

主成分分析和聚类分析法用于微晶纤维素分类

张南, 赵国巍, 钟绍金, 廖正根*, 梁新丽, 王春柳, 曹运朝
(江西中医学院现代中药制剂教育部重点实验室, 南昌 330004)

[摘要] **目的:**研究主成分分析法及聚类分析法在微晶纤维素分类中的应用。**方法:**测定了来源于 5 个不同厂家的 12 种 PH101 和 PH102 型微晶纤维素的粉体学等物理性质,采用主成分分析法及聚类分析法对微晶纤维素的分类进行研究。**结果:**主成分分析和聚类分析的分析结果一致,主成分分析揭示了 PH101 和 PH102 型微晶纤维素分类的依据。**结论:**主成分分析法和聚类分析法可以用于微晶纤维素的分类研究,其结果与传统生产应用中微晶纤维素分类一致。

[关键词] 微晶纤维素; 物理性质; 主成分分析; 聚类分析; 分类研究

[中图分类号] R283.6 **[文献标识码]** A **[文章编号]** 1005-9903(2011)19-0004-05

Classification of Microcrystalline Cellulose (MCC) by Principal Component Analysis and Cluster Analysis Method

ZHANG Nan, ZHAO Guo-wei, ZHONG Shao-jin, LIAO Zheng-gen*,
LIANG Xin-li, WANG Chun-liu, CAO Yun-chao

(Key Laboratory of Modern Preparation of Traditional Chinese Medicine, Ministry of Education,
Jiangxi University of Traditional Chinese Medicine, Nanchang 330004, China)

[Abstract] **Objective:** To study the application in classification of MCC by principal component analysis (PCA) and cluster analysis method. **Method:** Determined powder property and other physical properties of 12 kinds of PH101 and PH102 MCC types which obtained from 5 different manufacturers, then PCA and cluster analysis method were used to investigate the classification of MCC. **Result:** The same result was obtained from PCA and cluster analysis method, it revealed that the basis on classification of PH101 and PH102 MCC types by PCA. **Conclusion:** PCA and cluster analysis method can be applied to study the classification of MCC with the result was the same as classification of MCC in traditional production application.

[Key words] MCC; physical property; PCA; cluster analysis; classification study

微晶纤维素(MCC)是天然纤维素经稀酸水解至极限聚合度(LODP)的可自由流动的极细微的短棒状或粉末状多孔状颗粒,颜色为白色或近白色,无臭、无味,粒径大小一般在 20 ~ 80 μm ,极限聚合度(LODP)在 15 ~ 375;不具纤维性而流动性极强。

MCC 被广泛应用于医药、食品、化妆品以及轻化工行业。由于其优越的流动性和压缩成形性,被广泛地作为片剂的添加剂和直接压片辅料。

目前 MCC 主要是通过酸水解工艺制备,其产品来源和种类很多,性质因其来源和批次不同而有差异,从而影响到辅料的筛选及使用;在生产应用中对微晶纤维素辅料分类的标准并不明确,通常都是经验性的指标,而且各个厂家产品分类标准之间存在差异,缺乏统一的定性和定量化的指标。在众多微晶纤维素型号产品中,PH101 和 PH102 型是应用最为广泛的 2 种,具有优越的压缩成形性,其中 PH101 粒径较小,特别适合湿法制粒,PH102 粒径增大,流

[收稿日期] 20110429(004)

[基金项目] “重大新药创制”科技重大专项(2009ZX09310-005);国家自然科学基金项目(81160522)

[第一作者] 张南,硕士研究生,从事药物新剂型与新技术的研究,Tel:0791-7119190,E-mail:onlyzn@126.com

[通讯作者] *廖正根,博士,教授,从事药物新剂型与新技术研究,Tel:0791-7119190,E-mail:lyzlyg@163.com

动性明显改善,常用于直接压片,PH101与PH102兼具有黏合性,助流性,崩解性。PH102较PH101粒径增大,流动性明显改善,是传统两者分类的依据,然而粒径和其流动性差异是否是其两者的最主要差异,现行缺乏(数学统计)科学理论上的指导及论证。本文考察了来源于5个不同厂家的12种PH101和PH102型号MCC的粉体学等众多物理性质,应用主成分分析法对微晶纤维素进行分类,寻找PH101与PH102产品之间的主要物性差异,同时采用聚类分析法对主成分分析的结果做进一步的验证分析。

1 材料

ERWEKA-TRD8型多功能压片机(德国艾维卡公司),MA-45型卤素水分测定仪(德国赛多利斯公司),TriStar 3000型全自动比表面积及孔隙度测定仪(美国麦克公司),GZX-9140 MBE数显鼓风干燥箱(上海博讯实业有限公司医疗设备厂),Mastersizer2000型激光粒度测定仪(英国马尔文仪器有限公司),乌氏黏度计(上海良品玻璃仪器厂),动态接触角测定仪(德国GmbH设备有限公司),BT-1000型粉末综合性测定仪(丹东市百特仪器有限公司)。

12种MCC购自5个不同的厂家,分别为安徽山河药用辅料有限公司(A),山东聊城阿华制药有限公司(B),德国JRS公司(C),日本旭化成株式会社(D),美国国际特品有限公司ISP(E)。PH101型MCC分别用A-1,B-1(棉浆型),B-3(溶解浆型),C-1,D-1,E-1表示,PH102型MCC分别用A-2,B-2(棉浆型),B-4(溶解浆型),C-2,D-2,E-2表示。

2 方法及结果

2.1 粒度分布的测定 所有MCC样本粉体采用Mastersizer 2000型激光粒径测定仪测定粒径及其分布。

2.2 比表面积的测定 分别取待测粉体约1g于80℃流通氮气内干燥至恒重,置入比表面积及孔隙度测定仪中,利用BET法测定样品的比表面积和孔容。

2.3 休止角的测定 采用粉末综合性测定仪结合固定圆锥底法测定:取一定量的待测粉末,在一定振动频率(约100Hz)下使粉末通过漏斗均匀流出,直到获得最高的圆锥体为止,测量圆锥体斜面与平面的夹角即得,每次测量重复3次,取其平均值。

2.4 松密度和振实密度的测定 采用粉末综合性测定仪测定,取一已称重的量筒,用漏斗匀速注入一定体积的待测粉末,精确称重,并准确记录粉末体积数,重复测定5次,求出松密度。将上述装有一定体积粉末的钢筒,以250 Hz·min⁻¹振幅振动300次,记录振实后体积,重复测定5次,计出摇实密度,并根据公式算出压缩指数。 $C = (df - d_0) / df \times 100\%$,式中 df 为摇实密度, d_0 为松密度,压缩度反映粉末的可压性和填充性能,压缩度大的粉末可压性好,但流动性差;压缩度小的粉末填充性和流动性好,但可压性差。

2.5 水分测定 采用MA-45型卤素水分测定仪测定水分含量。取各样品约1g,分别平铺于红外水分测定仪中已恒重的托盘,105℃加热至粉末质量不再改变(维持时间大于30s),读取水分蒸发百分比(即粉末含水量)。

2.6 聚合度和分子量 取样品约1.3g,精密称定,置125mL带塞锥形瓶中,精密加水 and 1.0 mol·L⁻¹氢氧化乙二胺铜溶液各25mL,加塞,于电磁搅拌器上搅拌30min使完全溶解,转移该溶液适量至乌氏黏度计(毛细管内径0.84mm,已校正)中,在25℃水浴中平衡至少5min,记录溶液流经黏度计上下2个刻度时的时间 t_1 (以秒计),按如下公式计算溶液的运动黏度(V_1)。

$$V_1 = t_1(K_1), K_1 \text{ 为黏度计常数}$$

取适量1.0 mol·L⁻¹氢氧化乙二胺铜溶液与等量水混合,用乌氏黏度计毛细管内径0.63mm,已校正)测得流出时间 t_2 ,用以下公式计算溶剂的运动黏度(V_2)

$$V_2 = t_2(K_2), K_2 \text{ 为黏度计常数}$$

用以下公式计算供试品的相对黏度 η_{rel}

$$\eta_{rel} = \frac{V_1}{V_2} = \frac{t_1 \cdot K_1}{t_2 \cdot K_2}$$

根据计算得的相对粘度 η_{rel} 值,查特性黏度表(BP1998, P289),得特性黏度 $[\eta]_c$,用以下公式计算聚合度 P

$$P = \frac{95 \cdot [\eta]_c}{m[(100 - b)/100]}$$

其中 m 为供试品取样量(g), b 为供试品干燥失重率。乌氏黏度计毛细管内径0.84mm,黏度计常数0.03918 mm²·s⁻²;毛细管内径0.63mm,黏度计常数0.01254 mm²·s⁻²。

2.7 膨胀度^[1] 膨胀度是药品膨胀性质的指标,系指按干燥品计算,每1g药品在水或其他规定的溶

剂中,在一定的时间与温度条件下膨胀后所占有的体积。将样品置于 60 ℃ 真空干燥干燥至恒重,取 25 mL 膨胀管,称干燥样品约 2 g,加水 25 mL,密塞,振摇,静置。开始 1 h 内每 10 min 振摇 1 次,后静置 1 h,读取药物膨胀后的体积 V ,再静置 1 h,如上读数连续 2 次,读数差并不超过 0.1 mL 为止。平行测定 3 份。每个样品测定 3 份,求其平均数,按下式计算,即得供试品的膨胀度(准确至 0.1)。

$$S = \frac{V}{W}$$

S 为膨胀度, V 为药物膨胀后体积, W 为样品按干燥品计算的质量。

2.8 引湿性 药物的引湿性是指在一定温度及湿度条件下该物质吸收水分能力或程度的特性。实验结果可作为选择适宜的药品包装和贮存条件的参考。具体试验方法按照 2010 年版《中国药典》一部附录 XIXJ 药物引湿性实验指导原则进行。取干燥的具塞玻璃称量瓶于 25 ℃ 恒温干燥器(下部放置氯

化铵饱和溶液)平衡 24 h,精密称定质量(m_1),取样品平铺于上述称量瓶中精密称定质量(m_2),样品厚度约 1 mm,将称量瓶敞口,与瓶盖同置于上述恒温恒湿下,24 h 后称量吸湿后的瓶子和粉末的总质量(m_3),计算吸湿增重率 = $\frac{m_3 - m_2}{m_2 - m_1} \times 100\%$ 。

2.9 接触角 接触角的测量 基于 Washburn 公式,使液体通过毛细上升进入一种多孔介质,通过固气界面到固液界面的转换来实现测量^[2-3]。动态接触角测定仪通过高度法测定粉体的润湿接触角。接触角的测量分 2 次测试,第 1 次是检测毛细血管常数 C ,第 2 次是对所测液体进行接触角分析。

2.10 数据及分析 测量得到的样品粉体学等物性数据见表 1,对其结果进行主成分分析(表 2,3)和聚类分析,分析采用的统计分析软件 SIMCA-P 12.0 进行主成分分析和 SPSS 18.0 进行聚类分析。

表 1 MCC 样品和物性测定

样品	型号	松密度 /g·mL ⁻¹	振实密度 /g·mL ⁻¹	压缩度/%	含水量/%	聚合度	平均相对分子质量	膨胀度 /mL·g ⁻¹	接触角/°
A-1	PH101	0.319 ± 0.005	0.469 ± 0.008	32 ± 0.01	4.45 ± 0.07	215.2 ± 0.4	34 855 ± 69	3.88 ± 0.2	72.5 ± 1.1
A-2	PH102	0.352 ± 0.01	0.518 ± 0.014	32 ± 0.01	4.72 ± 0.16	189.5 ± 0.6	30 701 ± 105	3.56 ± 0.08	72.4 ± 0.8
B-1	PH101	0.415 ± 0.01	0.611 ± 0.014	32 ± 0	3.33 ± 0.17	198.2 ± 0.6	32 107 ± 104	3.04 ± 0.06	78.0 ± 0.9
B-2	PH102	0.357 ± 0.008	0.525 ± 0.011	32 ± 0.01	4.05 ± 0.27	202.5 ± 0.1	32 803 ± 14	3.10 ± 0.03	82.1 ± 0.4
B-3	PH101	0.319 ± 0.002	0.469 ± 0.002	32 ± 0	4.15 ± 0.18	266.7 ± 0.9	43 210 ± 142	3.78 ± 0.13	70.9 ± 2.0
B-4	PH102	0.314 ± 0.001	0.461 ± 0.001	32 ± 0	4.58 ± 0.07	221.7 ± 0.6	35 908 ± 101	3.79 ± 0.11	76.8 ± 2.2
C-1	PH101	0.334 ± 0.002	0.491 ± 0.004	32 ± 0.01	3.80 ± 0.10	224.7 ± 0.6	36 402 ± 91	3.38 ± 0.07	70.5 ± 2.3
C-2	PH102	0.336 ± 0.002	0.494 ± 0.003	32 ± 0.01	4.10 ± 0.09	225.2 ± 0.4	36 484 ± 68	3.35 ± 0.15	72.2 ± 2.4
D-1	PH101	0.334 ± 0.001	0.491 ± 0.002	32 ± 0.01	4.76 ± 0.14	227.3 ± 0.2	36 824 ± 35	3.42 ± 0.13	72.9 ± 1.2
D-2	PH102	0.323 ± 0.003	0.440 ± 0.008	26.67 ± 1.16	4.32 ± 0.15	233.0 ± 0.4	37 748 ± 57	3.54 ± 0.05	78.4 ± 1.4
E-1	PH101	0.303 ± 0.009	0.499 ± 0.020	39.19 ± 1.05	4.56 ± 0.12	239.2 ± 0.5	38 748 ± 75	3.18 ± 0.03	67.6 ± 0.4
E-2	PH102	0.314 ± 0.003	0.505 ± 0.004	37.86 ± 0.23	4.67 ± 0.10	238.8 ± 0.5	38 680 ± 80	3.24 ± 0.09	76.2 ± 0.8

样品	型号	表面积平均 粒径/μm	体积平均 粒径/μm	d_{50} /μm	d_{90} /μm	比表面积 /m ² ·g ⁻¹	孔容/cm ³ ·g ⁻¹	引湿性/%	流动性/°
A-1	PH101	39.344	78.477	64.108	160.549	1.053 8 ± 0.036 0	0.004 1 ± 0.000 1	7.35 ± 0.13	41.95 ± 1.55
A-2	PH102	74.476	125.946	113.189	228.802	1.195 2 ± 0.019 4	0.004 8 ± 0.000 4	6.33 ± 0.37	32.3 ± 1.21
B-1	PH101	35.276	67.523	55.393	135.657	0.662 6 ± 0.005 2	0.003 7 ± 0.000 2	5.58 ± 0.05	38.6 ± 1.85
B-2	PH102	49.499	90.532	77.274	180.581	0.597 1 ± 0.008 8	0.003 3 ± 0.000 0	5.25 ± 0.05	33.63 ± 1.09
B-3	PH101	47.551	82.517	71.515	157.144	1.055 2 ± 0.003 1	0.005 1 ± 0.000 2	6.94 ± 0.13	38.6 ± 1.85
B-4	PH102	92.587	146.506	138.021	249.838	0.951 9 ± 0.007 5	0.004 3 ± 0.000 1	5.69 ± 0.05	32.23 ± 2.31
C-1	PH101	39.002	71.285	60.061	142.526	1.264 9 ± 0.011 0	0.006 9 ± 0.000 2	6.13 ± 0.49	40.9 ± 0.87
C-2	PH102	70.896	136.03	126.825	252.903	0.980 0 ± 0.007 7	0.005 9 ± 0.000 2	6.55 ± 0.23	35.5 ± 1.04
D-1	PH101	44.207	70.33	62.212	130.047	0.999 7 ± 0.012 9	0.004 9 ± 0.000 1	6.40 ± 0.25	35.5 ± 1.04
D-2	PH102	73.696	132.143	122.035	240.551	1.160 3 ± 0.005 0	0.005 2 ± 0.000 1	7.30 ± 0.42	35.63 ± 1.1
E-1	PH101	35.413	74.528	58.726	154.137	1.481 3 ± 0.080 3	0.007 0 ± 0.000 4	7.31 ± 0.17	32.3 ± 1.21
E-2	PH102	64.007	131.678	117.661	255.996	1.339 5 ± 0.109 7	0.005 7 ± 0.000 3	7.0 ± 0.07	38.6 ± 1.85

表2 主成分分析初始统计值对原有变量总体描述情况

主成分	特征值	方差贡献率/%	累计方差贡献率/%
1	4.43	0.369	0.369
2	3.84	0.32	0.689
3	1.32	0.11	0.799
4	0.842	0.070 2	0.869

表3 主成分负荷矩阵

成分	主成分 1	主成分 2	主成分 3	主成分 4
比表面积	0.349 981	-0.064 314 3	0.261 357	0.166 307
流动性	0.174 838	-0.355 747	0.113 334	-0.028 681 1
松密度	-0.379 147	-0.068 952 8	0.064 073 3	-0.062 148 7
振实密度	-0.311 786	-0.166 311	0.330 782	-0.051 082 7
压缩度	0.135 181	-0.187 335	0.502 118	0.037 1117
孔容	0.293 645	-0.134 036	0.288 251	-0.134 861
膨胀度	0.185 277	0.159 326	-0.482 301	0.333 722
含水量	0.248 413	0.162 732	0.080 897 9	0.471 468
接触角	-0.265 456	0.210 977	-0.061 104 5	-0.277 264
聚合度	0.312 959	-0.082 785	-0.201 318	-0.487 582
相对分子质量	0.312 959	-0.082 768 6	-0.201 311	-0.487 615
表面积平均粒径	0.058 709 3	0.426 394	0.085 426 3	-0.033 385 9
体积平均粒径	0.087 463 2	0.408 48	0.200 546	-0.115 47
$d(0.5)$	0.082 848 4	0.413 47	0.171 662	-0.129 011
$d(0.9)$	0.094 396 6	0.386 919	0.257 235	-0.141 388
引湿性	0.334 992	-0.066 490 7	-0.054 894	0.091 477

3 讨论

主成分分析法广泛用于实验数据的统计分析,进行数据降维、变量提取与压缩、确定组分数、分类和聚类,以及与其他方法联用进行数据处理^[4-5]。在主成分分析中,数据的选择和数据的预处理是主成分分析成功与否的一个重要因素。标准化处理会保持原始指标数值的相对稳定性,本文应用 SIMCA-P 12.0 统计软件默认对原始数据进行标准化处理后进行主成分分析^[6]。

主成分的特征根及贡献率是选择主成分的依据,表1描述了主成分分析初始特征值对原有变量总体描述情况。从表1中可以看到总方差的86.9%的贡献来自前4个因子,即一个4因子模型解释了试验数据的86.9%,而主成分1与主成分2这2个分量已包含样本的较大部分信息(累计方差贡献率为68.9%)。表2中因子负荷矩阵反映了各指标对主成分贡献的大小,即各指标对主成分的解释作用,

因子负荷的符号表示各指标对改变主成分值的增减效果,由表中数据可知,对第二主成分起主要解释作用的是粒径指标和流动性指标(休止角),其贡献值显著大于其他各指标且符号相反,这也说明粒径大小与样本的流动性成负相关。图1展示了所有样品主成分分析中第一主成分和第二主成分得分的散点图,结果显示在95%的置信区间内没有特异值。样品 PH101 与 PH102 以 t_2 轴(第二主成分)很好的分开。而结合表3中表面积平均粒径,体积平均粒径, $d(0.5)$, $d(0.9)$ 及流动性对第二主成分的贡献最大,很好地说明了 MCCPH101 与 PH102 2 种型号的样品是以粒径和流动性为指标进行分类的,且 PH102 样品在得分图 t_2 轴上值均为正,PH101 则均为负,说明 PH102 的粒径大于 PH101 而流动性指标休止角小于 PH101,即流动性 PH102 也好于 PH101,这与传统的分类方法基本一致。

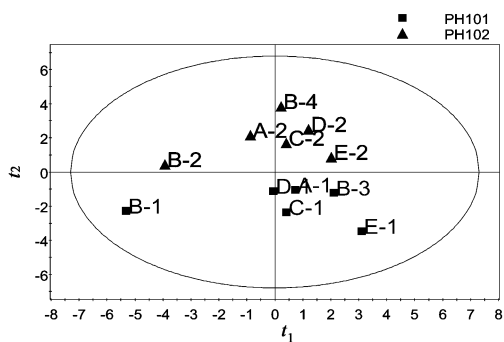


图1 PH101 和 PH102 型 MCC 样品主成分得分图

(t_1 为第一主成分, t_2 为第二主成分)

为了进一步验证主成分分析法对 MCC 样品分类的可行性,增加样品分类结果的可信度与科学性,本文再采用聚类分析法对 12 个 MCC 样品进行聚类研究。聚类分析作为一种探索性的分类方法,可以将一组数据按照本身的内在规律较合理地分为几类,大大缩小了以往全凭主观判断所造成的误差,使数据分析结果更具客观性^[7]。文中采用 SPSS 18.0 聚类分析程序对 12 个微晶纤维素样品进行聚类分组。在方法上采用欧氏距离测量,每两样本间用 Average linkage 法连结,12 个微晶纤维素样品用粒径及流动性作为聚类变量,可分成 3 大类,聚类结果表明,PH101 型号的 MCC 样品聚为一类,山东聊城阿华(棉浆)生产的 PH102 样品单独为一类,且与其他厂家的 PH102 距离接近。同时,通过聚类分析我们可以清楚的知道哪些厂家的产品性质(或说性

超临界 CO₂ 萃取云南香樟叶中右旋龙脑的工艺优选

缪菊连¹, 黄照昌^{2*}, 李红艳²

(1. 大理学院, 云南 大理 671000; 2. 大理药业, 云南 大理 671000)

[摘要] 目的:探讨超临界 CO₂ 萃取云南香樟叶中右旋龙脑的工艺条件。方法:采用气相色谱法测定右旋龙脑含量,设计 L₉(3⁴) 正交试验考察萃取压力、温度、流量,单因素考察时间对右旋龙脑萃取率的影响。结果:优选的萃取条件为压力 30 MPa,温度 45 ℃,二氧化碳流量 20 L·h⁻¹,萃取时间 90 min。结论:优选的最佳工艺萃取率高,是具有相当发展潜力的提取分离方法。

[关键词] 超临界 CO₂ 流体萃取法;香樟;右旋龙脑

[中图分类号] R283.6 **[文献标识码]** A **[文章编号]** 1005-9903(2011)19-0008-03

[收稿日期] 20110602(005)

[基金项目] 校企联合研究开发项目(20090426)

[第一作者] 缪菊连, 硕士, 讲师, 从事天然药物的开发与利用研究, Tel:15125282836, E-mail:sun007119@sina.com

[通讯作者] * 黄照昌, 学士, 工程师, 从事药物制剂研究, Tel:15125022320, E-mail:hzc@sina.com

能)最相接近,比如从图 2 的树状图可知,山东聊城阿华制药有限公司生产的棉浆型 PH101 与德国 JRS 公司和日本旭化成株式会社生产的同种型号产品十分接近,而安徽山河药用辅料有限公司生产的 PH101 则与美国国际特品有限公司 ISP 的同型号产品相差无几,从中我们可以看到,随着技术水平和生产设施水平的提高,现今国产辅料和进口辅料的差距越来越小,这为药学工作者和医药生产企业更好地选择质优价廉的辅料提供了科学的依据。

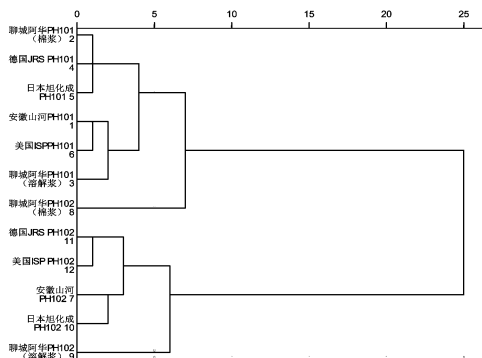


图 2 聚类分析树状图

[参考文献]

- [1] 中国药典.一部[S].2010:附录 56.
- [2] Bachmann J, Woche S K, Goebel M O, et al. Extended methodology for determining wetting properties of porous media[J]. Water Resour Res, 2003, 39:1353.
- [3] Goebel M O, Bachmann J, Woche S K, et al. Water potential and aggregate size effects on contact angle and surface energy[J]. Soil Sci Soc Am J, 2004, 68:383.
- [4] 陈军辉, 谢明勇, 王远兴, 等. 主成分分析法用于西洋参样品分类研究[J]. 天然产物研究与开发, 2006, 18(2):193.
- [5] 宋小妹, 杨新杰, 王薇, 等. 珠子参的 HPLC 指纹图谱及模式识别[J]. 中国实验方剂学, 2011, 17(11):59.
- [6] Rahul V H, Ingunn T, Annette B B. Multivariate analysis of relationships between material properties, process parameters and tablet tensile strength for a-lactose monohydrates[J]. Eur J Pharm Biopharm, 2009, 73:424.
- [7] 陈军辉, 谢明勇, 傅博强, 等. 西洋参中无机元素的主成分分析和聚类分析[J]. 光谱学与光谱分析, 2006, 26(7):1326.

[责任编辑 全燕]