms 1 and 2 (156.6δ; 176.1δ and 52.8δ; 23.5δ respectively for tests B and C (Table 2)). On the other hand, the ANNs were able to differentiate skeletons in which the great

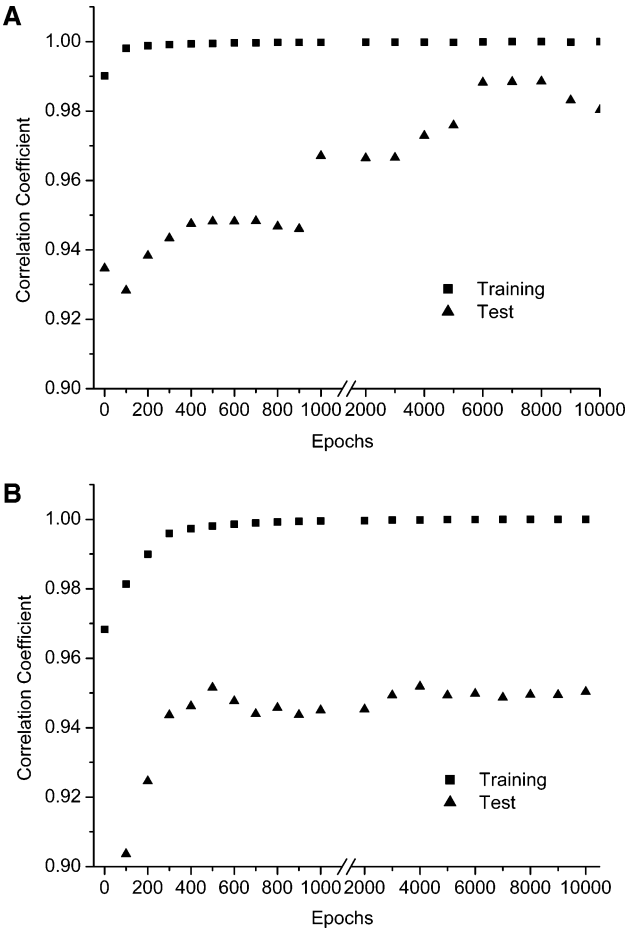

Figure 5. Epochs versus correlation coefficient (training and test) for Skeleton 01 of the aporphine alkaloid: A) MLP of 3 layers and B) MLP of 4 layers.

majority of the chemical shifts were similar, differentiating among them the presence or absence of the group methyl, as was observed in tests A and C (Table 2).

The number of epochs was increased gradually, in a range of 100−1000 (at a 100 scale) and then in a range of 1000−10000 (at a 1000 scale). By using this procedure, it was possible to observe the relation between the correlation coefficient and the number of epochs represented graphically in Figure 5.

Figure 5 shows two different behaviors for the training and testing of the ANN in consequence of the increase of the epochs number. The correlation coefficient of the

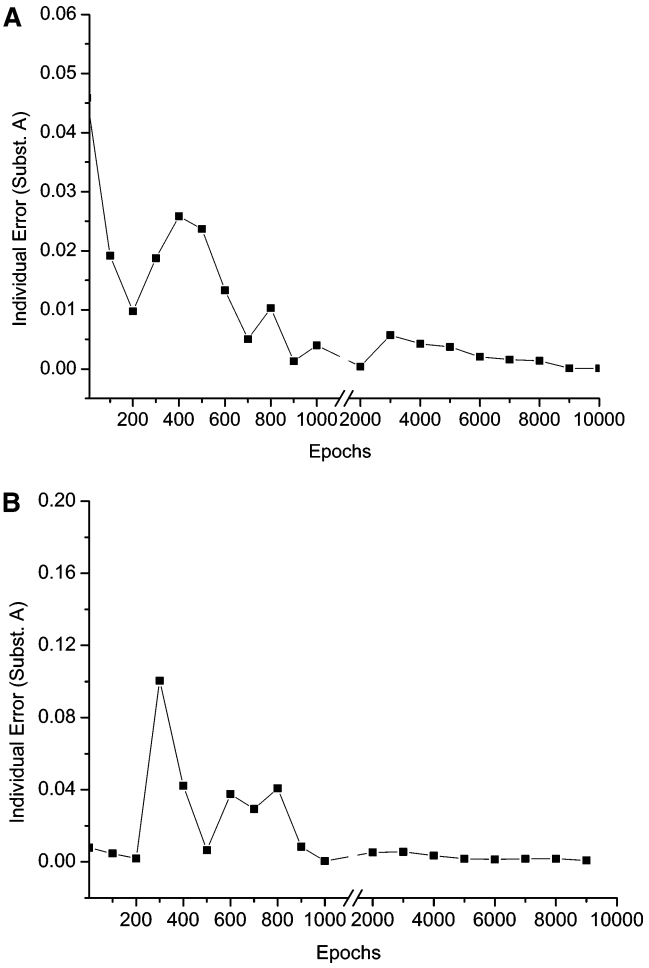

Figure 6. Individual error for test A versus Epochs: A) MLP of 3 layers and B) MLP of 4 layers.

trainings have kept constant and always present a good performance for both the nets (of 3 and 4 intermediate layers). Regarding the 3-layer MLP, it was observed that the correlation coefficient obtained in the tests increased from 0.93 to 0.99 when the epochs number was varied from 1 to 6000. The correlation coefficient was kept almost constant between 6000 and 8000 epochs, and it was decreased after 8000 epochs. The best correlation coefficient observed for the tests was in the training of the 3-layer MLP with 7000 epochs. In relation to the 4-layer MLP, the correlation coefficient for the tests was kept almost constant (close to 0.95) between 600 and 10000 epochs. These facts show that the results predicted by the 3-layer MLP are better than those obtained by the 4-layer MLP. The results predicted by both nets were the worst ones when the training with the number of the epochs was lower than 900 epochs. Thus, the best net architecture found for the training of Skeleton 01 was the 3-layer MLP of the type 22:22−5−1:1. This notation means that the net had 22 inputs (chemical shifts) and 1 output (skeleton), 3 layers of hidden neurons (hidden layers) with 22, 5 and 1 neurons, respectively.

Figure 6 shows that there is a certain relation between the number of epochs and the errors (0−1) presented by test A in the trainings of Skeleton 01.

One can observe that, from 100 to 800 epochs, the graph does not show a direct relation between the number of epochs and the individual errors presented for the identification of the test compound; however, over 2000 epochs, Figure 6 shows a constant reduction of the error for both 3- and 4-layer MLP, and the best results were also achieved by the 3-layer MLP networks.

4. CONCLUSIONS

For all trainings of the nets and in the graphs presented previously, it was possible to notice that the multilayer perceptron networks with three intermediate layers (hidden) showed better results than the multilayer perceptron nets with four intermediate layers.

The variations of values for the learning speed and the noise had little influence on the results. However, it could be observed that the increase in the number of epochs significantly improved the results supplied by the nets. The artificial neural nets were shown to be a simple and efficient tool to solve structural elucidation problems making use of ^{13}C NMR chemical shift data, even when a high degree of similarity occurs between the searched skeletons. Summarizing, artificial neural networks offer fast and accurate results for identification of skeletons and for assigning unknown compounds among distinct fingerprints (skeletons) of aporphine alkaloids. The computation method is much faster than the utilization of the traditional methods for skeleton prediction, which makes neural neworks ideal for selecting results for structure generators or checking the entries of a database. If a large number of skeletons has to be predicted or a fast and easy check of a structure is necessary, this approach is advantageous. Moreover, the large amount of the disk space for saving the database or long time for loading data from external computers are no longer necessary. It would also be possible to perform the training of the networks interactively, so that every researcher dealing with the skeleton identification of natural products could create a network specialized in groups of such complex substances.

Through this work, one can notice that the methodology here described and applied was useful in identifying different aporphine skeletons. The qualifying capacity of the artificial neural nets was tested with different compounds pertaining to diverse types of skeletons, presenting in all cases good results. Hence, it was possible to determine the correct skeleton of unknown compounds.

Artificial neural nets for the prediction of chemical data such as those from IR, MS, ^{1}H and ^{13}C NMR are being used increasingly. However, it is worthy of pointing out that this current work is a pioneering one, once it uses artificial neural nets for identification of aporphine skeletons based upon chemical data and furthermore demonstrates that this methodology may also be adapted for use in future structural elucidation studies.

ACKNOWLEDGMENT

A. R. Rufino acknowledges the financial support received from FAPESP (Grant N. 99/08242-4). This work was supported by grants from the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) and by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

REFERENCES AND NOTES

(1) Guinaudeau, H.; Leboeuf, M.; Cave, A. Aporphinoid alkaloids. 5. *J. Nat. Prod.* **1994**, *57*, 1033−1135.

(2) Guinaudeau, H.; Leboeuf, M.; Cave, A. Aporphinoid alkaloids. 3. *J. Nat. Prod.* **1983**, *46*, 761−835.
(3) Emerenciano, V. P. Determinação estrutural computadorizada − Princípios, métodos, perspectivas e mitos. *Quím. Nova* **1993**, *16*, 551−559.
(4) McCulloch, W. S.; Pitts, W. A logical calculas of the ideas immanent in nervous activity. *Bulle. Mathematical Biophys.* **1943**, *5*, 115−133.
(5) Hebb, D. O. *The Organization of Behavior*; Wiley: 1949.
(6) Chen, A. S.; Leung, M. T. Regression neural network for error correction in foreign exchange forecasting and trading. *Comput. Oper. Res.* **2004**, *31*, 1049−1068.
(7) Ye, Z. X.; Lu, P. L. Analysis of corporate finance using statistical and intelligent approaches. *Insurance Mathematics Econ.* **2003**, *32*, 487−488.
(8) Kodogiannis, V.; Lolis, A. Forecasting financial times series using neural network and fuzzy system-based techniques. *Neural Comput. Appl.* **2002**, *11*, 90−102.
(9) Davis, G. E.; Lowell, M. D. In *A Neural Network that predicts Psychiatric Length of Stay*; Davis, D., Davis, G. L., Eds.; MD Computing, NY, Springer-Verlag: 1993.
(10) Kallergi, M. Computer-aided diagnosis of mammographic microcalcification clusters. *Med. Phys.* **2004**, *31*, 314−326.
(11) Crawford, E. D. Use of algorithms as determinants for individual patient decision making: National comprehensive cancer network versus artificial neural networks. *Urology* **2003**, *62*, 13−19.
(12) Hsu, Y. L.; Wang, S. G.; Yu, C. C. A sequential approximation method using neural networks for engineering design optimization problems. *Eng. Optimiz.* **2003**, *35*, 489−511.
(13) Molga, E. J. Neural network approach to support modeling of chemical reactors: problems, resolutions, criteria of application. *Chem. Eng. Process.* **2003**, *42*, 675−695.
(14) Zupan, J.; Gasteiger, J. *Neural Networks for Chemists*: A*n Introduction*; VHC−Verlag: Weinheim, 1993.
(15) Meiler, J.; Maier, W.; Meusinger, R. Using neural networks for C-13 NMR chemical shift prediction-comparison with traditional methods. *J. Magn. Res.* **2002**, *157*, 242−252.
(16) Ferreira, M. J. P.; Brant, A. J. C.; Alvarenga, S. A. V.; Magri, F. M. M.; Rufino, A. R.; Emerenciano, V. P. Prediction of occurrences of diverse chemical classes in Asteraceae family through artificial neural networks. *Phytochem. Anal.* **2004**, *15*, 389−396.
(17) Capdevila, C.; Caballero, F. G.; de Andres, C. G. Prediction of martensite start temperature by neural network analysis. *J. Phys. IV* **2003**, *112*, 217−221.
(18) Kim, B.; Kwon, K. H. Qualitative modeling of silica plasma etching using neural network. *J. Appl. Phys.* **2003**, *93*, 76−82.
(19) Penza, M.; Cassano, G. Recognition of adulteration of Italian wines by thin-film multisensor array an artificial neural networks. *Anal. Chim. Acta* **2004**, *509*, 159−177.
(20) Haddadnia, J.; Karim, F.; Majid, A. A fuzzy hybrid learning algorithm for radial basis function neural network with application in human face recognition. *Pattern Recognit.* **2003**, *36*, 1187−1202.
(21) Turkoglu, I.; Arslan, A.; Ilkay, E. An intelligent system for diagnosis of the heart valve diseases with wavelet packet neural networks. *Comput. Biol. Med.* **2003**, *33*, 319−331.
(22) Aracena, J.; Goles, E. Complexity of perceptron recognition for a class of geometric patterns. *Theor. Comput. Sci.* **2003**, *299*, 65−79.
(23) Lawrence, J. *Introduction to Neural Networks, Design, Theory and Apllications*, 6th Edition; Luedeking, S., Ed.; July − 1994.
(24) Fraser, L. A.; Mulholland, D. A. A robust technique for the group classification of the C-13NMR spectra of natural products from Meliaceae. *Fresenius J. Anal. Chem.* **1999**, *365*, 631−634.
(25) Blinov, K. A.; Elyashberg, M. E.; Molodtsov, S. G.; Williams, A. J.; Martirosian, E. R. An expert system for automated structure elucidation utilizing H-1-H-1, C-13-H-1 and N-15-H-1 2D NMR correlations. *Fresenius J. Anal. Chem.* **2001**, *369*, 709−714.
(26) Militão, J. S. L. T.; Emerenciano, V. P.; Ferreira, M. J. P.; Cabrol-Bass, D.; Rouillard, M. Structure validation in computer-supported structure elucidation system: C-13 NMR shift predictions for steroids. *Chemom. Intell. Lab. Syst.* **2003**, *67*, 5−20.
(27) Meiler, J.; Sanli, E.; Junker, J.; Meusinger, R.; Lindel, T.; Will, M.; Maier, W.; Kock, M. Validation of structural proposals by substructure analysis and C-13 NMR chemical shift prediction. *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 241−248.
(28) Rumelhart, D. E.; Hilton, G. E.; Williams, R. J. Learning representations by back-propagating errors. *Nature* **1986**, *323*, 533−536.
(29) Rumelhart, D. E.; McClelland, J. L. *Parallel Distributed Processing, Vol. 1*; Foundations the MIT Press: 1986.
(30) Mittermayr, C. R.; Drouen, A. C. J. H.; Otto, M.; Grasserbauer, M. Neural networks for library search of ultraviolet-spectra. *Anal. Chim. Acta* **1994**, *294*, 227−242.
(31) Guinaudeau, H.; Leboeuf, M.; Cave, A. Aporphinoid alkaloids. 4. *J. Nat. Prod*. **1988**, *51*, 389−474.
(32) Statistica Neural Networks, Statsoft, Inc. Tulsa, OK, USA, 2000.

CI0498416