Available online at www.sciencedirect.com
ScienceDirect
Dyes and Pigments 74 (2007) 416–423
DYES and PIGMENTS
www.elsevier.com/locate/dyepig

# Quantum chemical and statistical study of hypocrellin dyes with phototoxicity against tumor cells

Hai-long Yang*, Fu-ru Huang

*School of Life and Environmental Sciences, Wenzhou University, Middle Xueyuan Road, Wenzhou 325027, PR China*

Received 23 February 2006; received in revised form 1 March 2006; accepted 3 March 2006
Available online 9 May 2006

**Abstract**

A set of molecular properties (variables) of 20 hypocrellin dyes (hypocrellins and their derivatives) with phototoxicity against tumor cells was calculated by the molecular orbital semi-empirical method AM1 and ChemPropStd. Pattern recognition techniques, principal component analysis (PCA) and hierarchical cluster analysis (HCA) were employed to reduce dimensionality and investigate which subset of variables could be more effective for classifying hypocrellins and their derivatives according to their degree of phototoxicities against tumor cells. The PCA and HCA studies showed that $\mu$ (dipole moment) and $Q_3$, $Q_4$, $Q_9$ and $Q_{10}$ (charges on atoms 3, 4, 9 and 10) were the most important variables for the classification between the hypocrellin dyes with higher and lower phototoxicities against tumor cells.


*Keywords:* Hypocrellin dyes; Phototoxicity; AM1; Principal component analysis; Hierarchical cluster analysis

## 1. Introduction

*Hypocrella bambusae* Sacc. is a parasitic fungus of bamboos, which grows abundantly in the northwestern mountains of Yunnan Province and in the southeastern region of Tibet Autonomous Region of China. A series of perylenequinoid pigments, including hypocrellin A (HA) and hypocrellin B (HB) are the metabolites of *H. bambusae*. Originally, hypocrellins were utilized as a potent therapeutic agent for white lesions of the vulva, keloid, vitiligo, psoriasis, tinea capitis and lichen amyloidosis. Modern researches revealed that hypocrellins showed excellent light-induced antitumor and antiviral activities [1–6]. Previous works had also reported a series of other activities related to hypocrellins, such as treatment of some vascular diseases [7].

Since separated in 1980s [8], hypocrellins have been receiving intensive interest in photodynamic therapy (PDT) due to their wide absorption band in the visible region and extremely high singlet oxygen ($^1O_2$) generation ability over the past two decades [9–12]. From the viewpoint of clinical applications, the water solubility and absorption intensity in the phototherapeutic window (600–900 nm) of the natural hypocrellins need to be improved for attaining ideal photodynamic efficacy. With these characteristics, as well as the ability to be chemically modified, a lot of derivatives of the parent hypocrellins A and B have been synthesized to optimize the photosensitizing properties [13–20].

Nowadays, structure–activity relationship (SAR) studies have been proven to be helpful in the understanding of the influence of molecular properties on the biological activity presented by several kinds of compounds [21]. In order to get high active photosensitizer, the relationship between the structure and phototoxicity of hypocrellins should be investigated. The quantum chemical parameters of molecules and even of the interacting molecular systems can, in principle, express all electronic properties related to the molecular interactions. Thus, SAR studies using quantum chemical parameters have become important in qualitative and quantitative analyses of three-dimensional molecular interactions [22,23].

* Corresponding author. Tel.: +86 577 88371274; fax: +86 577 89789101.
*E-mail address:* yanghl999@yahoo.com (H.-long Yang).

0143-7208/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.dyepig.2006.03.002

In the present work, we employ the semi-empirical AM1 method [24] to calculate selected quantum chemical molecular variables of 20 hypocrellin dyes reported in the literature as presenting a certain degree of phototoxicity against tumor cells [1,13–16]. The pattern recognition methods, principal component analysis (PCA) [25] and hierarchical cluster analysis (HCA) [26], have been employed to obtain a relationship between the calculated variables and phototoxicity against tumor cells.

## 2. Methodology

### 2.1. Compounds

The main structure and the numbering we have adopted to study the hypocrellin dyes are shown in Fig. 1. The structure and the number of the 20 compounds are shown in Fig. 2. The phototoxicities of the hypocrellin dyes are expressed by the numerical indicator, $IC_{50}$, which indicates the concentration to exert 50% growth inhibition against tumor cells, including EMT6/Ed murine tumor cell [1], human oral cavity epithelial carcinoma KB cell [13,14], human gastric adenocarcinoma MGC803 cell [15], and human cervix uteri tumor HeLa cell [16]. The respective $IC_{50}$ of all the 20 hypocrellins studied is shown in Table 1.

### 2.2. Calculation of the atomic and molecular descriptors

All the molecular structures of the tested hypocrellins (numbered from **1** to **20** in Fig. 2) were optimized by using the molecular mechanics method MM2 [27] and the semi-empirical method AM1 [24]. Then the following descriptors were calculated in this work:

- the energy of the HOMO (highest occupied molecular orbital energy) and LUMO (lowest unoccupied molecular orbital energy);
- Mulliken electronegativity ($\chi$): obtained from the following equation: $\chi = (E_{HOMO} + E_{LUMO})/2$;
- electron affinity (EA): obtained as $(-E_{LUMO})$;
- dipole moment ($\mu$), heat of formation ($\Delta H_f$), total energy ($E_T$), electronic energy ($E_{el}$);
- partition coefficient ($\log P$);
- Connolly molecular area (MA) and Connolly solvent-excluded volume (SEV);
- net atomic charge on atom N ($Q_n$), where $n = 2, 3, 4, 5, 8, 9, 10, 13, 14, 15$.

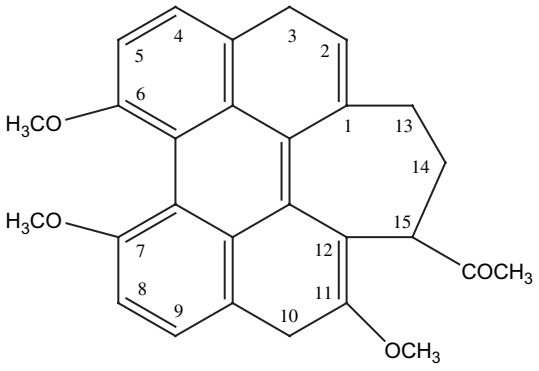



Fig. 1. Structural skeleton and numbering of the hypocrellins studied.

The calculated descriptors selected could represent electronic ($E_{HOMO}$, $E_{LUMO}$, $\chi$, EA, $\mu$, $\Delta H_f$, $E_T$, $E_{el}$ and $Q_n$), steric (MA and SEV) and hydrophobic ($\log P$) features of the compounds studied. The variables were calculated with the semi-empirical AM1 method or ChemPropStd combined in CS Chem3D Ultra 6.0 program [28]. The statistical analyses (PCA and HCA) were performed using the MATLAB 6.0 program [29]. Before applying the PCA and HCA methods, each variable was standardized so that they could be compared to each other on the same scale.

## 3. Results and discussion

### 3.1. Principal component analysis (PCA)

The main purpose of employing the PCA method is to reduce the number of variables used in the analysis. The method creates new variables as linear combinations of all the initial variables so that the first new variable contains the largest variance, the second new variable contains the second largest variance, and so on, until the last variable can be truncated. The PCA method also allows us to diminish the number of total variables in a data set.

In this work, after several attempts to obtain a good classification of the compounds, the best separation was obtained with five variables (see Table 1) out of the 21 we had initially. This suggests that the other 16 variables are not important for classifying these compounds.

The results of the PCA calculation show that the first three principal components (PC1, PC2, and PC3) describe 94.009% of the overall variance as follows (%): PC1 = 73.358, PC2 = 11.807, PC3 = 8.844 (see Table 2). The first two principal components (PC1 and PC2) describe 85.165% of the total variance. Table 3 shows the loading vectors for PC1, PC2, and PC3.

The plots of the score vectors of the principal components (PC1 × PC2) and (PC1 × PC3) are shown in Figs. 3 and 4. We can see from the figures that the hypocrellins studied are separated into two groups, A and B. Group A contains the hypocrellins (compounds **7**, **9**, **13**, **14**, **15**, **16**, **17**, **19**, **20** – see Table 1) with higher degree of phototoxicities against tumor cells, i.e. the molecules with $IC_{50} < 1.0\ \mu g\ mL^{-1}$ except compounds **14** and **15**. Group B consists of the hypocrellins (compounds **1**, **2**, **3**, **4**, **5**, **6**, **8**, **10**, **11**, **12**, and **18** – see Table 1) with lower phototoxicities against tumor cells, i.e. the molecules with $IC_{50} > 1.0\ \mu g\ mL^{-1}$ except compounds **8** and **18**. Taking into account the difference in tumor cell lines and the experimental error, we think that the classification of the 20 hypocrellins by PCA is consistent with their phototoxicities tested by the experiment.

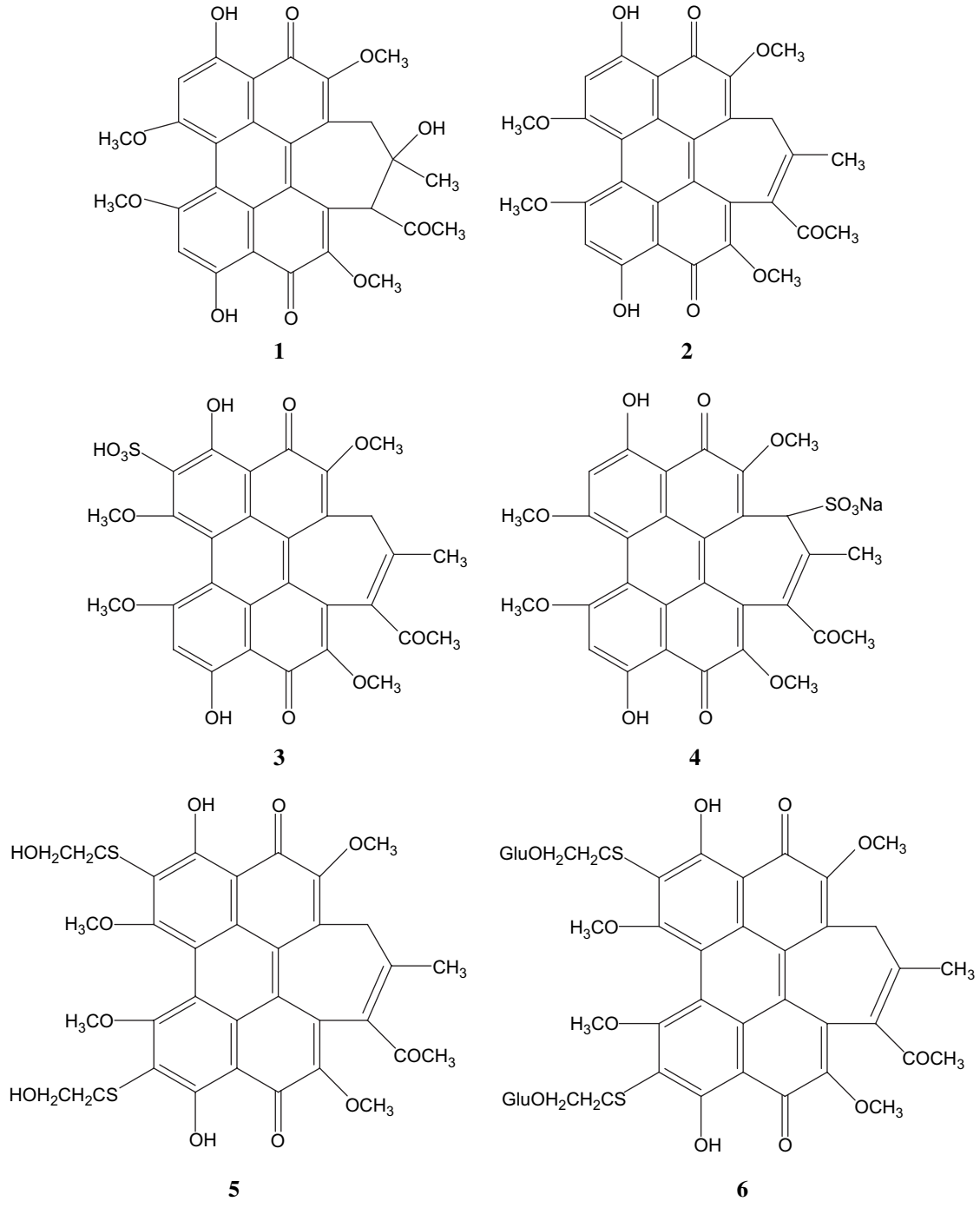


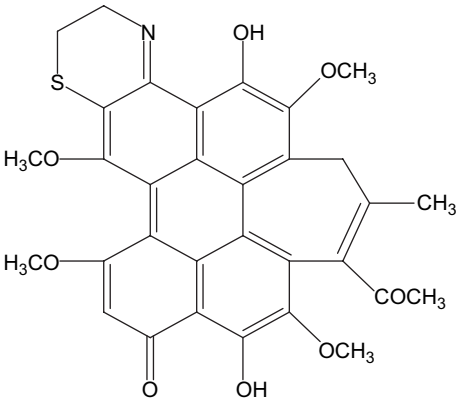

Fig. 2. Structure of 20 hypocrellins studied.

According to Table 3, PC1 can be expressed through the following equation:

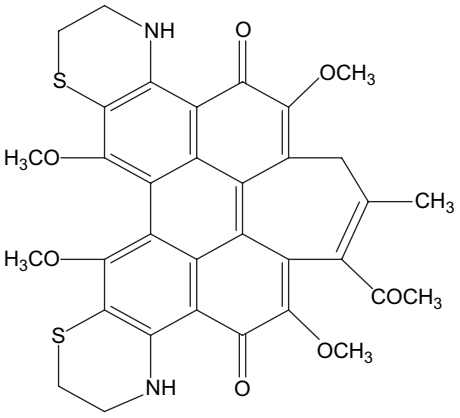
$$\mathrm{PC1} = 0.2128[\mu] + 0.4916[Q_3] - 0.2331[Q_4] - 0.3427[Q_9] + 0.735[Q_{10}].$$

From this equation, we can see that it needs to present small and negative values for dipole moment ($\mu$) and charges on atoms 3 and 10 ($Q_3$ and $Q_{10}$) along with large values for charges on atoms 4 and 9 ($Q_4$ and $Q_9$) for a hypocrellin derivative to be more active. These characteristics can be useful in the chemical structure modification of hypocrellin molecule with a higher phototoxicity against tumor cells.

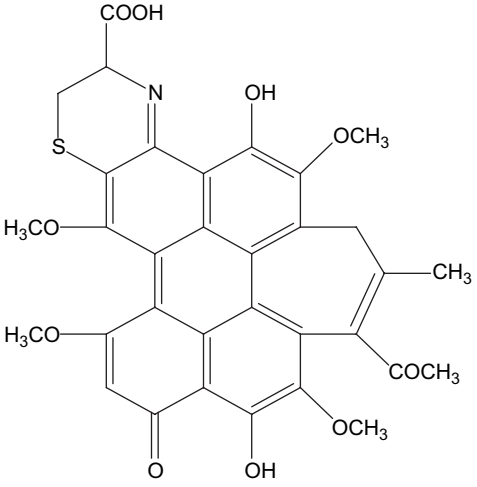
The variables responsible for the separation between higher active and lower active hypocrellins, i.e. dipole moment ($\mu$) and charges on atoms 3, 4, 9 and 10 ($Q_3$, $Q_4$, $Q_9$ and $Q_{10}$) are all electronic variables, therefore we can conclude that electronic effects have a very important role when one is trying to understand the phototoxicity of hypocrellins against tumor cells.



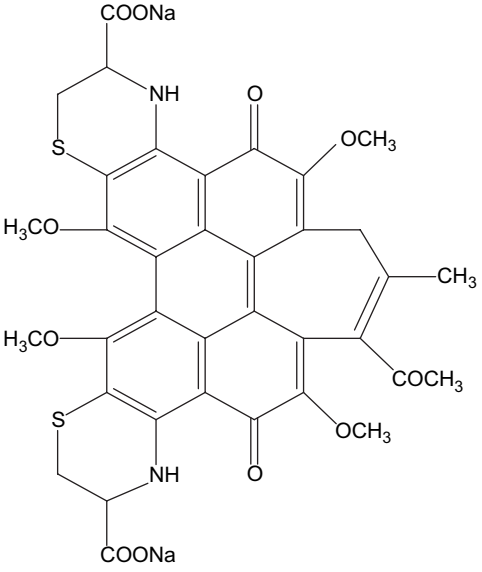

**7**



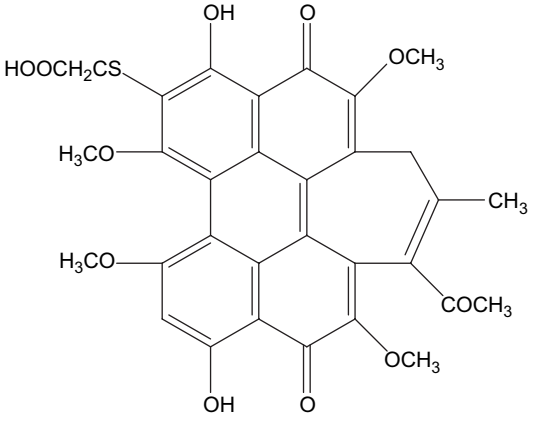

**8**



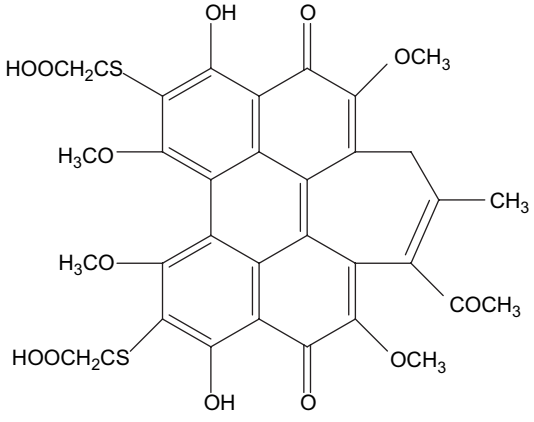

**9**




**10**




**11**




**12**

Fig. 2 (*continued*).

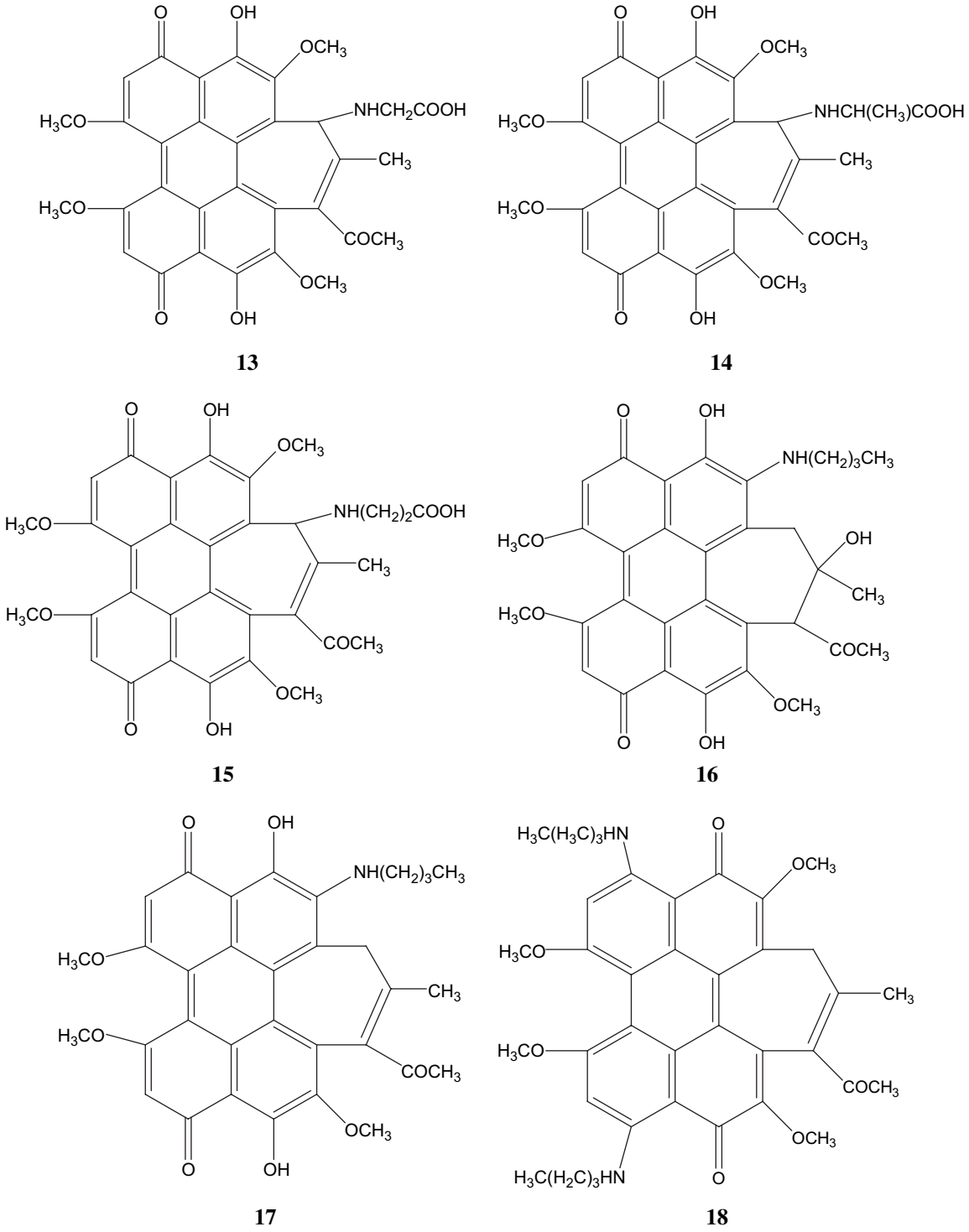
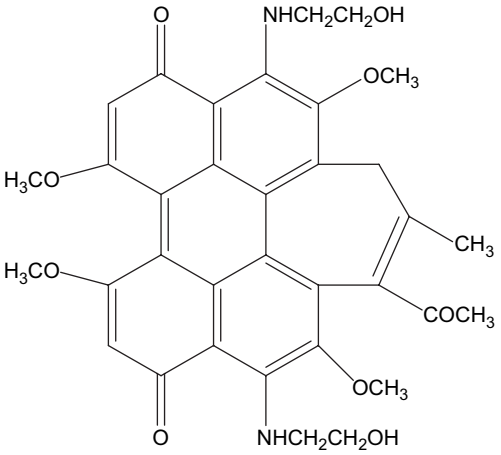



Fig. 2 (*continued*).

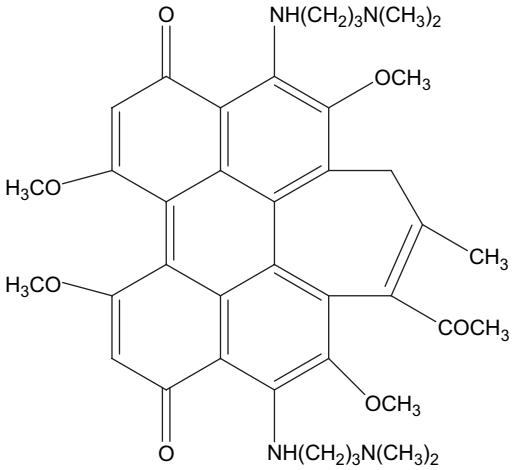
### 3.2. Hierarchical cluster analysis (HCA)

The HCA method is an excellent tool for preliminary data analysis and it is very useful for examining data sets for expected or unexpected clusters, including the presence of outliers. This technique examines the distances between the samples in a data set and represents this information as a two-dimensional plot called dendrogram [30]. It is informative to examine the dendrogram in conjunction with PCA results as they give similar information in different forms.

In the HCA analysis, each object (the 20 hypocrellins studied) is initially assumed to be a lone cluster and then the similarity matrix is analyzed. The most similar points are grouped forming one cluster and the process is repeated until all the points belong to one group [26].




**19**




**20**

Fig. 2 (*continued*).

The results obtained with the HCA analysis are similar to those obtained with PCA and are displayed in the dendrogram showed in Fig. 5. The dendrogram can be used to provide information on chemical behavior and verify the results obtained by PCA. In Fig. 5, the vertical lines represent the compounds and the horizontal lines represent the distances between pair of compounds, a compound and a group of compounds and between groups of compounds. From Fig. 5 we can see that 20

Table 1
Values of the five most important properties that classify the 20 hypocrellins studied and their phototoxicity $IC_{50}$

| Compound | $\mu$ | $Q_3$ | $Q_4$ | $Q_9$ | $Q_{10}$ | $IC_{50}$ ($\mu g\,mL^{-1}$) |
|---|---|---|---|---|---|---|
| **1** | 7.791 | 0.3676 | 0.2277 | 0.1945 | 0.2124 | 1.638–2.730[a] |
| **2** | 6.652 | 0.3572 | 0.2302 | 0.2454 | 0.3143 | 0.792–1.056[a] |
| **3** | 5.377 | 0.3863 | 0.2771 | 0.2383 | 0.3425 | 3.81[b] |
| **4** | 2.99 | 0.3464 | 0.2019 | 0.1869 | 0.2192 | >20[b] |
| **5** | 3.238 | 0.39 | 0.2252 | 0.2076 | 0.2459 | 1.58[b] |
| **6** | 4.822 | 0.3408 | 0.2373 | 0.2428 | 0.3454 | >20[b] |
| **7** | 4.912 | 0.1399 | 0.1245 | 0.3563 | −0.0037 | 0.09[b] |
| **8** | 6.909 | 0.3491 | 0.2311 | 0.2593 | 0.3413 | 0.27[b] |
| **9** | 5.292 | 0.1382 | 0.145 | 0.3908 | 0.1393 | 0.87[b] |
| **10** | 3.31 | 0.3442 | 0.224 | 0.2211 | 0.25 | >20[b] |
| **11** | 3.062 | 0.3465 | 0.2465 | 0.2353 | 0.3458 | 3.02[b] |
| **12** | 4.781 | 0.3903 | 0.2577 | 0.1651 | 0.2095 | >20[b] |
| **13** | 1.371 | 0.1444 | 0.3594 | 0.3601 | 0.1332 | 0.23[b] |
| **14** | 2.38 | 0.1596 | 0.3608 | 0.3531 | −0.0237 | 5.34[b] |
| **15** | 3.947 | 0.1692 | 0.3591 | 0.3495 | −0.0538 | 2.18[b] |
| **16** | 2.667 | 0.1165 | 0.362 | 0.3502 | −0.0385 | 0.147–0.588[c] |
| **17** | 3.553 | 0.1863 | 0.3687 | 0.3494 | −0.0492 | 0.142–0.284[d] |
| **18** | 9.925 | 0.3841 | 0.2041 | 0.1011 | 0.2309 | 0.128–0.383[a] |
| **19** | 3.996 | 0.1023 | 0.3531 | 0.3974 | 0.1692 | 0.092[a] |
| **20** | 3.381 | 0.2278 | 0.3633 | 0.3609 | 0.0494 | 0.342–1.027[a] |

[a] Phototoxicity against EMT6/Ed murine tumor cell [1].
[b] Phototoxicity against human oral cavity epithelial carcinoma KB cell [13,14].
[c] Phototoxicity against human gastric adenocarcinoma MGC803 cell [15].
[d] Phototoxicity against human cervix uteri tumor HeLa cell [16].

Table 2
Variances (eigenvalues) obtained for the first three principal components

| Component | Eigenvalue | Percentage (%) | Cumulative (%) |
|---|---|---|---|
| PC1 | 0.2835 | 73.358 | 73.358 |
| PC2 | 0.0456 | 11.807 | 85.165 |
| PC3 | 0.0342 | 8.844 | 94.009 |

Table 3
The loading vectors for the first three principal components

| Variable | PC1 | PC2 | PC3 |
|---|---|---|---|
| Dm | 0.2128 | −0.6421 | 0.1752 |
| C3 | 0.4916 | −0.1166 | −0.5319 |
| C4 | −0.2331 | 0.4483 | −0.5931 |
| C9 | −0.3427 | 0.2868 | 0.4709 |
| C10 | 0.7357 | 0.5393 | 0.3362 |

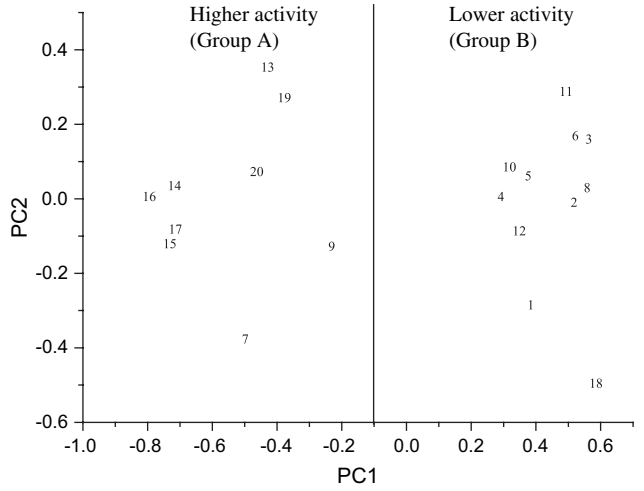



Fig. 3. Plot of the score vectors of the principal components (PC1 × PC2) for the 20 hypocrellins with phototoxicities against tumor cells. The PCA separates the compounds into two groups: higher activity (group A) and lower activity (group B).

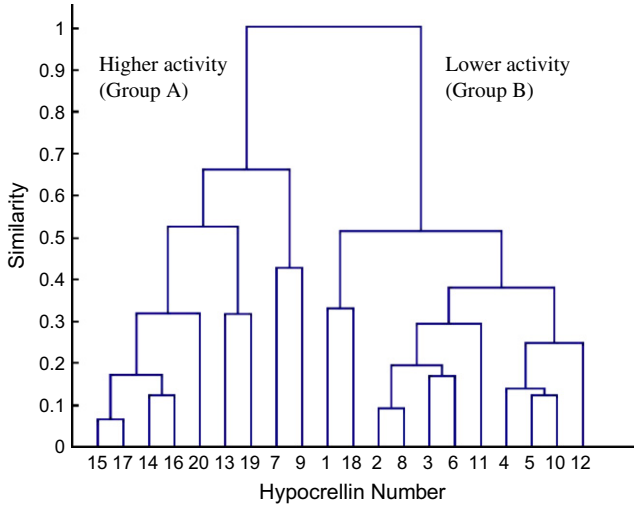



Fig. 5. Dendrogram obtained with hierarchical cluster analysis for the 20 hypocrellins with phototoxicities against tumor cells. The HCA classifies the compounds into two groups: higher activity (group A) and lower activity (group B).

hypocrellins are combined in two groups, group A and group B, according to the distance. The groups A and B in Fig. 5 correspond to the same groups A and B in Figs. 3 and 4 (PCA analysis). Both PCA and HCA methods classify the 20 hypocrellins studied into two groups. Based on the classification obtained with the PCA and HCA, we can say that $\mu$ (dipole moment) and $Q_3$, $Q_4$, $Q_9$ and $Q_{10}$ (charges on atoms 3, 4, 9 and 10) are the most important variables for the classification between the 20 hypocrellins with higher and lower phototoxicities against tumor cells.

## 4. Conclusion

PCA and HCA show that the 20 hypocrellins studied here can be classified into two groups: higher active (group A) and lower active (group B) according to their degree of phototoxicities against tumor cells. The variables, $\mu$ (dipole moment) and $Q_3$, $Q_4$, $Q_9$ and $Q_{10}$ (charges on atoms 3, 4, 9 and 10), are the most important ones for the classification between the molecules with higher and lower phototoxicities against tumor cells. The behavior of these five variables can be useful to design new hypocrellin derivative molecule with higher phototoxicities against tumor cells.

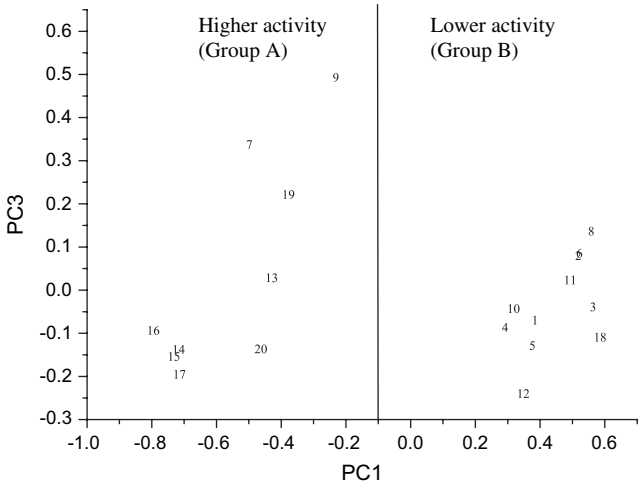



Fig. 4. Plot of the score vectors of the principal components (PC1 × PC3) for the 20 hypocrellins with phototoxicities against tumor cells. The PCA separates the compounds into two groups: higher activity (group A) and lower activity (group B).

## Acknowledgement

We gratefully acknowledge Education Bureau of Zhejiang province, China for financial support for this investigation (project no: 20051290).

## References

[1] Estey EP, Brown K, Diwu Z, Liu J, Lown JW, Miller GG, et al. Hypocrellins as photosensitizers for photodynamic therapy: a screening evaluation and pharmacokinetic study. Cancer Chemother Pharmacol 1996;37:343–50.

[2] Zhang J, Cao EH, Li JF, Zhang TC, Ma WJ. Photodynamic effects of hypocrellin A on three human malignant cell lines by inducing apoptotic cell death. J Photochem Photobiol B Biol 1998;43:106–11.

[3] Fehr MJ, Carpenter SL, Wannemuehler Y, Petrich JW. Roles of oxygen and photoinduced acidification in the light-dependent antiviral activity of hypocrellin A. Biochemistry 1995;34:15845–8.

[4] Deininger MH, Weinschenk T, Morgalla MH, Meyermann R, Schluesenera HJ. Release of regulators of angiogenesis following hypocrellin-A and -B photodynamic therapy of human brain tumor cells. Biochem Biophys Res Commun 2002;298:520–30.

[5] Zhang LZ, Tang GQ, Gao BW, Zhang GL. Spectroscopic studies on the excited-state properties of the light-induced antiviral drug hypocrellin A loaded in the mesoporous solid. Chem Phys Lett 2004;396:102–9.

[6] Hudson JB, Zhou J, Chen J, Harris L, Yip L, Towers GH. Hypocrellin, from *Hypocrella bambusae*, is phototoxic to human immunodeficiency virus. Photochem Photobiol 1994;60:253–5.

[7] Zhang L, Gu Y, Liu FG, Liu YY, Zhao XM, Zeng J, et al. Effect of photodynamic therapy on rat mesenteric microcirculation. Chin J Laser Med Surg 2003;12:69–75 [in Chinese].
[8] Wang X, Chen Y. Hypocrellin A – a new drug for photochemotherapy. Kexue Tongbao 1981;26:1040–2.
[9] Zang LY, Zhang ZY, Misra HP. EPR studies of trapped singlet oxygen ($^1O_2$) generated during photoirradiation of hypocrellin A. Photochem Photobiol 1990;52:677–83.
[10] He YY, An JY, Jiang LJ. pH effect on the spectroscopic behavior and photoinduced generation of semiquinone anion radical of hypocrellin B. Dyes Pigments 1999;41:79–87.
[11] Wang N, Zhang Z. Relationship between photosensitizing activities and chemical structure of hypocrellin A and B. J Photochem Photobiol B Biol 1992;14:202–17.
[12] Zhang Z, Wang N, Wang Q, Li M. EPR studies of singlet oxygen ($_1O^2$) and free radicals ($O_2^{\bullet-}$, $^{\bullet}OH$ and $HB^{\bullet-}$) generated during photosensitization of hypocrellin B. Free Radic Biol Med 1993;14:1–9.
[13] Jiang L, He Y. Photochemical, photophysical and photobiological properties of hypocrellins. Kexue Tongbao 2000;45:2019–33 [in Chinese].
[14] He YY, Liu HY, An JY, Han R, Jiang LJ. Photodynamic action of hypocrellin dyes: structure–activity relationships. Dyes Pigments 2000;44: 63–7.
[15] Zhang WG, Weng M, Pang SZ, Zhang MH, Yang HY, Zhao HX, et al. A novel photosensitizer, 2-butylamino-2-demethoxy-hypocrellin A (2-BA-2-DMHA) 1. Synthesis of 2-BA-2-DMHA and its phototoxicity to MGC803 cells. J Photochem Photobiol B Biol 1998;44:21–8.
[16] Yang H, Wu T, Zhang M, Zhang Z. A novel photosensitizer, 2-butylamino-2-demethoxy-hypocrellin B (2-BA-2-DMHB) – its photodynamic effects on HeLa cells: efficacy and apoptosis. Biochim Biophys Acta 2001;1540:22–31.
[17] Yu C, Xu S, Chen S, Zhang M, Shen T. Investigation of photobleaching of hypocrellin B in non-polar organic solvent and in liposome suspension. J Photochem Photobiol B Biol 2002;68:73–8.
[18] Zhao Y, Xie J, Zhao J. The photosensitization activity of a water-soluble carboxylate of hypocrellin B. Dyes Pigments 2005;66:61–8.
[19] Xu S, Chen S, Zhang M, Shen T, Zhao Y, Liu Z, et al. Butylamino-demethoxy-hypocrellins and photodynamic therapy decreases human cancer in vitro and in vivo. Biochim Biophys Acta 2001;1537:222–32.
[20] Zhao Y, Zhao J. Preparation of a novel hypocrellin derivative and its photochemical, photophysical properties. Dyes Pigments 2004;63:175–9.
[21] Arroio A, Honório KM, Silva ABF. A theoretical study on the analgesic activity of cannabinoid compounds. J Mol Struct (Theochem) 2004;709:223–9.
[22] Alves CN, Pinheiro JC, Camargo AJ, Souza AJ, Carvalho RB, Silva ABF. A quantum chemical and statistical study of flavonoid compounds with anti-HIV activity. J Mol Struct (Theochem) 1999;491:123–31.
[23] Camargo AJ, Mercadante R, Honório KM, Alves CN, Silva ABF. A structure–activity relationship (SAR) study of synthetic neolignans and related compounds with biological activity against *Escherichia coli*. J Mol Struct (Theochem) 2002;583:105–16.
[24] Dewar MJS, Zoebisch EG, Healy EF, Stewart JJP. AM1: a new general purpose quantum mechanical model. J Am Chem Soc 1985;107:3902–9.
[25] Fukunaga K, Koontz WLG. A criterion and an algorithm for grouping data. IEEE Trans Comput 1970;C-19:917–23.
[26] Kolwalski BR, Bender CF. Pattern recognition, a powerful approach to interpreting chemical data. J Am Chem Soc 1972;94:5632–9.
[27] Allinger NL. Conformational analysis 130. MM2. A hydrocarbon force field utilizing V1 and V2 torsional terms. J Am Chem Soc 1977;99: 8127–234.
[28] Cambridge Corporation. ChemOffice 6.0. Cambridge, MA: Cambridge Corp; 2001.
[29] Little J. MATLAB. Natick, MA: Mathworks Inc; 1998.
[30] Souza J, Santos RH, Ferreira MMC, Molfetta FA, Camargo AJ, Honório KM, et al. A quantum chemical and statistical study of flavonoid compounds (flavones) with anti-HIV activity. Eur J Med Chem 2003;38:929–38.