

密级：

论文编号：

# 中国农业科学院

# 学位论文

## 植酸酶产品近红外指纹图谱库的构建 与应用研究

**Research of the Construction and Application of Phytase Fingerprinting  
Database by Near Infrared Spectroscopy**

硕士研究生：杨海锋

指导教师：秦玉昌 研究员

申请学位类别：农学硕士

专业：动物营养与饲料科学

研究方向：饲料质量评价技术

培养单位：中国农业科学院研究生院

中国农业科学院饲料研究所

提交日期 2007 年 6 月

Secrecy:

No:

# **Chinese Academy of Agricultural Sciences**

## **Master Dissertation**

### **Research of the Construction and Application of Phytase Fingerprinting Database by Near Infrared Spectroscopy**

Ms. Candidate: Yang Hai-feng

Advisor: Professor Qin Yu-chang

Major: Animal Nutrition and Feed Science

Specialty: Feed Quality Evaluation Technology

Chinese Academy of Agricultural Sciences

June 2007

## 独 创 性 声 明

本人声明所呈交的论文是我个人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得中国农业科学院或其它教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

研究生签名：

时间： 年 月 日

## 关于论文使用授权的声明

本人完全了解中国农业科学院有关保留、使用学位论文的规定，即：中国农业科学院有权保留递交论文的复印件和磁盘，允许论文被查阅和借阅，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。同意中国农业科学院可以用不同方式在不同媒体上发表、传播学位论文的全部或部分内容。

(保密的学位论文在解密后应遵守此协议)

论文作者签名：

时间： 年 月 日

导师签名：

时间： 年 月 日

**中国农业科学院**  
**硕士学位论文评阅人、答辩委员会名单表**

论文题目		植酸酶产品近红外指纹图谱库的构建与应用研究				
论文作者		杨海锋	专业	动物营养与 饲料科学	研究方向	饲料质量安全 评价
指导教师		秦玉昌	培养单位(研究所)			饲料研究所
姓名		职称	硕(博) 导师	单 位	专业	签 名
评 阅 人	谯仕彦	教授	硕导□ 博导□	中国农业大学 动物科学院	动物营养	
	张萍	研究员	硕导□ 博导□	中国农业科学院 饲料所	动物营养	
答辩主席	韩鲁佳	教授	硕导□ 博导□	中国农业大学 工学院	农产品加工	
答 辩 委 员	杨曙明	研究员	硕导□ 博导□	中国农业科学院 质检所	检测技术	
	谯仕彦	教授	硕导□ 博导□	中国农业大学 动物科学院	动物营养	
	田吉顺	高级工程师	硕导□ 博导□	北京水产研究所	水产学	
	干小英	研究员	硕导□ 博导□	中国农业科学院 饲料所	化工合成	
会议记录(秘书)						
论文答辩时间地点		2007年06月17日上午 中国农业科学院饲料研究所				

## 摘要

伴随着生物学、发酵工程等基础学科和应用技术的飞速发展，具有高效、强特异性生理功能的生物酶制剂已成为新型饲料添加剂的研发焦点，其生产使用规模也不断扩大。针对生物制品安全风险具有隐蔽性强、潜伏期长和危害性大等特点，如何确保和提高饲用酶制剂的安全水平，以及如何科学利用饲用酶制剂显著提升饲料工业和畜牧业生产效益空间、推动饲料工业和畜牧业可持续健康发展的技术优势已成为亟待解决的问题。本论文主要收集来自同一生产企业的植酸酶产品，其中喷雾干燥植酸酶样本 134 个和吸附干燥植酸酶样本 121 个，在 950-1650nm 谱区范围内，对试验样本进行了近红外光谱扫描，通过构建植酸酶的近红外光谱指纹图谱库，以及利用其建立的植酸酶产品真伪鉴别模型、植酸酶酶活和水分预测模型，研究探索将近红外光谱指纹图谱库引入植酸酶质量安全管理，发现其不仅利于生产者提高产品质量的稳定性和保护产品品牌，而且利于管理者建立高效的植酸酶可追溯管理体系，确保提高市场植酸酶整体安全水平。同时，本论文还利用高效液相色谱、差示扫描量热仪，以及设计开展植酸酶喷雾干燥实验、恒温恒湿储藏实验，研究分析了不同植酸酶产品的成分组成和功能特性，进一步探讨了植酸酶的近红外光谱信息与监控植酸酶质量安全水平和鉴别区分植酸酶产品的关键指标之间的关系。具体研究结果如下：

- (1) 研究建立了能够清晰区分不同生产企业、不同产品形式植酸酶产品的近红外光谱指纹图谱库，为植酸酶产品的真伪鉴别和品牌保护提供科学合理的技术方法；
- (2) 利用主成分分析定性判别程序，在 950-1650nm 范围内提取来自不同生产企业，不同形态的植酸酶产品的漫反射近红外光谱的主成分，取主成分 1 和主成分 2 作二维图。结果表明，在不同的模式中，前两个主成分对信息量的累积贡献率均在 90% 以上；相同生产企业，相同形态的植酸酶产品均能很好的聚为一类；采用主成分分析（PCA）所建立的喷雾干燥和吸附干燥植酸酶定性判别模型对 15 个未知样品进行判别，准确率均达到 100%，说明应用主成分分析所建立的定性分析模型可用来定性判别植酸酶产品的真伪，通过马氏距离法结合设定相应的阈值，还可以对其产品质量的好坏和稳定性作进一步的判别。
- (3) 在最优的波数范围内，选择 MSC 做散射校正处理、一阶导数和平滑组合为光谱预处理方法，采用内部交互验证法验证，根据定标集交互验证的标准差 RMSECV 值，确定了最佳主因子数，采用偏最小二乘法回归的方法（PLS-1）从而建立了植酸酶酶活的 NIRS 定量分析模型。植酸酶酶活的化学分析值与 NIRS 定标模型预测值之间的决定系数  $R^2$  达到 0.90 以上，相对标准差 RSD 均小于 10%，相对分析误差 RPD 大于 3。喷雾干燥植酸酶的决定系数  $R^2$  为 0.937，相对标准差 RSD 为 5.23%，相对分析误差 RPD 为 3.64；吸附干燥植酸酶的决定系数  $R^2$  为 0.926，相对标准差 RSD 为 6.21%，相对分析误差 RPD 为 3.45。结果表明，利用近红外光谱分析技术能够准确地检测喷雾干燥植酸酶和吸附干燥植酸酶酶活含量。
- (4) 建立了快速跟踪喷雾干燥加工、储藏过程中植酸酶质量变化的近红外光谱分析方法。

**关键字** 植酸酶，近红外光谱技术，指纹图谱库，定量测定，定性判别

## Abstract

With the rapid development of biology and fermentation engineering, the biotechnological enzymes which possessing high efficiency and specifically physiological function have become a research focus of new feed additives, and the production and application of the biotechnological enzymes are growing continuously. Due to the risk of biotechnological products, which always have been with a long incubation period and the serious damage, how to ensure and improve the safety level of feed enzyme, and how to scientifically utilize the super-advantage of feed enzymes in feed industry, has become a serious problem. This paper had collected 134 spray drying phytase and 121 adsorption drying phytase samples from the same manufacturers, the samples were scanned at the near infrared reflectance spectroscopy (NIRS) region of 950-1650nm. With constructing phytase NIR fingerprint database, the product authenticity discrimination model, the phytase enzyme activity and water quantitative determination model, we exploited to introduce the NIR fingerprint database into the phytase quality and safety management, and we found that it not only benefited the producers to improve the quality stability and protect the product brands, but also helped managers to establish an efficient phytase traceability system. Meanwhile, the paper also utilized high performance liquid chromatography, differential scanning calorimetry, spray drying experiments, and constant temperature and humidity storage experiments to investigate the composition and function characteristics of different phytase products, to further explore the relationship between the near infrared spectral information and the key indicators of quality. Specific results were as follows:

- (1) We investigated and established the phytase product NIR fingerprint database which could clearly distinguish the phytase products from different manufacturers, different form. The phytase product NIR fingerprint database could provide scientific and rational methods for authenticity identification and brand protection.
- (2) With the Principal Component Analysis (PCA) to construct two-dimensional map, we extracted the near-infrared diffuse reflectance spectra principal components from the phytase of different manufacturers at the NIR range of 950-1650nm. The results showed that the different models, the cumulative contribution rate which reflecting the sample information was over 90% from the first two principal components. The same manufacturers, the same pattern of phytase products could gather into one group; an established NIR qualitative model used to discriminate 15 unknown samples, the accuracy rate was 100%, this showed that principal component analysis could be used to establish a qualitative model for evaluating the authenticity of phytase products. With Mahalanobis distance and corresponding threshold, it could also make further discrimination for their quality and stability.
- (3) Multiplicative Scatter Correction (MSC) was chosen as scatter correction, and the first derivative and smooth combination chosen as spectral pretreatment methods. In the optimal wave scope, using internal cross-validation, based on Root Mean Square Error of Cross Validation (RMSECV), the optimal factor was determined. The calibration models to predict the phytase activity and moisture content were developed using Partial Least Squares (PLS-1) technique. The coefficient of determination in calibration

( $R^2$ ) of moisture, activity in phytase were all over 0.90, the Relative Standard Deviation (RSD) all less than 10%, and the Relative Predictive Determination (RPD) all over 3; the  $R^2$  of spray drying phytase was 0.937, the RSD and RPD of spray drying phytase were 5.23% and 3.64 respectively; the  $R^2$  of adsorption drying phytase was 0.926, the RSD and RPD of adsorption drying phytase were 6.21% and 3.45 respectively. The results showed that NIRS analysis technique could be adopted to correctly measure the activity in spray drying phytase and adsorption drying phytase.

(4) We also exploited to establish a NIR method for fast-tracking the quality changing of the spray drying and storage process of phytase products.

**Key word:** Phytase    NIR Spectroscopy    Fingerprinting Database    Quantitative Determination  
Qualitative Discrimination

# 目 录

第一章 绪 论 .....	1
1.1 植酸酶在饲料工业中的地位 .....	1
1.2 近红外光谱分析技术及在饲料工业中的应用 .....	2
1.2.1 近红外光谱技术应用发展史 .....	2
1.2.2 近红外光谱分析技术的主要应用领域 .....	4
1.2.3 近红外光谱技术在饲料工业中的应用 .....	5
1.2.4 近红外光谱技术在监控生物发酵过程中的应用 .....	6
1.3 近红外指纹图谱库技术 .....	6
1.3.1 指纹图谱库的原理 .....	6
1.3.2 近红外指纹图谱在农产品和中药安全管理中的应用 .....	7
1.3.3 近红外指纹图谱与饲料产品注册管理 .....	7
1.3.4 近红外指纹图谱与饲用酶制剂质量安全管理 .....	8
1.4 课题的选题意义 .....	8
1.5 本研究的主要内容及预期目标 .....	8
第二章 植酸酶产品特性研究 .....	10
2.1 植酸酶蛋白组成特性研究 .....	10
2.1.1 概述 .....	10
2.1.2 SDS-PAGE分析植酸酶蛋白组成 .....	10
2.1.3 反相高效液相色谱分析植酸酶蛋白组成 .....	12
2.1.4 小结 .....	15
2.2 植酸酶变性温度特性研究 .....	16
2.2.1 概述 .....	16
2.2.2 差示扫描量热仪分析植酸酶热稳定性 .....	16
2.2.3 小结 .....	18
第三章 植酸酶产品近红外指纹图谱库的建立 .....	19
3.1 近红外光谱分析技术测定原理 .....	19
3.1.1 近红外光谱分析的基础 .....	19
3.1.2 近红外光谱中的化学计量学方法 .....	20
3.1.3 评价数学模型效果的指标 .....	23
3.1.4 奇异点的判别与处理 .....	24
3.1.5 近红外光谱分析技术流程 .....	25
3.2 植酸酶产品近红外指纹分析图谱的构建 .....	26
3.2.1 指纹图谱库的构建原则 .....	26
3.2.2 实验样本的收集 .....	26
3.2.3 近红外反射光谱的采集 .....	26

3.2.4 异常样本的剔除 .....	27
3.2.5 近红外光谱特征信息的提取 .....	28
3.2.6 植酸酶产品指纹图谱库的建立 .....	30
3.3 小结 .....	31
<b>第四章 植酸酶酶活定量分析模型的建立与优化.....</b>	<b>32</b>
4.1 实验材料、仪器与实验方法 .....	32
4.1.1 实验样品 .....	32
4.1.2 参考分析方法及仪器 .....	32
4.1.3 近红外光谱分析方法 .....	32
4.1.4 数据处理 .....	32
4.2 结果与分析 .....	32
4.2.1 参考方法分析结果 .....	32
4.2.2 近红外光谱采集 .....	33
4.2.3 建模回归方法的确定 .....	34
4.2.4 光谱预处理方法的确定 .....	35
4.2.5 定标模型的建立与验证 .....	37
4.2.6 在校准曲线中加入温度补偿样品 .....	39
4.3 小结 .....	40
<b>第五章 植酸酶产品定性分析模型的建立与优化.....</b>	<b>41</b>
5.1 实验部分 .....	42
5.1.1 实验仪器和样品 .....	42
5.1.2 光谱采集 .....	43
5.2 结果与分析 .....	44
5.2.1 主成分分析 .....	44
5.2.2 主成分得分结合马氏距离法建模预测 .....	44
5.3 小结 .....	47
5.4 液态植酸酶近红外光谱分析 .....	47
5.4.1 仪器设备 .....	47
5.4.2 样品制备 .....	48
5.4.3 光谱采集及数据处理 .....	48
5.4.4 液态植酸酶近红外光谱分析 .....	49
5.4.5 小结 .....	50
<b>第六章 植酸酶近红外指纹图谱库的应用研究.....</b>	<b>51</b>
6.1 植酸酶喷雾干燥过程与近红外指纹图谱库 .....	51
6.1.1 概述 .....	51
6.1.2 喷雾干燥实验 .....	51
6.1.3 喷雾干燥过程中近红外光谱信息研究.....	52

6.1.4 小结 .....	54
6.2 植酸酶储藏过程与近红外指纹图谱库 .....	55
6.2.1 概述 .....	55
6.2.2 恒温恒湿储藏试验 .....	55
6.2.3 恒温恒湿储藏过程中近红外光谱信息研究 .....	56
6.2.4 小结 .....	57
第七章 结论及建议 .....	58
7.1 主要结论 .....	58
7.2 建议 .....	59
参考文献 .....	60
致    谢 .....	65
作者简历 .....	66

## 英文缩略表

英文缩写	英文全称	中文名称
A	Absorbance	相对吸光度
ANN	Artificial Neural Networks	人工神经网络
CLS	Classical Least Squares	经典最小二乘法
COE	Constant Offset Elimination	常偏移量消除
DSC	Differential Scanning Calorimetry	差示扫描量热仪
GA	Genetic Algorithm	遗传算法
$\Delta H$	Enthalpy of Transition	变性热焓值
HPLC	High Performance Liquid Chromatography	高效液相色谱仪
ILS	Inverse Least Squares	逆最小二乘法
LW-NIR	Long Wavelength Near Infrared Spectroscopy	长近红外波段
LWR	Locally Weighted Regression	局部权重回归
MCM	Multiple Chain Method	多链方法
M-Distance	Mahalanobis Distance	马氏距离法
MLR	Multivariate Linear Regression	多元线性回归
MSC	Multiplicative Scatter Correction	多元散射校正
NIT	Near Infrared Transmittance Spectroscopy	近红外透射光谱
NIR	Near Infrared Reflectance Spectroscopy	近红外反射光谱
OD	Optical Density	吸光度值
OSC	Orthogonal Signal Correction	正交信号校正
PC	Principal Component	主成分
PCA	Principal Component Analysis	主成分分析
PCR	Principle Component Regression	主成分回归
PLS	Partial Least Square	偏最小二乘法
PRESS	Prediction Residual Error Sum of Squares	预测残差平方和
R <sup>2</sup>	The Coefficient of Determination	决定系数

英文缩写	英文全称	中文名称
RMSEC	Root Mean Square Error of Calibration	校正标准差
RMSECV	Root Mean Square Error of Cross Validation	内部交互验证均方差
RMSEP	Root Mean Square Error of Predication	预测均方差
RPD	Relative Predictive Determination	相对分析误差
RP-HPLC	Reversed Phase-High Performance Liquid Chromatography	反相高效液相色谱
RSD	Relative Standard Deviation	相对标准差
SDS -PAGE	Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis	十二烷基硫酸钠-聚丙烯酰胺凝胶电泳
SLS	Straight Line Subtraction	直线相减
SMLR	Stepwise Multiple Linear Regression	多元线性回归
SNR	Signal-to-Noise Ratio	信噪比
SNV	Standard Normal Variate	标准正态变换
SNVT	Standard Normal Variate and De-Trend	标准正态变换与去趋势变换
SW-NIR	Short Wavelength Near Infrared Spectroscopy	短近红外波段
T <sub>c</sub>	Conclusion Temperature	终止温度
T <sub>o</sub>	Onset Temperature	起始温度
T <sub>p</sub>	Peak Temperature	峰值温度
WT	Wavelet Transform	小波变换
1-VR	Coefficient of Determination in Cross Validation	交互决定系数

# 第一章 绪 论

随着世界经济和人口的增长，为快速发展的畜牧业提供充足、高质量的饲料产品将是农业科学家和政府决策者长期面对的巨大挑战；而高效利用现有资源、充分利用非常规饲用资源和以生物制品为代表的新资源也成为确保饲料工业和畜牧业可持续健康发展的基础支撑。近年来饲料工业中涌现出越来越多的生物制品，在一定程度上推动了饲料工业和畜牧业的发展，但这些产品在带来丰厚经济利益的同时，也将世界范围共同面对的生物安全风险问题直接带入了饲料和畜牧生产体系。生物制品将是饲料新资源开发的重点领域，涉及到生产使用者、科研开发机构、农业产业，乃至国家整体经济利益；而生物制品安全隐患隐蔽性强、潜伏期长，很难在短时期内全面准确判别这些新型产品的安全特性；因此，对饲用生物制品需要更为智慧的安全监管机制、建立健全的可追踪管理体系、严格规范饲用生物制品生产和加强饲用生物制品安全跟踪评价。

植酸酶是近年来新出现的一种饲料添加剂，用于提高畜禽动物对有机磷元素的吸收利用效率、减少动物粪便对环境的磷污染。随着人们环保意识的增强，商品植酸酶已在养殖业和饲料业得到广泛应用，约占饲用酶市场的 1/3。我国是饲用植酸酶主要的生产和出口国，但由于生产企业技术水平参差不齐，导致市场植酸酶产品质量良莠不齐，鱼目混珠现象严重。近红外光谱技术不仅能够反映绝大多数有机化合物的组成和结构信息，更为重要的是近红外光谱技术能够瞬时获取准确反映产品个体特征的数据信息，而用这些数据信息描述产品特性远比用单一化学成分指标和语言描述更为科学全面。针对近红外光谱技术特点和植酸酶安全监管面临的难点，我们将研究探索将近红外光谱技术引入饲用植酸酶产品的质量安全管理。

## 1.1 植酸酶在饲料工业中的地位

磷在动物配合饲料中占举足轻重的地位，参与动物牙齿、骨骼及细胞膜构成，是遗传物质核酸的组分，参与重要生长代谢过程。磷缺乏严重影响动物生长，导致动物的生产性能低下。单胃动物（猪、鸡、鸭）和水产类（鱼、虾）消化道内缺乏植酸酶，无法有效利用玉米、大豆等饲料原料中含量丰富的植酸磷，需要在饲料中添加磷酸氢钙来满足动物生长对磷的需求，而磷酸氢钙生产过程耗能高、污染重。动物未加利用的植酸磷形成的高磷粪便对环境的污染已经成为我国饲料工业和养殖业可持续发展中不可忽视的问题。我国畜禽粪便年产生量约为 17.3 亿吨，含磷 363 万吨，是化肥磷流失量的 1.3 倍，不断在饲料行业中推广使用能显著提高畜禽吸收利用磷元素的植酸酶具有深远的社会意义。

植酸酶是性质优良的饲料添加剂，约占饲用酶制剂市场的 1/3 份额，其具体生理功能包括：（1）使饲料中植物有机磷得到有效利用，在猪和鸡的养殖过程中，可以采用植酸酶部分或全部替代无机磷，降低饲料成本；（2）减少磷污染，采用植酸酶替代无机磷可以使猪禽粪便中磷含量减少 25%~65%；（3）显著改善畜禽生长性能；（4）在低磷或正常磷含量的日粮中添加植酸酶，不仅可提高植酸磷的生物利用率，而且增加钙、镁、锰、铜、锌、铁以及氮的生物利用率，促进动物矿质营养平衡；（5）促进动物对蛋白质、氨基酸及

碳水化合物的消化吸收。鉴于植酸酶显著的综合效益，特别是环保效益，欧盟各成员国、加拿大以及美国等发达工业国家为此从上世纪九十年代就制定了养殖业强制使用植酸酶的政策，东南亚、日本和韩国近期也将出台法律，把它作为一种“绿色磷”用以取代传统的无机磷酸盐，在环境保护中发挥了巨大作用。自 90 年代以来，芬兰、德国和我国等先后成功开发出商品植酸酶，特别是近年来，随着新菌种的选育和基因修饰技术的应用，提高了商品植酸酶的活性，显著降低了生产成本。目前饲用植酸酶绝大多数是通过微生物发酵方式生产，而近期中国工程院院士范云六先生开发了一项节能、环保、低成本利用玉米种子生物反应器生产植酸酶的新技术，将进一步推广饲用植酸酶的生产应用。总之，植酸酶已经成为应用广泛、生理功能明显、生产工艺发展迅速、商品化和市场化较成功的饲用生物酶之一。

生物安全性和严格的生产加工工艺是决定饲用酶制剂质量安全水平直接因素，首先，为了提高饲用酶制剂的生理功能、降低生产成本，以及针对饲用酶制剂产品直接应用对象为畜禽动物，饲用酶制剂生产工业成为最新分子生物学理论、基因克隆和分子操作技术的集中应用领域和试验场，而科学技术本身是把“双刃剑”，新技术不仅带来巨大生产力，而且会带来具有强破坏力的安全隐患。其次，生物技术在饲用酶制剂研发生产中主要应用途径包括利用重组微生物反应器高效表达目的酶，降低生产成本，以及利用基因工程技术改良饲用酶制剂，提高酶的质量与效率；尽管目前通过基因工程手段，将酶蛋白的基本结构改变，强化酶在某方面的功能特性的这一做法已成为饲用酶制剂商业推广的成功典范，但这种做法具有明显的安全隐患，对经改造，尤其是经过基因技术改造的酶制剂必须经过合理的、必要的生物安全性评价才能工业化生产和应用。再有，利用微生物发酵是当前生产饲用酶制剂的最主要生产方式，尽管这种酶的生产方法具有产量高、生产成本低、便于提高酶制品获得率等诸多优点，但微生物发酵法生产的酶制品中常会带入一些细菌自身的生理活性物质，这些生理活性物质往往对动物体有害，因此必需进行毒性实验验证产品的安全性。最后，微生物发酵生产工艺条件和参数亦直接影响饲用酶制剂产品质量安全水平，包括培养基、培养条件、纯化过程、载体、其他添加剂和成分等；如果改变这些工艺参数，除非能证明终产品与用原生产方法生产的产品没有区别，则需将其产品视为是新的加工和纯化方法重新进行安全性评价。

## 1.2 近红外光谱分析技术及在饲料工业中的应用

### 1.2.1 近红外光谱技术应用发展史

近红外光谱区域是人们发现的第一个非可见光谱区域，于 1800 年第一次被Herschel所发现<sup>[1]</sup>，距今已有近 200 年的历史。到 19 世纪前期，近红外谱区的分析在经典定量、定性分析方面已得到相当的发展，但由于分子在该谱区的波频和吸收信号均较弱，且谱带多相互重叠，信息解析相当困难，近红外光谱技术一直没有得到实际应用。19 世纪末，Abney 和Festing在NIR短波区域首次记录了有机化合物的近红外光谱<sup>[2]</sup>。由于缺乏可靠的仪器基础，在 20 世纪 50 年代以前，NIR的研究工作只局限于为数不多的几个实验室中，很少有实际应用。以致Wheeler于 1960 年称近红外谱区为“被遗忘的谱区”<sup>[3]</sup>。此后，随着简易

近红外光谱仪的出现及美国农业部的Karl Norris等人所做的工作，使近红外光谱技术在农副产品分析中得到广泛应用，如他们成功采用NIR测量了农副产品（包括谷物、饲料、蛋类及动物体等）中的水分、油脂及蛋白含量<sup>[4-5]</sup>。为了排除样品中的脂肪、蛋白质等其他成分干扰水分的测定，他们采用了类似光谱分析中的多组分测定技术：同时测定几个波长处样品的透射率或反射率，并用统计学方法分别计算出样品中的水分、蛋白质、脂肪等含量。为了克服固体样品透射率低的困难，他们采用了漫反射技术；为了克服样品状态对测定带来的影响，他们用反射吸光度之差作为光谱参数，来计算样品中各成分的含量<sup>[6]</sup>；并发展了一套检验方法，用以决定某种计算方法的舍取。这些方法与传统的（研究光谱、确定谱峰归属，再进行定性、定量分析）光谱分析方法有很大不同。因此，当时这一领域的工作没有受到分析界的足够重视。

1970年美国的一家公司率先研制了近红外品质分析仪器。该仪器使用了六个波长的窄带干涉滤光片，主要用于分析农产品中水分、蛋白质等的含量。但NIR在分析测试中表现出灵敏度低、抗干扰性差等弱点，从而使得人们淡忘了该技术的应用，在此后约20年中，NIR技术几乎处于一种徘徊不前的状态。1983年，Wetzel称之为“光谱技术中的沉睡者”。20世纪80年代，仪器分析特别是光谱分析领域的有关专家开始注视近红外光谱分析技术的发展。著名的应用光谱杂志自80年代以来刊载许多相关论文，以及一系列评论，其中包括有机物的近红外光谱分析<sup>[7]</sup>、近红外定性定量分析<sup>[8]</sup>、近红外药物分析<sup>[9]</sup>以及近红外反射光谱分析进展<sup>[10]</sup>。由于光谱界、分析界与近红外应用领域人员的合作，近红外光谱分析进一步采用了化学计量学中多元校正方法，以及现代光学、计算机数据处理技术，使近红外光谱分析技术发展为现代近红外光谱分析技术。1987年在挪威召开了第一届近红外光谱分析国际会议，1988年成立了国际近红外光谱协会（ICNIRS）<sup>[11]</sup>。此后，近红外光谱应用领域不断扩展，一些国际著名的光谱仪器公司纷纷推出了近红外光谱分析仪器。20世纪90年代国际分析界逐渐形成了近红外光谱分析的热潮。近年来近红外光谱分析技术一直是匹兹堡会议的热点<sup>[12-14]</sup>。McLure在1994年发表了一篇题为“The giant is running strong”的论文。1998年，Davies撰文讨论了近红外光谱技术的潜在用途和发展趋势，并将其描述为光谱领域中“从沉睡者变为了启明星”的技术<sup>[15]</sup>。

我国近红外光谱技术的应用研究始于农业领域，我国粮食与农业系统在1977年前后才开始从国外引入近红外光谱分析仪，但由于建立数学模型的困难，到20世纪80年代才发表了我国第一批有关研究论文<sup>[16-17]</sup>。我国农业仪器分析工作者在20世纪80年代初就关注此项技术。1983年前后开始在大型通用傅里叶变换红外光谱仪上进行近红外光谱分析技术在农业分析的理论研究和应用研究<sup>[18]</sup>。我国石油科学与药物科学等领域有许多研究报道<sup>[19-20]</sup>。我国光谱学界也很重视此项技术，在全国光谱学术报告会上发表了一批近红外光谱应用的论文<sup>[21]</sup>。在著名的PITTCON会议上也发表了一些近红外光谱应用的论文<sup>[22-23]</sup>。20世纪90年代我国开始了近红外光谱仪器的研制<sup>[24-25]</sup>。但建立定量分析数学模型的困难一直是影响该技术在我国广泛应用的重要原因。目前我国一些高产值的产业已经率先广泛运用近红外光谱分析技术，如我国的石油化学、烟草、乳品、医药与农业等行业，近年来都已经进口了一大批高精度近红外光谱仪，其中石油化学与农业领域还应用了一批国产仪器，它们在保证和提高产品质量方面起了重要作用。我国石油化学与药物生产领域开始了近红外光谱在线分析的应用<sup>[26-27]</sup>。

## 1.2.2 近红外光谱分析技术的主要应用领域

现代数学和计算机技术在分析化学中的应用,使得化学计量学这门新兴学科逐渐形成。而近红外光谱定量分析技术是化学计量学在光谱分析中应用的结果。从50年代起,化学计量学和光谱分析的结合揭开了分析技术史上崭新的一页,随后此种技术在数十类农业、生物样品上,测试过几百种项目,测定的精度可满足农业研究的要求,其中有的项目已被AACC, AOAC, ICC所采用<sup>[28-30]</sup>。

近红外光谱分析技术诸多优点决定了它应用领域的广阔,使其在许多行业中都能发挥积极作用,并逐渐扮演着不可或缺的角色。下面提供的是有关文献归纳的近红外光谱分析技术的应用领域、分析对象以及各自的分析指标,以供参考(见表1-1)。

表1-1 近红外光谱分析技术应用领域<sup>[31]</sup>

Table 1-1 Application field of Near Infrared Spectroscopy

应用领域	分析对象	分析指标
食品	酒制品	产地、真伪鉴别等
	葡萄酒	乙醇,含糖量,有机酸,含氮值,pH值等
	白酒	原料中的水分,淀粉,支链淀粉;pH值,淀粉和残糖等
	啤酒	大麦原料中的水分,麦芽糖;啤酒中的乙醇和麦芽糖等
	饮料	咖啡因,糖分,酸度,果汁真伪鉴别等
	调味品(酱油、醋等)	蛋白质,氨基酸总量,总糖,还原糖,氯化钠,总酸,总氮,品质分级,真伪鉴别等
	乳制品(牛奶等)	乳糖,脂肪,蛋白质,乳酸,灰分,固体含量等
	玉米浆,蜂蜜	果糖,水分,葡萄糖,多醣类,偏振参数等
	食用油(花生、豆油和菜籽油等)	原料中油分含量;食用油中的脂肪酸,水分,蛋白质,过氧化值,碘值,真伪鉴别等
	烘焙食品	脂肪,蛋白质,水分,淀粉,面筋等;方便面油分等
	肉类(猪、牛、鸡肉,鱼类,香肠等)	蛋白质,脂肪,水分,各种氨基酸,脂肪酸,纤维素等,以及新鲜及冷冻程度,产品种类,真伪鉴别等
	大麦,小麦,豆类,水稻,甘薯,面粉,及其它谷类	脂肪,蛋白质,水分,纤维量,淀粉,产地、产季鉴别,品质等级,谷物成熟度,病虫害等
	饲料	干物质,粗蛋白,粗纤维,粗灰分,消化能,代谢能,氨基酸,棉酚,植酸磷,哇乙醇含量,抗营养物质,预混料中维生素等
	烟草	尼古丁,水分,总糖,还原糖,多酚类,总氯,添加物,产地鉴别,等级分类等
农牧	咖啡	咖啡因,水分,绿原酸,种类、产地鉴别,品质分级等
	水果,蔬菜	糖分,酸度,维生素,纤维素,品质分级,成熟度,硬度
	茶叶	老嫩度,氨基酸、茶多酚、咖啡碱,水分,总氮,品质分级,真假识别,品种鉴定等
	土壤	水分、有机质和总氮含量,土壤分类等
	其它	堆肥的品质、腐熟度等
石油炼制	原油	密度,实沸点蒸馏,浊点,油气比;油砂中沥青含量等
	成品汽油	辛烷值(RON、MON),密度,芳烃,烯烃,苯含量,MTBE等

	催化裂化汽油	辛烷值, PIONA (直链烷烃、异构烷烃, 芳烃和烯烃), 馏程等
	裂解汽油	辛烷值, 二烯、二甲苯异构体含量等
	石脑油	POINA, 密度, 分子量, 馏程, 乙烯潜收率, 结焦指数等
	柴油	十六烷值, 密度, 折光指数, 凝点, 馏程, 芳烃组成等
	航煤	冰点, 芳烃, 馏程等
	润滑油	族组成, 基础油粘度指数, 粘度, 添加剂等
	重油	API 度, 渣油中 SARA 族组成; 沥青中蜡含量等
高分子	原料	纯度, 水分, 羟基含量等
	加工过程	聚合度, 动力学、热力学性质测定, 添加剂含量等
	产品	共混、共聚物的组成, 分子量, 密度, 熔融指数, 残余单体、溶剂, 添加剂含量, 粒度分布, 力学性能, 回收废塑料类别鉴定等
制药	原料	原料药的主要活性成分, 结晶状态、粒径、旋光性和密度, 鉴别中药材的真伪、产地和品质分级等
	加工过程	混合均匀性, 干燥过程水分, 注射用产品灭菌, 膜衣厚度, 粒径, 溶出度, 药物中微生物定性定量监测等
	产品	主要成分, 硬度, 包装材料的鉴定, 稳定性, 真伪鉴定等
其它	临床医学	全血或血清中血红蛋白载氧量、胆固醇、蛋白质、葡萄糖、尿素等含量; 无创血糖监测; 尿液中尿素、肌氨酸酐和蛋白质; 烧伤伤口分类; 组织氧含量; 脑氧饱和度和血流动力学; 细胞病理如癌细胞鉴别等
	纺织	混纺织品中各成分含量, 纤维的染色性, 棉织物丝光度, 羊毛髓化度, 二醋酸纤维素醋化值, 地毯纤维类别鉴定等
	造纸	纸浆中木素含量, 卡伯值, 纸页水分、涂层含量等
	煤炭	水分, 挥发份、灰份、含热量, 品质分级等
	生物化工	生物发酵过程中乙醇、葡萄糖、乳糖、氨基酸、乳酸盐和谷氨酸盐等含量, 细胞密度, 反应动力学跟踪, 菌种鉴定
	制糖	蔗汁、碎蔗、蔗渣、原糖、成品糖的旋光度、锤度、糖度、色度、浊度、粒度、固体物和水分含量等
	日用品	原料纯度, 香料, 油脂混合物分析, 蜡成分鉴别, 均匀度, 牙膏中氟含量, 表面活性剂含量, 真伪鉴定等
	油漆和墨水	原料分析, 溶剂纯度, 色素品质等
	环保	海洋石油、土壤污染源鉴定; 废水 pH、BOD、COD 等

### 1.2.3 近红外光谱技术在饲料工业中的应用

在饲料分析方面, 近红外光谱技术不仅能测定饲料中的常规成分, 如水分、粗蛋白、粗纤维、粗灰分, 而且能测定饲料中的微量成分, 如氨基酸、维生素、有毒有害物质(棉酚, 总葡萄糖等)。现在近红外光谱技术已经向饲料营养价值评定方面发展, 如测定可消化的氨基酸、消化能、代谢能、预测动物的反应(动物的采食量、消化率)等, 应该说, 近红外光谱分析技术几乎涵盖了所有评价饲料营养价值的指标。从近红外分析原理来看饲料中需要测定的有机成分都可以用近红外进行测定; 而从测试准确性来看, 在测定原料水分、大多数谷物及牧草的蛋白质含量、动物性副产品、豆粕及大多数饲料原料、脂肪含量低于5%饲料的脂肪及油分方面, 均可达到较好的预测效果。因此, NIR 技术在饲料分析中有着广阔的应用前景。

除此之外, NIR还已成功用于饲料产品质量及原料的在线质量监控。在线质量监控是保证饲料产品质量安全的有效途径之一, 因为通过在线监控, 可以更加科学、准确及时地获取饲料产品的质量信息, 从而尽快对生产参数做出调整, 确保饲料质量的稳定, 减少损失, 降低成本。但由于用传统化学分析方法获取检测结果往往时间长, 存在严重的滞后性, 无法满足现代工业生产中对维持产品质量稳定和实时监控生产过程的要求; 而近红外光谱分析技术可在复杂背景下瞬时分析样品的组分信息, 且通过与光纤技术结合还可实现在不破坏样品和不影响生产的前提下, 可直接准确地获取生产过程中的实时信息, 达到对各种工业生产过程进行在线监控的要求。因此, 探索将近红外光谱分析技术用于在线监控饲料原料、半成品、成品的营养成分等参数的变化, 有效提高饲料产品质量控制的效果, 在饲料质量监控中有着同样的研究价值和应用潜力。

#### 1.2.4 近红外光谱技术在监控生物发酵过程中的应用

20世纪90年代以来, 近红外光谱分析技术在生物安全管理中的应用得到越来越多的关注, 许多生物反应过程也尝试着使用近红外光谱分析技术进行监控。尤其是在生物发酵过程质量控制中, 近红外光谱分析技术已广泛用于监控和优化整个发酵过程<sup>[32]</sup>。由于在酶制剂微生物发酵过程中, 随着微生物总量的不断生长, 营养水平将逐渐趋于下降, 为优化发酵过程, 需要实时监控发酵液的营养状态, 使之维持在一个最佳的水平; 另外, 微生物总量也需要被实时监控, 以避免过度生长和抑制生长; 而近红外光谱分析技术可以快速地测定发酵液的营养成分浓度和微生物总量累计程度, 因此, 已被成功应用于各类发酵过程的控制。如在酶制剂生产发酵过程中测定生产的碱性蛋白酶的活性, 在放线菌发酵过程中实时分析培养液营养物质中糖、谷氨酸盐、甲基硫酸盐的浓度等。值得注意的是, 在发酵过程中, 随着生物细胞的增长, 发酵液将由透明状态逐渐变成高散射度的不透明状态, 因此, 光谱采集方式可以采用透射(用以初始的透明状态)和反射方式(用于后期的不透明状态)两种方式分别进行, 而在总体趋势上, 反射吸光度光谱的基线随发酵时间和生物量的增长而提高, 但也受到发酵液其他组分的影响。另外, 近红外光谱分析技术也逐步应用在生物催化反应过程监控, 这为改进催化反应工艺开辟更为便捷的新途径。

### 1.3 近红外指纹图谱库技术

#### 1.3.1 指纹图谱库的原理

“指纹图谱”是起源于犯罪学与法医学刑侦技术中的专业术语。在19世纪末20世纪初指纹鉴定技术已广泛应用于刑事案件的侦破过程。人们已认识到人类的指纹具有共同特征, 即基本可以分为拱形、环形和螺纹形三大类; 但每一个人的指纹在微小的细节构造上则各有不同, 这种唯一的指纹特征则是相对固定不变的, 可用来鉴别犯罪现场的痕迹, 追踪犯罪分子。刑侦专家的工作就是在共性的指纹特征中寻找出犯罪者个性指纹的“唯一”特征。同理, 对于同一类型的生物制品, 由于载体相同, 那么必然有群体共有的特征, 这样就在它们的光谱图(指纹图谱)上就会表现出相同的特征。但由于菌种来源、发酵条件、

加工工艺、储存方式的不同而可能产生较为明显的差异，在其指纹图谱中又会反映出其与其它群体的不相似性。指纹图谱分析就是利用这一特点，通过分析同一产品，或者同一产品不同批次间的指纹图谱的差异，反映出其内在的均一性和稳定性，差异越小，说明稳定性越好。因此，指纹分析的最终目的就是判定原料或产品的质量是否稳定，确定其真伪及优劣。

### 1.3.2 近红外指纹图谱在农产品和中药安全管理中的应用

目前，近红外指纹图谱已成功应用于中药成分的鉴定和作物品种的鉴定和分类。中药的近红外指纹图谱是一种综合的、整体的鉴定手段，它可准确且又可量化地对药材进行真伪鉴别和质量评价，其功用范围在于鉴别真伪，以及评价药材的质量均一性和稳定性。李峰<sup>[36]</sup>的研究表明，应用红外光谱建立金银花的红外指纹图谱，可以用于反映和鉴别金银花成分，以及区别不同来源和产地的金银花药材。黄宝康等<sup>[37]</sup>应用近红外漫反射指纹图谱聚类分析鉴别不同产地楮实子的真伪，发现不同产地楮实子之间指纹图谱存在一定差异，与伪品差异较大。周群，孙素琴等采用傅立叶变换红外光谱法（FTIR）并结合聚类分析法（SMICA）对不同产地的枸杞进行聚类分析，所建立方法可对不同产地的枸杞进行快速、准确的鉴别。孙素琴等<sup>[38]</sup>利用红外指纹图谱并结合计算机辅助对比辅助软件技术，对9个省份的45个山药样品进行了道地性与非道地性的鉴别研究，发现各产地山药的相关性有一定的差异性，该方法快速、简便、可靠。荆森等<sup>[39]</sup>应用近红外漫反射光谱技术建立了甘草的指纹图谱库，对不同产地、不同部位、不同级别的甘草加以鉴别，取得了良好的效果。

### 1.3.3 近红外指纹图谱与饲料产品注册管理

20世纪，“疯牛病”、“二恶英”、污水饲料等事件的接连发生促使各国加快实施“从农场到餐桌的全程质量安全监管”以保证食品链的安全，而饲料工业作为食品链的重要一环，在保证和控制食品质量中起着关键的承上启下的关键作用；另外，从饲料工业自身情况出发，饲料工业每年要消费大量的安全特性尚未明了的生物产品，如转基因作物、各种微生物资源等，对这些产品实施有效的安全评价和跟踪管理是十分必要且相当紧迫的。为了增强饲料产品质量控制的透明度和效果，提高饲料工业和食品行业的整体安全水平，有必要采取有效措施，严格饲料产品的注册管理、建立完善的饲料产品可追溯管理制度。

中国目前注册饲料企业产品种类和规格非常繁多复杂，质量也是良莠不齐，有效管理和控制饲料产品的重要手段之一就是强化和完善产品登记注册制度及其内容。饲料主管部门通常要求饲料生产企业或进口饲料经营商提供饲料产品的有效组分、理化性质、功能、生产工艺条件、安全性、稳定性、产品来源等相关个体特征信息；而建立饲料产品的近红外光谱库可以快速、有效、准确的反映有关饲料产品的上述信息和特性。因此，在饲料产品登记注册中建立一一对应的每种产品的近红外光谱信息库，作为产品注册信息的一部分，不仅能更为全面、准确地反映产品的个性特征信息，而且有利于快速鉴别饲料产品掺假，遏制和打击假冒伪劣产品，维持公平竞争的市场秩序；同时，也将有助于饲料标签制度和其他相关饲料法规的贯彻执行，最终全面实现产品注册管理的目标。

### 1.3.4 近红外指纹图谱与饲用酶制剂质量安全管理

由于生产饲用酶制剂的微生物菌种的差异，及发酵工艺条件和纯化工艺等的差异，产品组成也是复杂多变，从而导致其质量等级千差万别，要系统规范的管理这些产品，亟待针对性建立各类产品的近红外光谱信息库。尽管近红外光谱分析技术在监控饲用酶制剂的发酵过程已有一定的应用基础，但是由于生物酶制剂是一个复杂的混合物体系，所含各化合物吸收强度的叠加具有难以解析的复杂性，发酵产品的主体成分相似，造成近红外光谱图有较大的相似性，很难直接对不同产品的近红外光谱信息进行鉴别。而由于相同微生物菌属的发酵产物，只要发酵工艺和纯化工艺等存在差异，就会导致其化学成分含量不同，近红外的光谱就会有差异，因此借助其近红外指纹图谱库的差异，运用化学计量学判别法对光谱图进行解析，建立相关的模式识别法，就可实现微生物酶制剂类别的快速鉴别。因此，利用近红外光谱法建立植酸酶的指纹图谱库，必将在其产品质量监控中展现出诸多优势，检测成本低廉，应用简便快捷，利于产品品牌保护，一旦图谱库建立完毕，对样品伪劣真伪鉴别只需数分钟。

## 1.4 课题的选题意义

由于生产饲用酶制剂的微生物菌种、发酵工艺条件和纯化工艺等的差异，造成产品组成复杂多变，如果其发酵工艺和纯化工艺等发生改变，就会导致其化学成分含量的变化，要系统规范的管理这些产品，亟待建立更加科学合理的生物制品质量安全管理指标体系。本课题将通过研究构建植酸酶产品的近红外光谱指纹图谱库，为植酸酶产品质量控制、安全监管、真伪鉴别和品牌保护提供科学的技术方法；而逐步推动近红外光谱指纹图谱库在饲用生物制品、甚至是饲料添加剂注册管理中的应用，不仅利于生产者提高产品质量的稳定性，而且利于管理者建立高效的生物制品可追溯管理体系，不断提高生物产品整体安全水平，同时又在一定程度上保护生产企业的知识产权，激发企业创新积极性，最终确保实现生物制品的生产经营者、消费者、和管理者三方面的利益。

针对植酸酶生产管理和市场监管中对快速、简便、经济、准确的酶活检测方法的需要，本课题还将建立快速测定植酸酶活性的近红外光谱分析方法，提高植酸酶质量管理的效率。另外针对近红外光谱技术已成功用于植酸酶等生物酶制剂的发酵生产过程的质量安全监管，而在后加工和储藏环节几乎没有应用的现状，但这些环节也在一定程度上影响决定植酸酶产品的质量，因此，本课题还将建立快速跟踪植酸酶产品喷雾干燥过程和储藏过程中质量变化的近红外光谱分析方法。总之，本课题将近红外光谱指纹图谱库引入到植酸酶的质量安全管理中是一个新颖并且具有实际意义的课题，不仅在一定程度上拓展了近红外光谱技术在生物制品安全管理领域中的应用模式，而且也将有力的推动饲用生物制品安全监管机制的创新。

## 1.5 本研究的主要内容及预期目标

### 主要研究内容：

- (1) 借助反相高效液相色谱(RP-HPLC)和SDS-PAGE电泳研究植酸酶蛋白组成；
- (2) 利用差示扫描量热仪(DSC)研究植酸酶产品的变性温度和热稳定特性；

- (3) 研究建立植酸酶产品的近红外指纹图谱库;
- (4) 利用所建立的植酸酶近红外指纹图谱库来定性判别植酸酶产品的真伪;
- (5) 研究建立快速准确预测植酸酶酶活的近红外光谱检测方法;
- (6) 通过喷雾干燥实验, 分析干燥条件对植酸酶近红外光谱信息的影响, 探讨近红外光谱信息在监控植酸酶喷雾干燥过程的应用模式;
- (7) 设计饱和平衡盐恒温恒湿储藏实验, 研究分析植酸酶近红外光谱信息变化与储存品质变化之间的关系。

#### 预期目标:

- (1) 研究分析不同植酸酶产品的成分组成和功能特性, 明确监控植酸酶质量安全水平和鉴别区分植酸酶产品的关键指标, 探讨植酸酶的近红外光谱信息与上述指标之间的关系;
- (2) 研究建立能够清晰区分不同生产企业、不同产品形式植酸酶产品的近红外光谱指纹图谱库, 为植酸酶产品真伪鉴别和品牌保护提供科学合理的技术方法;
- (3) 建立快速检测植酸酶酶活的近红外光谱分析方法;
- (4) 探索建立快速跟踪喷雾干燥加工过程、储藏过程中植酸酶质量变化的近红外光谱分析方法;
- (5) 逐步将近红外光谱技术引入生物制品的注册管理, 促进生产者提高产品质量的稳定性, 协助管理者建立高效的生物制品可追溯管理体系, 提高生物产品整体安全水平, 从而确保实现生物制品的生产经营者、消费者、和管理者三方面的利益。

#### 课题研究技术路线图

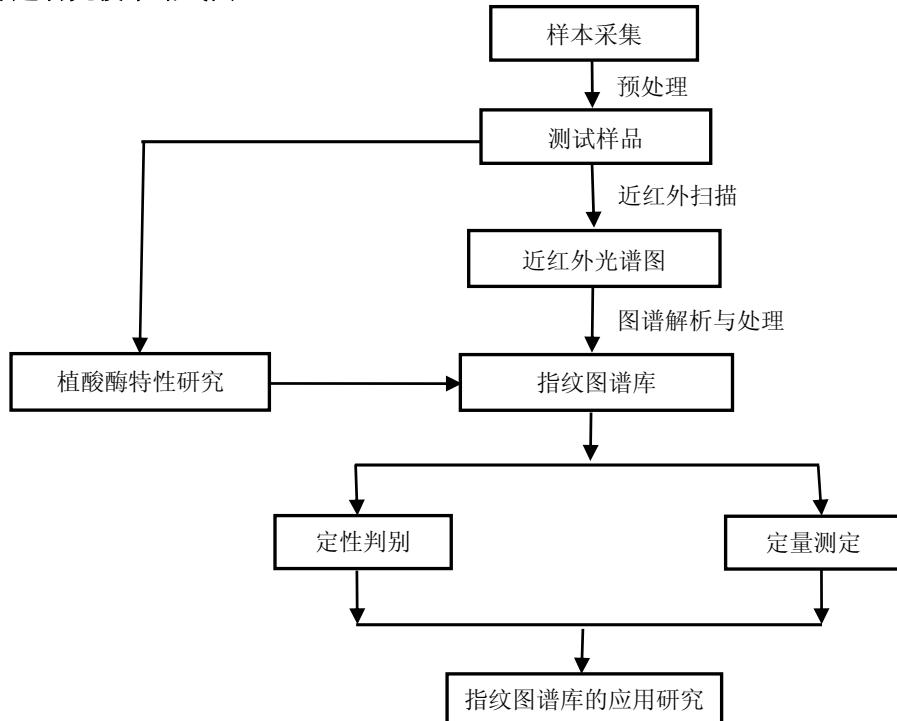


图 1-1 课题研究技术路线图

Figure 1-1 Flow chart of the research

## 第二章 植酸酶产品特性研究

### 2.1 植酸酶蛋白组成特性研究

#### 2.1.1 概述

植酸酶产品中蛋白质成分的组成直接决定着产品的酶活和质量安全水平,微生物表达体系、发酵工艺、纯化工艺是影响植酸酶产品中蛋白质组成变化的重要因素。本小节主要是利用 SDS-PAGE 和 RP-HPLC 研究不同植酸酶产品的蛋白组成。SDS-聚丙烯酰胺凝胶电泳 (SDS-PAGE) 和高效液相色谱 (HPLC) 是通用的,也是经典的蛋白质组分分析技术。蛋白质颗粒在各种介质中的迁移率通常取决于他所带的静电荷以及分子大小和形状等因素。由于 SDS (十二烷基硫酸钠) 是阴离子,使多肽链覆盖上相同密度的负电荷,且电荷量远超过蛋白质分析原有的电荷量,完全掩盖了不同种类蛋白质间原有的电荷差别,因此, SDS-聚丙烯酰胺凝胶电泳中蛋白质的迁移率主要取决于它的分子量,而与原来所带的电荷和分子形状无关,即这种方法也就成为了蛋白质分子量大小的经典检测方法。反相高效液相色谱法 (RP-HPLC) 是另一种通用性的蛋白质分离技术。它是利用蛋白质表面的非极性区域同固定相表面的非极性配体相互作用,按照疏水作用力大小进行分离的一种方法。在分离时,RP-HPLC 由于使用酸性极性有机溶剂作为流动相,使用疏水性极强的非极性烷链作为固定相的配体。在 RP-HPLC 分离体系中,固定相的变化对分离效果影响不大。对分离效果影响较大的是流动相的组成和洗脱方式的选择。蛋白质在分离时的洗脱方式有梯度洗脱和等浓度洗脱,其中梯度洗脱最为常用。

本节试验利用 SDS-PAGE 蛋白电泳和反相高效液相色谱 (RP-HPLC) 来准确分析不同植酸酶产品的组分构成和各组分蛋白分子量的分布,确定植酸酶产品蛋白组成与产品功能特性之间的关系;实验中植酸酶样品是来源于三个不同生产企业(分别命名为 A、B、C)的液态植酸酶产品(大肠杆菌属来源)和植酸酶标准品(购于 Sigma, 曲霉来源)。

#### 2.1.2 SDS-PAGE 分析植酸酶蛋白组成

##### 1. 样品制备

取三个生产企业的液态植酸酶产品各1.0ml,以及植酸酶标准品0.1g,分别放入到5ml的离心管中,参考2.1.3进行样品制备,并于4℃贮存,备用。

##### 2. SDS-聚丙烯酰胺凝胶电泳

取制备好的植酸酶标准品和三个不同生产企业的液态植酸酶溶液,加入80 $\mu$ l的 1×SDS 凝胶缓冲液,涡旋混匀,以5000r/min离心5min。上样前将样品以12000r /min离心5min,取10 $\mu$ l上样进行SDS-PAGE。分离胶浓度12.5%,浓缩胶浓度4.4%。分离胶长13cm,浓缩胶长2.5cm,胶板厚度为1mm。在常温下进行,当指示剂迁移至距胶板下缘1cm时停止电泳,电泳全过程约5~6 h,电泳结束,迅速进行染色和固定。

##### 3. 凝胶染色

采用0.25%考马斯亮蓝R-250（45%甲醇+45%双蒸水+10%乙醇+0.25%考马斯R-250）进行染色<sup>[40]</sup>，染色时间约为4~5 h，染色和固定同时进行。然后进行脱色，多次更换脱色液（45%甲醇+45%双蒸水+10%乙醇），至背景清晰为止。

#### 4. 凝胶成像

采用Alpha Imager 2200凝胶成像系统对脱色后背景清晰的凝胶进行成像。

#### 5. 结果与分析

电泳图谱左边的M泳道为分子量标准蛋白质。植酸酶产品经SDS-PAGE得到了较好的分离效果（如图2-1所示）。实验结果清晰表明不同生产企业的植酸酶产品具有不同的蛋白电泳图谱，四个植酸酶蛋白电泳图谱共分离出不同条数的蛋白带，不同生产企业的电泳谱带在数目、迁移率及谱带的强度等方面均存在显著差异。来源于三个生产企业的植酸酶产品的植酸酶分子量均在45KD左右，而来源于曲霉的植酸酶标样的分子量为65KD，表明三个生产企业的微生物表达体系相近，但由于发酵工艺参数的差别，植酸酶分子量略有差别，这与前人的研究结果相一致，即不同的培养基和发酵条件将产生不同分子量大小的植酸酶产品。另一方面，从图谱中可以清晰看到不同品牌的植酸酶产品中蛋白质组成之间存在显著差异，其中来源于C生产企业的植酸酶产品的蛋白质成分组成单一，而来源于A、B生产企业的植酸酶产品蛋白质成分组成复杂，表明不同生产企业的植酸酶纯化工艺存在明显的差别。

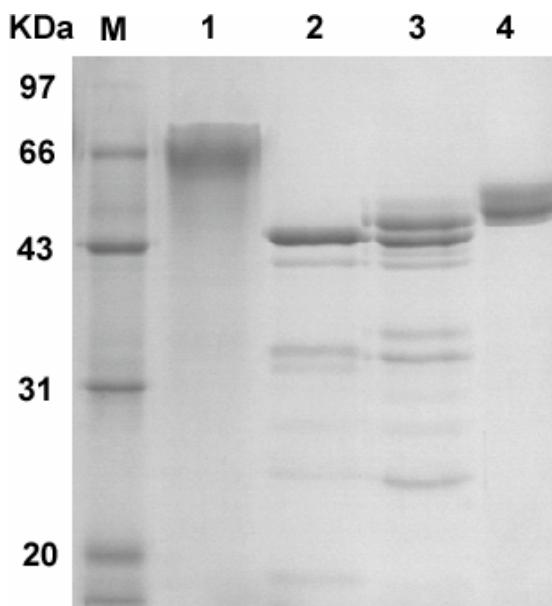


图 2-1 植酸酶样品 SDS-PAGE 图 (1: 标准品, 2: 产品 A, 3: 产品 B, 4: 产品 C)

Figure 2-1 SDS-PAGE analysis of phytase samples (1: standard, 2: product A, 3: product B, 4: product C)

#### 6. 讨论

研究分析发现不同植酸酶产品间的SDS-PAGE图谱差异明显，不同产品都有各自特异的蛋白质组成，不同品牌植酸酶有效成分的组成和含量各不相同。植酸酶产品内蛋白质组分的变化直接影响产品的质量安全水平，因为杂蛋白的存在不仅会直接影响功能蛋白的活性和热加工过程的再折叠作用，更为重要的是这些杂蛋白存在极大的安全隐患，因此，各国在生物制品的安全管理中均规定当产品中的蛋白质组成发生变化时，产品要进行重新的安全评价和申请注册。总之，在饲用植酸酶的安全管理中，有必要也必须跟踪监控植酸酶产品中的蛋白组成及其变化。

### 2.1.3 反相高效液相色谱分析植酸酶蛋白组成

#### 1. 仪器

Agilent1100高效液相色谱仪(美国); Agilent Chemstation工作站(美国安捷伦科技公司); 检测器: Agilent DAD; 色谱柱: ZORBAX 300SB C18柱 (4.6×150mm, 粒度5um); 检测波长254nm; 5KQ-250B型超声清洗仪、电子天平、离心机、酸度计、透析袋(D36mm, 截流范围: 8000-15000)。DF-C型恒压恒流电泳仪(北京东方仪器厂), DF-10型圆盘电泳槽(北京东方仪器厂), 小玻璃管(16cm×Φ5.5mm)若干。

#### 2. 试剂

双蒸水(新制备); 乙腈(ACN, 色谱纯, Fisher公司); 三氟乙酸(TFA, 分析纯); 植酸酶标准品(购自Sigma公司); 聚乙二醇8000(PEG8000, 上海凯必特化工有限公司); 丙烯酰胺(Acr, C. P., 临海化学厂, 经重结晶处理); 甲叉双丙烯酰胺(Bis., 瑞士Fluka公司, 重结晶); 四甲基乙二胺(TEMED, 生化试剂, 上海前进试剂厂); 三羟甲基氨基甲烷(Tris, Buffer grade, 进口分装); 十二烷基硫酸钠(SDS, C.P., 上海化学试剂厂); 其他试剂皆为分析纯。

#### 3. HPLC 测定方法研究

##### (1) 仪器、试剂的选择

进行植酸酶HPLC指纹图谱研究, 必须具备一台双泵的HPLC色谱仪, 才能进行梯度洗脱研究; 具备DAD(二极管阵列)检测器, 才能较好的进行不同波长、不同梯度洗脱的优选研究; HPLC仪器还应配有功能较多的色谱工作站, 才能进行图谱的分析处理。

本试验选用Agilent1100高效液相色谱仪(美国); Agilent Chemstation工作站(美国安捷伦科技公司); 检测器: Agilent DAD。测试方法为: 分别将供试液溶液和对照品溶液用微孔滤膜过滤后转入样品瓶内, 设定HPLC色谱仪的测定参数: 在线脱气, 自动进样, 每次10μl, 流速1.0ml/min, 柱温40℃, 记录70min的色谱图。试验所用的试剂应为分析纯或优级纯, 蒸馏水或超纯水应为新制备的, 使用前均应过滤。

##### (2) 样品溶剂及流动相的选择

在HPLC分离前, 多种溶剂可用于使蛋白溶解和/或变性。通常, 样品浓度如达到0.1~2mg/ml, 用于分离已足够。样品最好溶于可作为初始流动相的溶剂中[如0.05~0.5%三氟乙酸(TFA)-水作反相分离]。6M盐酸胍或尿素缓冲液具有: 增溶效果明显, 能与多数流动相混溶, 可使许多蛋白质有效地变性等优点<sup>[41]</sup>。由于植酸酶的成分复杂, 试验中选用单一流动相难以很好的分离, 而采用梯度洗脱可以达到很好的分离效果, 梯度洗脱是流动相中含有两种或两种以上的不同极性的溶剂, 在洗脱过程中连续或间断改变流动相组成, 以调节它的极性使每个流出的组分获得最理想的分离效果并使样品中的所有组分可在最短的分析时间内以适宜的分离度获得圆满的选择性分离。经梯度洗脱分离的各共有色谱峰峰型尖锐, 分离较好。试验中我们先后考察了不同比例的流动相(乙腈:水)等, 最后确定了具有良好分离效果的梯度洗脱条件。

##### (3) 色谱柱的选择

不同规格品牌的色谱柱在填充材料、粒径、孔径、键合相种类、密度、封端技术、色谱柱尺寸、适用范围等方面均有差异。根据大量的文献报道<sup>[42-43]</sup>, 本试验选用ZORBAX

300SB-C18 柱 ( $4.6 \times 150\text{mm}$ , 粒度  $5\mu\text{m}$ ) 作为分离柱。

#### (4) 检测波长的选择

研究中我们利用 Agilent DAD 检测器具有的所有模式储存采集光谱图的特点, 通过从光谱图数据中筛选信号, 获得分离色谱峰信息较多的检测波长, 再以此波长为指纹图谱的检测波长, 进行新的试验研究。 $254\text{nm}$  为检测波长的指纹图谱中峰最高, 分离的较好; 所以我们选择  $254\text{nm}$  为植酸酶指纹图谱的检测波长。

### 4. 标准品溶液制备

精密称量植酸酶标准品  $10\text{mg}$ , 置入  $25\text{ml}$  烧杯中, 加入  $7\text{ml}$  乙酸缓冲液 (称取  $34.02\text{g}$  三水乙酸钠,  $0.1\text{g}$  牛血清白蛋白于  $1000\text{ml}$  容量瓶中, 加入  $900\text{ml}$  水溶解, 用乙酸调节 pH 至  $5.50 \pm 0.01$ , 并用蒸馏水定容到  $1000\text{ml}$ , 室温下存放 2 个月有效), 在磁力搅拌器上高速搅拌  $30\text{min}$ , 然后用乙酸缓冲溶液定容至  $10\text{ml}$ 。摇匀, 在离心机上以  $5000\text{r}/\text{min}$  离心  $10\text{min}$ , 取上清液备用。

### 5. 样品制备

本试验所收集的三个植酸酶生产企业的液态植酸酶产品分别被编号为 Phytase A, Phytase B, Phytase C (市售产品)。

(1) 透析除盐: ①透析袋的处理方法, 商品透析袋常涂有甘油以防破裂, 并含有微量硫化物、重金属和一些具有紫外吸收的杂质。它们对蛋白质和其它生物活性物质有害, 用前应除去, 试验中采用 McPhie<sup>[44]</sup>建议的方法去除杂质, 先用  $50\%$  乙醇慢慢煮沸一小时, 再分别用  $50\%$  乙醇、 $0.01\text{mol/L}$  碳酸氢钠溶液、 $0.001\text{mol/L}$  EDTA 溶液依次洗涤, 最后用蒸馏水浸洗 3 次, 处理好的透析袋贮存在  $4^\circ\text{C}$  蒸馏水中。②透析过程及条件: 用透析袋夹好透析袋一端, 灌注去离子水检查是否有漏水地方, 如果不漏, 则将去离子水倒出, 向透析袋中加入液态植酸酶溶液, 但不能装满, 大约流一半左右的空间, 再用透析袋夹密封好袋口, 用棉线将透析袋悬挂于装有  $600\text{ml}$  去离子水的大烧杯中, 将烧杯放在磁力搅拌器上进行恒速搅拌, 透析 24 小时。

(2) 浓缩: 将透析袋 (装有经过透析的液态植酸酶样品) 置入  $200\text{ml}$  烧杯中, 再在透析袋的表层撒上聚乙二醇 8000 (PEG8000), 以全部覆盖透析袋为准, 然后将烧杯置入  $4^\circ\text{C}$  下过夜。

(3) 离心除杂: 将经过浓缩的植酸酶样品置入离心管中, 以  $12000\text{r}/\text{min}$  离心  $10\text{min}$ , 去除不溶杂质, 取上清液以备用。

### 6. HPLC 色谱条件<sup>[45]</sup>

固定相: ZORBAX 300SB C18 反相液相色谱柱 ( $4.6 \times 150\text{mm}$ ,  $5\mu\text{m}$ ); 流动相 A:  $0.1\%$  TFA 的水溶液, 流动相 B:  $0.09\%$  TFA 的 ACN 溶液; 流速:  $1.0\text{ml}/\text{min}$ ; 进样量:  $10\mu\text{l}$ ; 检测波长:  $254\text{nm}$ 。

### 7. 梯度洗脱条件的确定

梯度洗脱的方法建立可以通过与等度分离的系统方式同样的步骤进行, 梯度方法建立的具体步骤如下:

(1) 以等度分离的相同方式选择初始条件: 色谱柱, 流动相组成, 流速, 温度等 (见表 2-1); 等度方法建立能以强流动相开始 ( $80\sim 100\%B$ ), 但首次梯度试验应采用较宽的梯度范围 (如  $5\rightarrow 100\%B$ );

- (2) 再调节梯度范围以缩短运行时间，去除色谱图开始和结束时无用的空间；
- (3) 如出现谱峰重叠或运行时间太长，则需改变选择性 ( $\alpha$ )；
- (4) 作为备选) 考虑采用非线性梯度方式，以进一步改善分离；
- (5) 优化峰间距后，改变条件以改善分离度和/或运行时间。

表 2-1 反相HPLC方法建立的初始条件<sup>[41]</sup>

Table 2-1 The initial conditions of RP - HPLC

分离变量	最佳的初始选择
柱填料	C8 或 C18; 弱酸性硅胶; 若设计温度>50℃, 选择更稳定的空间保护填料
柱结构	15×0.46cm 柱, 5μm 填料
流速	1.0ml/min
流动相	乙腈-水 (中性样品) 或乙腈-缓冲液 (离子型样品); 对初始试验, 可用 60min 内 5→100%B 的梯度
温度	35 或 40℃
进样量	<50 μl; 50~100 μg

## 8. 结果与分析

文献报道的蛋白质、多肽的分离一般用反相高效液相色谱法，流动相常为乙腈-水-酸（三氟乙酸，TFA）的混合物。实验中以乙腈-水-酸为流动相按不同比例混合等度洗脱时，样品分离效果不理想，故选用梯度洗脱分离样品。本文比较不同植酸酶产品以乙腈（含 0.09%TFA）-水（含 0.1%TFA）流动相系统梯度洗脱分离样品的效果，得到植酸酶产品 A 和 B 的 RP-HPLC 图（图 2-2 所示）。

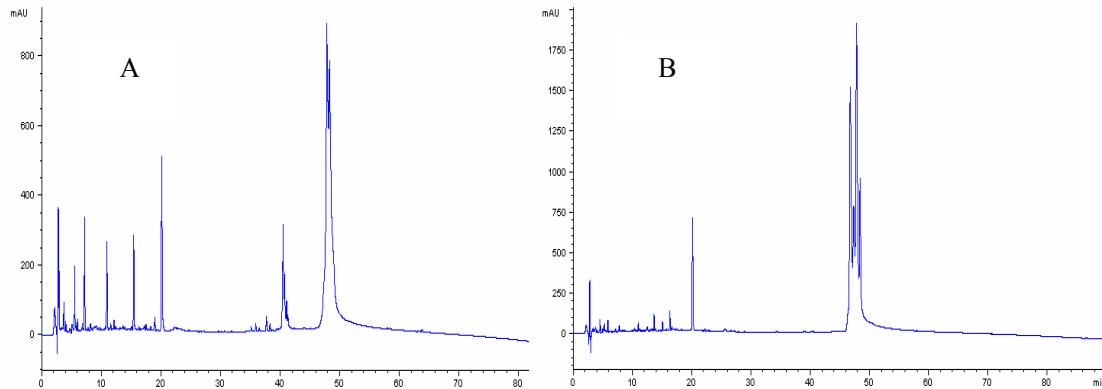


图 2-2 样品 A, B 反相高效液相色谱分析图

Figure 2-2 RP - HPLC analysis of sample A and B

由图 2-2 可知当洗脱浓度 (B%) 设为 5%~95% 时，能够将产品中的主要蛋白成分洗脱出来，但是 20~40min 这段时间为无效保留洗脱时间。因此经过对洗脱条件进行优化（见表 2-2），得到图 2-3。从图 2-3 可以看出通过优化梯度洗脱时间，三个生产企业植酸酶产品的主要组分都能够在反相高效液相色谱图上有所体现。植酸酶蛋白的保留时间一致，并且与植酸酶标品的保留时间一致。

表 2-2 RP-HPLC 优化梯度洗脱表

Table 2-2 Optimization gradient elution table of RP-HPLC

洗脱时间(min)	A%: (H <sub>2</sub> O+0.1%TFA)	B%: (ACN+0.09%TFA)
0	100	0
21	79	21
23	61	39
30	58	42
31	54	46
45	45	55

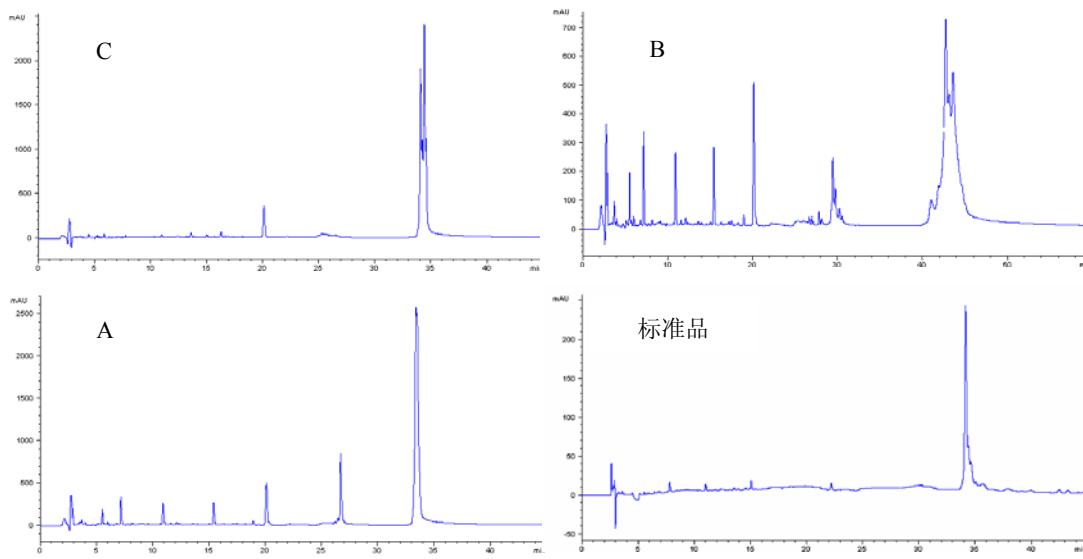


图 2-3 植酸酶反相高效液相色谱洗脱图

Figure 2-3 RP - HPLC analysis of phytase samples

## 9. 讨论

利用 RP-HPLC 可以更加清晰的区分不同品牌植酸酶产品中蛋白质组成特性，三种商品植酸酶中蛋白质种类存在明显差别，而且产品中的主成分植酸酶蛋白的峰形也存在一定差别，直接反映了发酵工艺和纯化工艺对植酸酶产品中蛋白组成的影响。由于 RP-HPLC 对产品中的蛋白质洗脱分离作用是利用蛋白质表面的非极性区域同固定相表面的非极性配体相互作用，按照疏水作用力大小进行分离的一种方法，我们发现尽管三种商品植酸酶的分子量大小与来源于曲霉的植酸酶存在明显差异，但它们的洗脱时间几乎相同，即这些来源于不同微生物的植酸酶蛋白的疏水作用相似，这种现象产生的具体机制值得进一步去探索。

### 2.1.4 小结

综合上述分析可以看出，在适当的梯度洗脱条件下，利用反相高效液相色谱可以清楚地解析植酸酶产品的蛋白组成，同时通过与植酸酶标准品进行对比分析，发现不同生产企业的植酸酶产品由于生产工艺、发酵条件等方面的差异，致使植酸酶产品中的蛋白含量不仅在种类上存在明显差异，而且在与目标蛋白的百分比上也存在明显差异，各个产品有着自己独特的反相高效液相色谱图。通过 SDS-PAGE 蛋白电泳，验证了 RP-HPLC 的试验结果。

## 2.2 植酸酶变性温度特性研究

### 2.2.1 概述

饲料加工过程中的膨化和制粒工序要求植酸酶产品具有较好的热稳定性，植酸酶的变性温度是评价植酸酶品质的重要指标，蛋白质的变性温度与植酸酶产品中的蛋白质组成和纯化工艺相关，另外不同菌种来源的植酸酶变性温度也不同。虽然研究大分子结构的方法很多，但是差示扫描量热分析仪（Differential Scanning Calorimetry, DSC）仍可给出一些与其他方法不同的信息。例如，蛋白质热变性过程中用DSC测定与蛋白质结构变化有关的热效应<sup>[46-48]</sup>，可以分析其结构变化过程，还可以研究影响蛋白质结构稳定性的各种因素。温度变化会对生物分子的结构造成扰动，可引起生物大分子结构的转变。在一定的温度下，不同结构的生物大分子会达到某种平衡状态，这种平衡状态随温度的变化可以提供关于自由能的各种参数，如焓、熵、热容量等信息。系统的焓是温度的函数，从而得到一个关于系统的完备的热力学描述。由于DSC可以提供这些信息，因此它也提供了一种确定蛋白质热力学的可能性。主要是通过DSC测定其变性温度，即酶蛋白的变性温度（主要包括：喷雾型，吸附型），从而研究其在制粒工艺过程中的热稳定特性。

### 2.2.2 差示扫描量热仪分析植酸酶热稳定性

#### 1. 仪器设备

差示扫描量热仪（DSC）：DSC6200，日本 Seiko Instruments 公司。

分析天平：梅特勒-托利多 AL104-IC 系列；

#### 2. DSC 样品制备及操作

将准确称量并密封的样品及标准参照物放入设备中（如图 2-4 左图），放置的顺序为：  
左边标准参照物，右边待测试样（如图 2-4 右图）；

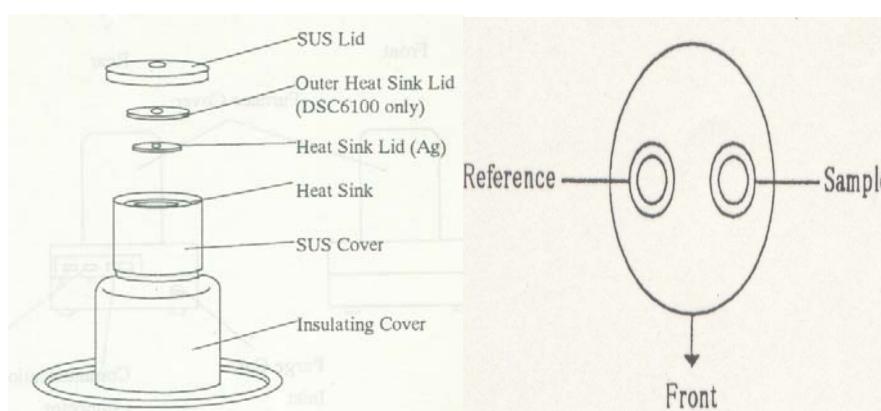


图 2-4 DSC 分析热天平结构解剖图

Figure 2-4 Structure anatomical drawings of DSC

称取 5.0mg 左右的喷雾干燥植酸酶样品于铝盒中，压盖密封。室温放置 30min 后，用差示扫描量热仪测定其变性温度。其次要进行样品的设定以及操作温度的设定：样品的设定

包括样品名称、样品重量、参照物名称、参照物重量，实验结果文件的路径及名称。操作温度的设定包括起始温度、终止温度、升温速率（一般为：10℃/min，食品样品一般为：5℃/min）、保持时间（0min）、取样的频率（一般0.2sec）。

### 3. 差热分析（DSC）

用日本Seiko Instruments公司DSC-6200型差示扫描量热分析仪对植酸酶热变性和热化学变性过程进行了研究。系统配有一个液氮冷却装置，在DSC-6200型DSC的样品室里有两个样品池，一个放样品一个放空白参比。在实验过程中两个样品池之间的热量差被仪器测定，可获得一个热容( $C_p$ )对温度(T)的热力学曲线。具体步骤为：用十万分之一电子天平精确称取5mg植酸酶样品，置铝盒中，然后用配套铝盖密封，以空铝盒为空白参比，室温平衡1 h后将两个铝盒置入样品室中。以10℃/min的加热速率使铝盒温度从30℃上升到150℃。以密封空白铝盒作为对照。每样品重复测定2次。采用配套分析程序Pyris Software记录并计算吸热曲线上的起始温度(Onset Temperature,  $T_o$ )、峰值温度(Peak Temperature,  $T_p$ )、终止温度(Conclusion Temperature,  $T_c$ )和变性时的热焓变化(Enthalpy of Transition,  $\Delta H$ )。所有的实验重复次数不低于3次，并取其平均值作为结果计算。

### 4. 植酸酶的热稳定特性研究

来源于A生产企业和B生产企业喷雾干燥植酸酶产品和植酸酶标准品的DSC曲线参数如表2-3所示：

表2-3 不同植酸酶产品的DSC曲线参数

Table 2-3 DSC curve parameters of different phytase samples

样品	$T_o$ (℃)	$T_p$ (℃)	$T_c$ (℃)	$\Delta H$ (mJ/mg)
标准品	70	102	146	187.46
植酸酶B	65	108	162	216.37
植酸酶A	68	114	158	245.21

变性温度和变性过程中的热焓变化DSC分析植酸酶蛋白变性曲线呈不太典型的V型，且不完全对称。从图2-5中可看出，不同生产企业植酸酶的DSC曲线差异主要表现在，蛋白吸热过程中发生热能“跃变”时的温度、峰值温度和变性时热焓变化的不同。

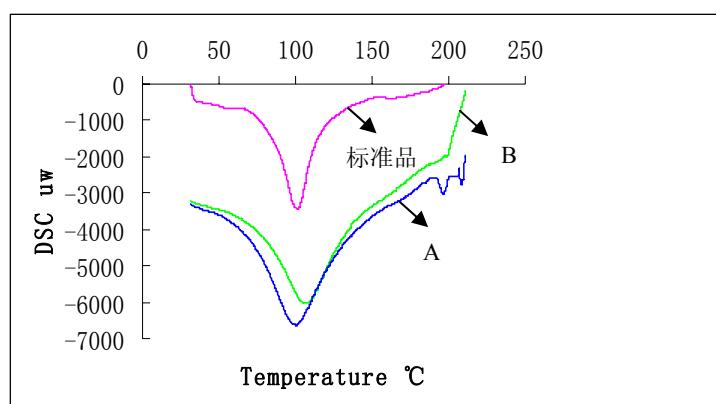


图2-5 喷雾干燥植酸酶产品DSC曲线图

Figure 2-5 DSC curve of spray drying phytase samples

为了获得更多的关于植酸酶变性温度方面的信息,开展了实验部分所描述的差示扫描量热仪实验。图 2-5 所示为植酸酶蛋白随温度变化而发生的吸收热焓值变化图,植酸酶标准品的起始温度是 70℃,峰值温度为 102℃,而喷雾干燥植酸酶样品 B 和 A 的起始温度和峰值温度分别为 65℃,68℃和 108℃,114℃。通过比较植酸酶标准品的 DSC 曲线,与喷雾干燥植酸酶 A 和 B 的 DSC 曲线之间的差异,可以发现,植酸酶 B 在 114℃(植酸酶 A 在 108℃)出现一个转变峰,而植酸酶标准品的变性温度为 102℃,这主要是由于通过基因工程的手段,增加了植酸酶的热稳定性,从而导致其变性温度也随之升高。同时,从表 2-3 可以发现,由于不同生产企业其菌种来源不一样,所以其植酸酶产品的变性温度和吸收热焓也不一样。从图中结果显示植酸酶标准品在 100℃附近有一尖锐吸热峰,天然的植酸酶蛋白是由大小亚基组成单体,由单体进一步构成高聚态。因为维持蛋白高级结构主要是靠亚基之间的疏水键,键能较低,最容易打开。所以,100℃附近的吸热峰应该是由于蛋白质高聚态解聚所致。而植酸酶 B 较高温度区的小吸热峰可能是由蛋白质分子的进一步变构而导致的。

### 2.2.3 小结

本文采用了差示扫描量热法(DSC)分析植酸酶蛋白变性特性,操作中所需样品少,重复性好,所得到的变性温度、热焓变化等参数较全面地反映了植酸酶蛋白变性特性。同时通过与植酸酶标准品进行对比分析,发现不同生产企业的植酸酶产品由于生产工艺、发酵条件等方面的差异,致使植酸酶产品中的酶蛋白的变性温度和吸收热焓都存在着明显的差异。这也说明某一特定形式的植酸酶产品对应特定的变性温度和吸收热焓。因此,可以利用差示扫描量热仪这一独特的角度,鉴别植酸酶产品的真伪。而且 DSC 分析结果与 SDS-PAGE 电泳、RP-HPLC 的分析结果存在一致性,以上三种技术分别从不同侧面说明,不同厂家的植酸酶产品由于种属来源、加工工艺、纯化工艺等存在差异,从而导致其植酸酶产品的蛋白组成、变性温度、热稳定性也存在明显差异。

### 第三章 植酸酶产品近红外指纹图谱库的建立

#### 3.1 近红外光谱分析技术测定原理

##### 3.1.1 近红外光谱分析的基础

NIR分析的化学基础：近红外谱区是电磁波谱中很窄的区段。正如我们所见到的不同的物质在可见光区段显示不同的颜色一样，不同的物质在近红外光谱区段也具有不同的近红外光谱。NIR谱区主要是含H基团如：C—H，N—H，O—H等伸缩振动的各级倍频和这些基团的伸缩振动与弯曲振动的合频吸收，这是NIR定量分析的化学基础<sup>[49]</sup>，而这些基团是组成有机物的一些最重要基团，NIR谱区丰富的结构与组成信息构成了NIR分析技术广泛应用的基础。由于分子合频的组合方式很多，NIR谱区谱峰重叠非常严重，谱峰比较宽，一般无法确定谱峰对应的基团；另外由于倍频和合频跃迁需要的能量较高，这两种跃迁出现的几率比上一级（或基频）出现的几率低一个数量级左右，吸收的强度比中红外的基频要弱很多，可以穿透样品内部，如果用传统的光谱分析技术，NIR谱区将无法进行分析；但是计算机技术和统计算法的应用使NIR可以分析测定与这些基团有关的成分以及物理、化学性质，如物质的密度、粘度、颗粒的大小以及有关样品的电学、热学、力学性质等相关指标。

NIR分析的物理基础：光的反射有两种，一种是镜面反射，另一种是漫反射。所谓镜面反射，就是光线照射到光滑物体表面上被有规则地反射出来的现象，它未与样品内部相互作用，没有携带入射光与样品相互作用的信息。因此，NIR进行固体样品的定量分析时，一般采用漫反射技术，避免镜面反射。所谓漫反射，就是光线照射到粗糙物体表面被无规则反射的现象。当光线照射到由一定厚度颗粒物质组成的样品层时，一部分被吸收，一部分被样品内部各层次反射、吸收、折射后最终反射出来，反射出来的光线反映了样品的内部特性，更多地携带有样品的信息。在样品具有相当厚度的情况下，用近红外光照射有机物，样品的透射光可忽略不计，入射光主要被样品吸收或是被样品反射<sup>[50-51]</sup>。依据“Kubelka-Munk”，定律，固体样品的漫反射吸光度A为：

$$A = -\log(1 + K/S - [(K/S)^2 - 2(K/S)]^{1/2}) \quad (3-1)$$

其中：A 为固体样品的漫反射吸光度；

K 为固体样品的吸收系数，它决定于固体样品的化学组成；

S 为固体样品的散射系数，它决定于固体样品的物理特性（如样品的粒度等）。

当待测组分的含量在一定范围内时，A 与 K/S 可用一截距不等于零的直线来拟合。

$$A = a + b(K/S) \quad (3-2)$$

当散射系数 S（样品物理状态）保持不变时，吸收系数 K 与样品组份含量 C 成正比，

即：

$$A = a + b \times C \quad (3-3)$$

也就是说，漫反射吸光度A与样品组份含量C成线性关系。但要注意，只有在样品浓度较低、散射系数S保持不变时，这一关系式才成立。漫反射吸光度是近红外漫反射光谱分析中最常用的光谱参数<sup>[52]</sup>。

NIR 分析的数学基础：NIR 分析的数学基础主要是化学计量学，而化学计量学是综合使用数学、统计学和计算机科学等方法从化学测量数据中提取信息的一门新兴的交叉学科。大量化学计量学方法被写成软件，并成为分析仪器（尤其是近红外光谱仪）的重要组成部分。近红外光谱定量分析的基础模型涉及多元线性回归模型和多元线性回归模型最小二乘估计。近红外光谱定量分析模型的评价涉及回归方程和回归系数的显著性检验，回归模型预测效果的检验，和模型异常点与高杠杆点的判定等。而优化定量分析模型的逐步回归法则涉及逐步回归过程，Sweep 变换，变量的选入和剔除变量的显著性检验和逐步回归的计算步骤等数学基础方面的内容。近红外光谱定量分析中主要应用多元校正的方法，包括多元线性回归法、逐步回归法、主成分回归法和偏最小二乘法等一些涉及化学计量学的方法。在近红外光谱定性分析中应用到数学中的聚类方法、贝叶斯判别分析的方法和人工神经网络的方法。

### 3.1.2 近红外光谱中的化学计量学方法

化学计量学是综合运用数学、统计学和计算机科学等方法从化学测量数据中提取信息的科学。光谱化学计量学方法研究在现代近红外光谱分析中占有非常重要的地位，近红外光谱化学计量学方法主要涉及以下几个方面：

#### 3.1.2.1 光谱预处理方法

在测量得到的近红外光谱数据中，不可避免地含有各种各样的噪声。这些噪声的存在或大或小地影响了数据的分析和模型的建立。光谱预处理方法的目的是净化谱图信息，剔除异常样品、消除光谱噪声及其他干扰因素影响。为减少甚至消除各种噪声对光谱的干扰，需要采用合适的方法对光谱进行预处理，提高所建立的近红外光谱模型预测的准确性和精度，常用的预处理方法主要有以下几种：

##### (1) 平滑处理

常用的平滑方法主要有平均窗口（Moving Averages）平滑方法、中位值平滑方法、Savitsky-Golay平滑方法等<sup>[53]</sup>。Savitsky-Golay平滑方法是目前使用最广泛的去除噪声的平滑方法，采用多项式在最小二乘法下拟合原数据，采用Hamming窗或Hanning窗来代替平均窗口平滑方法中的矩形窗口，从而不会造成峰形失真，该方法还可以用于数据的微分<sup>[54]</sup>。

##### (2) 散射校正

在近红外光谱分析中，测量对象的固体颗粒度、晶形等物理性质的不同，也会导致谱图出现差异。这种差异是进入固体内部的散射光经过的光程和被吸收程度不同而引起的。消除

散射效应最常用的两种方法是多元散射校正（MSC）和标准正态变换（SNV）。多元散射校正是用来消除样品表面性质差异和颗粒尺寸大小不同所带来的光谱漂移常用的一种方法<sup>[55-56]</sup>。假设所有样品具有相同的散射系数，那么就可以用一个理想光谱（通常用平均光谱代替）来消除由于光的散射对样品近红外光谱的影响。标准正态变换法<sup>[57]</sup>也是一种用来消除样品尺寸与均匀性对近红外光谱影响的方法，是将光谱数据矩阵按行进行归一化而消除散射效应。

### （3）基线校正

基线校正是化学计量学中一个很古老的话题，常用的基线校正方法主要为以下几种：①常偏移量消除法（Constant Offset Elimination, COE），即在选择的波段区域里所有光谱减去最低的Y值，用来消除基线偏移。②直线相减法（Straight Line Subtraction, SLS），是一种典型的基线校正方法，在每一个被选中的波段用最小二乘法拟合一条直线，然后从光谱中减去该直线，达到校正基线的目的。

### （4）微分

微分是光谱分析中一个非常经典也很常用的方法。它常被用来增进复杂光谱中低分辨率组分的分辨率，也可以应用于基线校正<sup>[58]</sup>。常用于基线校正的一阶微分和二阶微分。一阶微分可以消除光谱中的基线漂移，二阶微分可以消除光谱中基线的倾斜。但微分在除去低频的基线的同时，还会放大高频的噪音。因此，在微分之前，通常都进行平滑处理以滤去噪音。

### （5）小波变换

小波变换（Wavelet Transform, WT）是近年来新出现并不断获得推广应用的一种高效信号处理方法<sup>[59]</sup>。该方法可把各种频率组成的混合信号按不同分辨尺度分解成一系列不同频率的块信号。根据这一特点，依先验知识可对特殊频率范围的噪声或背景进行滤波处理，如高频随机噪声的滤波，可设定一个阈值，对高频部分的信号进行重构，即可滤掉随机噪声；同样，也可以对噪声或漫漂移进行扣除。如果需要，加以特殊“修饰”，甚至可对倍频等其它干扰进行平滑处理。小波变换技术应用面广，除过滤噪声之外，还可用于特征提取和光谱数据压缩。

### （6）正交信号分解

正交信号校正（Orthogonal Signal Correction, OSC）是1998年Wold等提出的一种新的预处理方法<sup>[60]</sup>，在近年得到了非常广泛的应用。它的基本思想是利用正交化方法，除去光谱中与预测值无关的部分，从而得到“纯净”的光谱。

## 3.1.2.2 波长选择方法

在近红外光谱分析校正模型建立过程中，若采用全谱计算时，计算量很大。实际上，在某些光谱区域，样品的光谱信息很弱，与其样品指标性质间缺乏相关性。研究证明，选择适当的光谱区域，可以有效的提高预测模型的精度<sup>[61]</sup>。

在早期，人们往往根据化学信息来选择适当的光谱区域。这类方法很容易理解，也很符合分析化学工作者一贯的思维。然而有时候凭简单的化学知识，很难确定适合的光谱区域，于是就发展出来了很多种计算的方法。这些计算方法主要分为两类：一类是简单的基于阈值的判定方法，另一类是将波长选择看成一个组合优化的模型，通过一些搜索方法来选择最佳的波长子集。基于阈值的方法比较简单，它们往往是基于相关系数，或者基于信噪比，或者是某种巧妙设计的阈值指标。但阈值方法往往很难得到一个最优的结果。基于搜索的方法主

要有两类，一类是逐步算法，速度很快，但很容易陷入局部最优。另一类是全局优化算法，能找到最优解，但耗时很长。搜索算法一般均以波长选择之后建立的偏最小二乘法（PLS）或主成分回归（PCR）模型的预测平均偏差（RMSEP）作为优化的目标函数。常用的逐步算法主要包括顺序前进选择、顺序后退选择、逐步选择三种方法。常用的全局搜索方法有遗传算法（Genetic Algorithm, GA）和多链方法（Multiple Chain Method, MCM）等，其中 GA 算法的研究应用较为广泛。

### 3.1.2.3 定性分析

光谱的定性分析主要是依靠已知样品和未知样品的光谱数据比较来完成。已有的一些光谱的比较方法可以用于近红外光谱的定性分析。由于近红外的谱带较宽且灵敏度较差，吸收峰重叠严重，从人眼考察近红外图的相似性难以判别，因此定性分析一般须借助计算机模式识别方法方可完成，常用的模式识别方法有如下几种：

#### （1）相似系数法

相似系数法是最广泛使用的相似性判别方法，它通过求取已知的平均光谱与样品光谱特征值之间的夹角余弦值。理论上，若两个谱图完全相同，则相似系数为 1；但由于测量误差和噪音的影响，相似系数总是接近 1。使用相似系数进行判别分析的优点是，它与谱图个数以及浓度变化无关。

#### （2）马氏距离法

马氏距离法可以成功用于光谱的聚类分析，马氏距离法对校正集数据的内部变化敏感，它不仅给出较好地分类结果，而且也能给出能保证测量结果优劣的统计结果<sup>[62]</sup>。马氏距离法计算较直观，但是，参加计算马氏矩阵的光谱数据正确与否将直接影响结果。

#### （3）主成分分析结合马氏距离法

主成分分析（PCA）法可将光谱矩阵进行降维，得到光谱的主成分得分数据。首先对多维光谱矩阵用主成分分析进行降维处理，再用各光谱的主成分得分数据计算马氏距离，这样，既利用了主成分分析降维不丢失信息的优点，同时又可以利用马氏距离便于确定定量阈值的优点。通过对主成分的选择可以有效地选择代表样本的特征，滤除干扰信息。使用主成分得分代替原始光谱来计算马氏距离，不仅能反映全谱数据信息，而且能压缩参加计算马氏距离的变量数目，并能避免出现共线性问题。

#### （4）SIMCA 法

SIMCA（Soft Independent Modeling of Class Analogy）法是基于类模型基础上的有监督的模式识别方法<sup>[63]</sup>。该法基于这样的假设：同一类样本具有相似的特征，在一定的特征空间内，属于同一类的样本就会聚集在某一特定的空间区域，而对不同类的样本，则分布在不同的区域，通过因子分析法分别针对训练集中的每类样本建立类模型。类模型建好后，对预测集中的样本，计算其到各类模型的 SIMCA 距离，从而判别该样本属于哪一已知类，或同时属于哪一已知类，或归于新类。

#### （5）人工神经网络

人工神经网络（ANN）具有突出的非线性映射能力，已发展为化学信息学方法中对非线性体系进行校正的重要方法之一<sup>[64]</sup>。近年来在近红外光谱模式识别领域也得到广泛应用。

### 3.1.2.4 定量分析

光谱仪器实际采集的光谱数据不仅包括了被测样品的组成和结构信息，而且还包括了噪音，即测量误差与不同组分之间的干扰等。化学计量学方法可有效地剔除这些噪音。化学计量学方法通过数学方法对原始光谱处理，得到光谱的主成分和得分，根据一定的规则选取一定数目的主成分光谱重建光谱，该重建光谱最大限度地反映了被测样品的组成和结构信息，而最小限度地包含噪音，能够克服经典方法的缺点，并保留了其优点。近红外光谱分析常用的化学计量学方法为多元校正方法<sup>[11]</sup>，包括多元线性回归（MLR）、主成分分析、主成分回归、偏最小二乘法、拓扑学方法和人工神经网络方法等。

光谱定量分析方法的核心在于建立稳定、可靠的定量分析模型。早期的近红外定量研究较常使用 MLR 方法，因当时的光谱仪是采用滤光片分光，波长数相对较少。当新一代近红外光谱仪可以在近红外全部波段采集光谱后，PCR 与 PLS 方法较 MLR 方法获得更广泛的应用。由于仪器的非线性响应，以及固体样品粒子尺寸不均匀，线性校正方法通常很难发挥作用，这时须使用人工神经网络（ANN）、局部权重回归（LWR）和 PCR 与 PLS 的二次多项式回归等非线性建模方法。由于本论文主要采用了 PLS，下文将简要介绍该分析方法。

偏最小二乘法是将因子分析和回归分析结合起来，主要用于建立多因变量与多自变量的统计关系。不仅对光谱矩阵进行主成分分解，而且对浓度矩阵也进行主成分分解，然后用光谱主成分得分矩阵与浓度主成分得分矩阵进行回归的方法。在回归分析中，它组合了经典最小二乘法（CLS）的平均效应与逆最小二乘法（ILS）的部分组分回归特性；分解与回归在一步中完成，模型更稳健；保证了主成分一定与感兴趣组分浓度相关，是在与感兴趣组分浓度最相关的方向投影，而不是简单的在方差变化最大的方向投影；只需要知道少数几种我们感兴趣的组分浓度，适于复杂混合物的组分浓度预测。但自计算慢，模型复杂，需要大量样品建模，必须避免组分浓度共线性等问题。

PCR与PLS建模关键在于如何确定主成分数目（f）。随着主成分数目增加，载荷向量对建模的重要程度将减小（PCR是方差减小，PLS是权重减小），到一定程度，载荷向量将变成模型的噪声。如果F过大会出现过拟合现象，即将一些代表噪声的主成分加入模型，使模型预测能力下降，反之，如果f过小会出现欠拟合现象，即太少的主成分不足以反映未知样品被测组分产生的光谱变化。目前通常采用的方法是PRESS法。一般取PRESS最小时的主成分作为建模最佳主成分f<sub>opt</sub>，计算公式如下：

$(1) \text{PRESS} = \sum_{j=1}^m (\text{Differ}_j)^2$	$(2) f_{opt} = f_k / \min(\text{PRESS}_{fk})$	(3-4)
---	---	-------

### 3.1.3 评价数学模型效果的指标

近红外预测模型效果评价指标包括：使用定标集决定系数（R<sup>2</sup>）和交互验证均方差（RMSECV）分析模型的拟合能力，用预测均方差（RMSEP）评判模型的预测能力，以及综合交互验证决定系数（1-VR）、预测决定系数（r<sup>2</sup>）、相对标准差（RSD）等指标进行评定。分别用公式表示为：

(1) 决定系数R<sup>2</sup>

$$R^2 = \left(1 - \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_m)^2}\right) \times 100\% \quad (3-5)$$

n: 样品个数;

$\hat{y}_i$ : 第i个样品的预测值;

$y_i$ : 第i个样品的参考值(真值);  $\bar{y}_m$ : 样本真值的平均值;

R<sup>2</sup>接近100%表示预测浓度值接近真值,若R<sup>2</sup>=1则说明存在完全拟合。当预测残差平方和大于总平方和(真值的方差),R<sup>2</sup>成为负值,表示拟合效果极差。

## (2) 交互验证标准差(RMSECV)

$$RMSECV = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (3-6)$$

$\hat{y}_i$ : 第i个样品的预测值;

$y_i$ : 第i个样品的参考值(真值);

## (3) 校正标准差(RMSEC)与预测标准差(RMSEP)

$$RMSEC = \sqrt{\frac{1}{n-R-1} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (3-7)$$

R: PLS主成分个数;

$\hat{y}_i$ : 第i个样品的预测值;

$y_i$ : 第i个样品的参考值(真值);

$$RMSEP = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (3-8)$$

## (4) 相对标准差(RSD)

相对标准差RSD%是由校正标准差(RMSEC)与预测标准差(RMSEP)除以各自样本均值 $y_{mc}$ 和 $y_{mp}$ 得到的。它反映模型对某一组分的总体预测效果。

校正相对标准差: RSD% = RMSEC/ $y_{mc}$  × 100%

预测相对标准差: RSD% = RMSEP/ $y_{mp}$  × 100%

### 3.1.4 奇异点的判别与处理

奇异点是指落在总体之外的数据点,主要是由于实验条件的改变、样品性质变化以及仪器测量误差和人为测量误差导致。近红外的计量方法是根据已知样品建立拟合模型,而且通常为线性模型,如果某样品的光谱与定标样品差别较大,则该样品也是异常样品。一类异常样品产生的原因是误操作或仪器异常,发现后可以通过简单地重新扫描得到纠正;而另一类

异常样品是由于样品本身所产生，不能简单地通过重新扫描得到纠正，该样品的预测值是否可靠取决于其光谱变异与模型的拟合程度。在使用主成分回归(PCR)或偏最小二乘法(PLS)对近红外光谱数据建立光谱预测模型时，常常会出现浓度预测值与标准值间的相关系数很低的情况，其主要原因是校正集样品中混有不适合放在一起参与建模的异常样，一般分为光谱异常和化学值异常，异常值的判别和处理是提高近红外分析质量的一个重要步骤<sup>[65]</sup>。

校正集中奇异点剔除最常见的方法主要包括：主成分得分图、预测残差图、马氏距离法、以及杠杆值与学生T检验等，其依据原理都是判断某统计量是否超过一定的分布下（如正态分布等）的临界值。如果光谱数据中只有一个奇异点，这些方法都很有效，但对于多个奇异点的情况，一般都没有满意结果，主要原因是多奇异点改变了数据的重心和离散度。

### 3.1.5 近红外光谱分析技术流程

本研究所采用的近红外光谱分析流程图如图 3-1 所示：

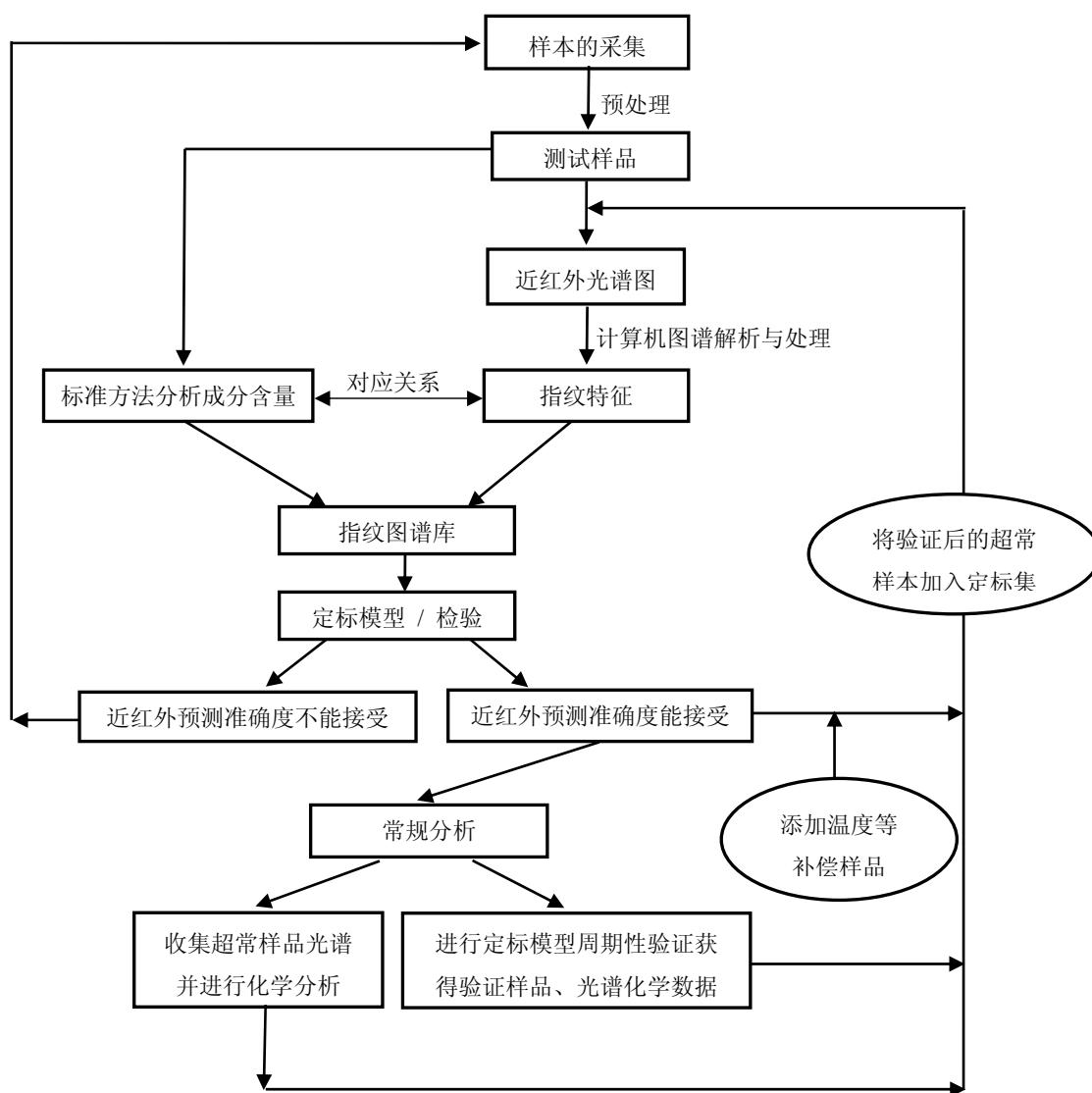


图 3-1 近红外分析技术流程图

Figure 3-1 Flow chart of the NIR analysis

## 3.2 植酸酶产品近红外指纹分析图谱的构建

### 3.2.1 指纹图谱库的构建原则

指纹图谱必须能反映出产品的“共性”及“个性”。共性是指能体现出不同生产企业、不同加工工艺、不同批次的同一种产品共同具有的某些特征（共有峰）。“个性”是指能反映出该产品特有的结构和组成信息，以及能反映出同种产品由于生产企业或由于加工工艺不同造成的差异，对于成品，则能反映出同一产品不同批次间的差异。近红外光谱是一种权威的鉴定方法<sup>[66]</sup>，这是对纯物质而言，对于多组分组成的混合物，那么它的近红外光谱是混合物内各组分近红外光谱的叠加。只要其内在的各种化学组分和量相对稳定，并且样品的处理按统一要求进行，则其近红外光谱应该是相对稳定的，也就是说，其近红外光谱应具有一定的客观性和可重复性，如果使用过程中掺假，则其成品的近红外光谱与正品的近红外光谱图有明显差异。因此，近红外指纹图谱库技术可用于产品品质的溯源管理。建立近红外指纹图谱的基本程序为：(1)研究分析产品参数，收集并确证标准品；(2)采集标准产品光谱图；利用化学计量学的方法剔除异常样品，并对获得的图谱库进行比较分析，寻找分辨率高，重现性好的图谱；(3)进行图谱辨认和评价，找出具有规律性的“共性”并加以量化；(4)对所获得的结果进行校验；(5)建立指纹图谱库。

指纹图谱属于质量标准中的“鉴别项目”，一旦建立，就是终身伴随产品的检验标准，因此必须保证其可行性、重复性和实用性。构建中应注意以下几个问题：(1)设立对照。设立对照是构建指纹图谱的主要原则之一，可将大宗进口、国内大型企业生产的产品作为标准对照，构建该品种的“标准指纹图谱”；(2)样本应具有代表性。若用于鉴别某种产品的真伪，所选的样本应包括不同批次、不同加工工艺条件下的产品，确保所选样本能代表该品种的总体；(3)样本要适当。样本量越大，越接近样本总体的真实情况，代表性越高，结论更可靠，但花费及工作量越大。样本量小，则易发生偏差。因此，样本量要适当。

### 3.2.2 实验样本的收集

根据指纹图谱库的构建原则，选取国内两家大型植酸酶生产企业的产品作为标准对照，构建该品种的标准指纹图谱。连续3个月收集A生产企业和B生产企业不同批次的喷雾干燥和吸附干燥植酸酶产品，共收集到A生产企业的134个喷雾干燥植酸酶样品和121个吸附干燥植酸酶样品，共收集到的B生产企业88个喷雾干燥植酸酶样品和72个吸附干燥植酸酶样品，所选样本能代表该产品的总体，因此所收集的样品具有广泛的代表性。

### 3.2.3 近红外反射光谱的采集

图3-2为A生产企业134个喷雾干燥植酸酶样本的近红外反射光谱图。横坐标为光谱波长(nm)，纵坐标为与本底相比的相对吸光度A，即 $A = -\log(I/I_0)$ ，其中I为样本的反射光强度， $I_0$ 为本底的反射光强度。图3-3为A生产企业和B生产企业喷雾干燥植酸酶样品的原始光谱图和一阶导数光谱比较图。由于近红外光谱区与有机分子中含氢基团(C-H、O-H、N-H)

振动的合频与各级倍频的吸收一致，因此通过扫描样品的近红外光谱，可以得到样品中有机分子含氢基团的特征振动信息。无论是喷雾干燥植酸酶还是吸附干燥植酸酶产品中含有各种含氢基团，所以通过对近红外光谱进行分析可以测定其组分的含量。由于篇幅有限，A生产企业和B生产企业的吸附干燥植酸酶采集近红外光谱图略。

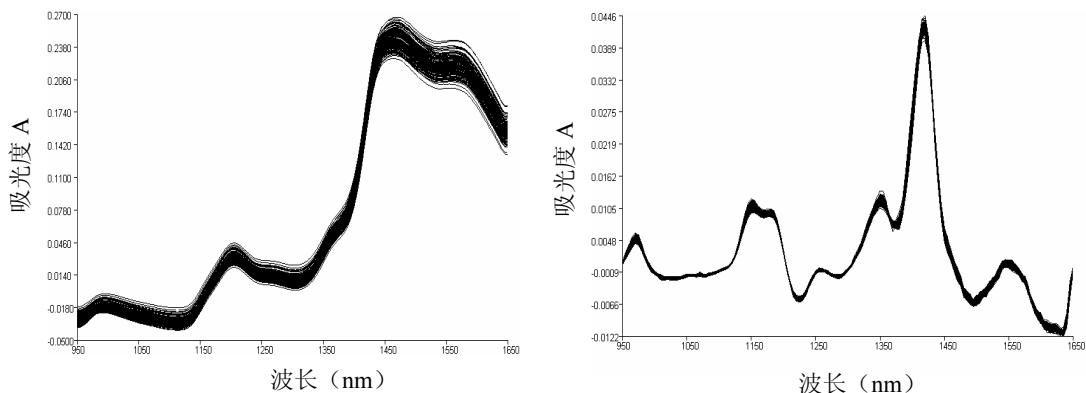


图3-2 A喷雾干燥植酸酶近红外原始光谱图和一阶导数光谱图

Figure 3-2 NIR original and first derivative spectra of spray drying phytase from manufacture A

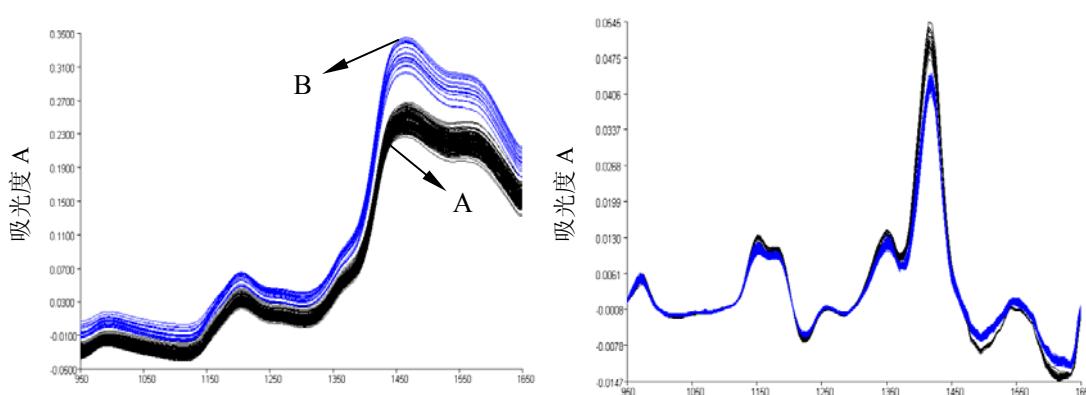


图3-3 A、B喷雾干燥植酸酶近红外光谱图

Figure 3-3 NIR spectra of spray drying phytase from manufacture A and B

### 3.2.4 异常样本的剔除

本研究采用Grams32PLS+IQ化学计量软件中的偏最小二乘回归法（PLS）和交互验证法（Cross-validation）对A生产企业喷雾干燥和吸附植酸酶产品光谱进行异常值检验。计算各个成分的浓度残差值（Concentration Residual）、光谱残差值（Spectral Residual）、光谱杠杆值（Sample Leverage）和学生氏残差值（Student Residual），得到各化学成分异常值检验结果见图3-4，图3-5所示（仅列出部分插图以说明方法）。光谱残差值和光谱杠杆值的阈值分别设为样本平均误差的2倍，如果大于该阈值，则可判断该样本光谱异常。浓度残差值和学生氏残差的阈值分别设为样本平均误差的3倍，如果大于该阈值，则可判断该样本化学值异常<sup>[67]</sup>。

从图3-4中可以看出，11#，45#样本浓度残差异常而光谱残差正常，这些是属于好的样本；而36#样本的浓度残差正常，而光谱残差异常，属于“坏”的异常样本，应从样本集

中剔除。图 3-5 为吸附干燥植酸酶酶活异常值检验结果。从图 3-5 中可以看出，11#，58#样品的学生氏残差值异常，而光谱杠杆值正常，故应予保留，而 43#样品的学生氏残差值正常，而光谱杠杆值异常，故应剔除。总之，异常样品的剔除，是综合多方面的综合考虑。最后用于构建 A 生产企业喷雾干燥产品近红外指纹图谱库的样本只有 132 个，A 生产企业吸附干燥产品的近红外指纹图谱库的样本只有 118 个。

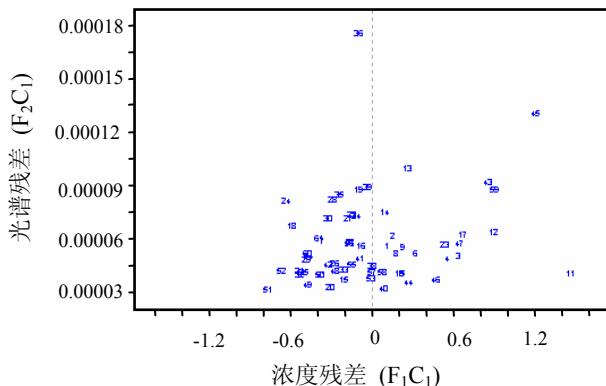


图 3-4 喷雾干燥植酸酶浓度残差 Vs. 光谱残差

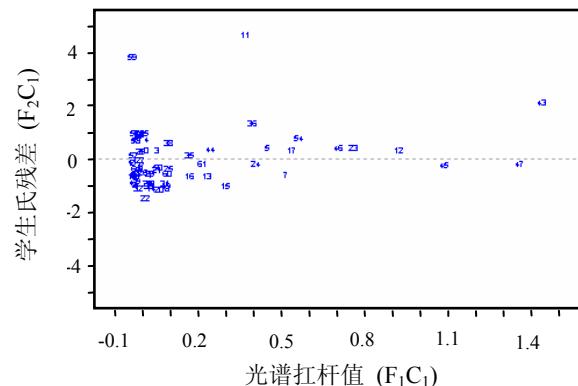


图 3-5 吸附干燥植酸酶杠杆值 Vs. 学生氏残差

Figure 3-4 Concentration residual vs. Spectral residual of spray drying phytase samples

Figure 3-5 Sample leverage vs. Student residual of adsorption drying phytase samples

### 3.2.5 近红外光谱特征信息的提取

在近红外结合偏最小二乘法 (PLS) 建模中，传统观点认为PLS具有较强的抗干扰能力，可全波长参与多元校正模型的建立<sup>[68]</sup>。随着对PLS方法的深入研究和应用，通过特定方法筛选特征波长或波长区间有可能得到更好的定量校正模型<sup>[69]</sup>。波长选择可以简化模型，更主要的是由于不相关或非线性变量的剔除，可以得到预测能力强、稳健性好的校正模型。按实验方法，在950–1650nm范围内扫描A生产企业喷雾干燥和吸附干燥植酸酶得到的近红外反射光谱在不同波长段的相关系数和决定系数图（见图3-6和图3-7）。

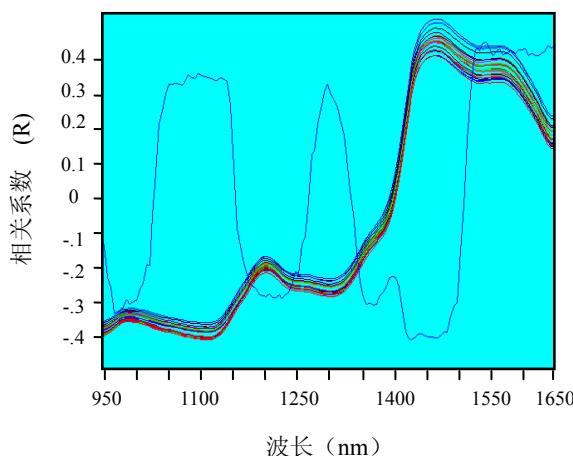


图3-6 喷雾干燥在不同波长处的相关系数

Figure 3-6 The correlation coefficient at different wavelengths of spray drying phytase

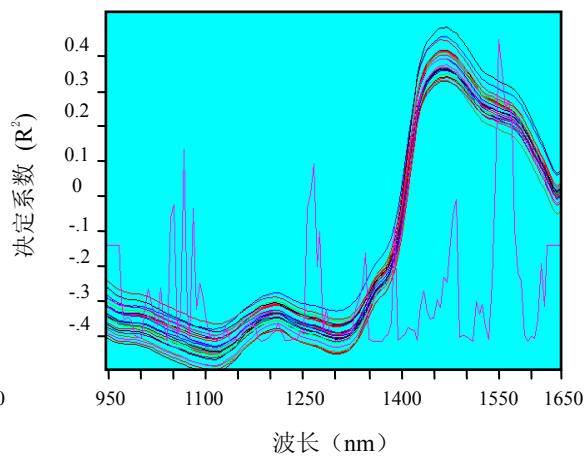


图3-7 吸附干燥在不同波长处的决定系数

Figure 3-7 The coefficient of determination at different wavelengths of adsorption drying phytase

目前，在多元校正分析中，波长选择方法主要有相关系数法（决定系数法）、方差分析法、逐步回归法、遗传算法等。相关系数法（决定系数法）和方差分析法相关系数法是将校正集光谱阵中的每个波长对应的吸光度向量 $X_i$ 与浓度阵中的待测组分浓度向量 $Y_i$ 进行相关性计算，得到波长-相关系数（决定系数法） $R$  ( $R^2$ ) 图<sup>[70]</sup>。对应相关系数越大的波长其信息应越多，因此，可结合已知的化学知识给定一阈值，选取相关系数大于该阈值的波长参与模型建立。方差分析法是通过对校正集光谱阵在各波长下的方差分析，得到波长-标准偏差图，对应标准偏差越大的波长，其光谱变动越显著。与相关系数法相似，可给定一阈值来选择波长区间。由于方差分析法不是针对待测组分优化选取波长，一般不用于定量分析模型，但该方法特别适合于定性分析中的波长选取。不同研究领域的工作者<sup>[71-72]</sup>采用相关系数法选取波长，建立了NIR常规成分的定量模型，与全谱方法相比，各组分的RMSEP分别下降了。值得注意的是，由于相关系数法是基于线性统计方法建立的，对于非线性相关及校正集样本分布不均匀问题，通过该方法选取的结果往往不可靠。

通过光谱特征信息的提取和分析，得到A生产企业植酸酶产品的酶蛋白谱峰和水分谱峰归属如表3-1与表3-2所示<sup>[73]</sup>。

表 3-1 A 植酸酶蛋白谱峰归属

Table 3-1 Protein peaks attribution of phytase from manufacture A

蛋白峰（波长 nm）	谱峰归属
910	C-H 伸缩振动，蛋白
975	O-H 伸缩振动， ROH
1014	二级C-H伸缩振动+三级C-H扭转， CH <sub>3</sub>
1062	N-H伸缩振动， RNH <sub>2</sub>
1092	二级 C-H 伸缩振动+二级 C-C 伸缩振动
1147	C-H伸缩振动， CH <sub>3</sub>
1183	C-H伸缩振动， CH <sub>3</sub>
1219	C-H伸缩振动， CH <sub>2</sub>
1372	二级C-H伸缩振动+C-H扭转， CH <sub>3</sub>
1401	二级C-H伸缩振动+C-H扭转， CH <sub>2</sub>
1428	N-H伸缩振动， CONH <sub>2</sub>
1461	N-H伸缩振动， CONH <sub>2</sub>
1497	N-H 伸缩振动， NH
1530	N-H伸缩振动， RNH <sub>2</sub>
1559	N-H伸缩振动， CONH <sub>2</sub>
1584	O-H 伸缩振动
1616	C-H伸缩振动， =CH <sub>2</sub>

表 3-2 A 植酸酶水分谱峰归属

Table 3-2 Water peaks attribution of phytase from manufacture A

水分 (波长 nm)	谱峰归属
969	O-H伸缩振动, H <sub>2</sub> O
976	O-H伸缩振动, H <sub>2</sub> O
1005	O-H伸缩振动, H <sub>2</sub> O
1033	O-H伸缩振动, H <sub>2</sub> O
1168	O-H伸缩振动, H <sub>2</sub> O
1204	O-H伸缩振动, H <sub>2</sub> O
1239	O-H伸缩振动, H <sub>2</sub> O
1210	O-H伸缩振动, H <sub>2</sub> O
1415	O-H伸缩振动, H <sub>2</sub> O
1454	O-H伸缩振动 <sup>1</sup> , H <sub>2</sub> O
1490	O-H伸缩振动 (分子内的氢键) <sup>2</sup> , H <sub>2</sub> O
1529	O-H伸缩振动 (分子内的氢键) <sup>2</sup> , H <sub>2</sub> O
1578	O-H伸缩振动 (分子间的氢键) <sup>1, 3</sup> , H <sub>2</sub> O
1618	O-H伸缩振动, H <sub>2</sub> O

注: 1. 同样包含淀粉的-OH 基团; 2. 同样包含纤维素的-OH 基团; 3. 同样包含糖的-OH 基团。

Note: 1. Similarly containing -OH groups from starch; 2. Similarly containing -OH groups from cellulose; 3. Similarly containing -OH groups from sugar.

### 3.2.6 植酸酶产品指纹图谱库的建立

通过研究发现, 最终反映植酸酶产品好坏优劣的指标为植酸酶的酶活, 因此针对这一参数和条件, 本小节以 A 生产企业喷雾干燥和吸附干燥植酸酶产品为例构建其指纹图谱库。对所获得的图谱库进行比较分析, 应用图谱分析软件 (Managespetra) 和化学计量学软件 (Grams32PLS+IQ) 中的浓度残差值、光谱残差值、光谱杠杆值和学生氏残差值剔除指纹图谱库中的异常点, 寻求分辨高, 重现性好的图谱 (如图 3-2 中所示)。从图 3-2 可以看出, 对于同一剂型的植酸酶产品, 由于载体相同, 所以存在群体共有的特征, 这样就在光谱图上表现出相同的特征峰。从图 3-3 可知, 由于菌种来源、发酵条件、加工工艺、储存方式的不同而在同一剂型的植酸酶产品中产生较为明显的差异, 在其指纹图谱上也会反映出不相似性。图 3-8 为 A 生产企业喷雾干燥和吸附干燥植酸酶产品的近红外指纹图谱库, 图 3-8(1)、(3)为喷雾干燥植酸酶产品的近红外指纹图谱库; 图 3-8(2)、(4)为吸附干燥植酸酶产品的近红外指纹图谱库。

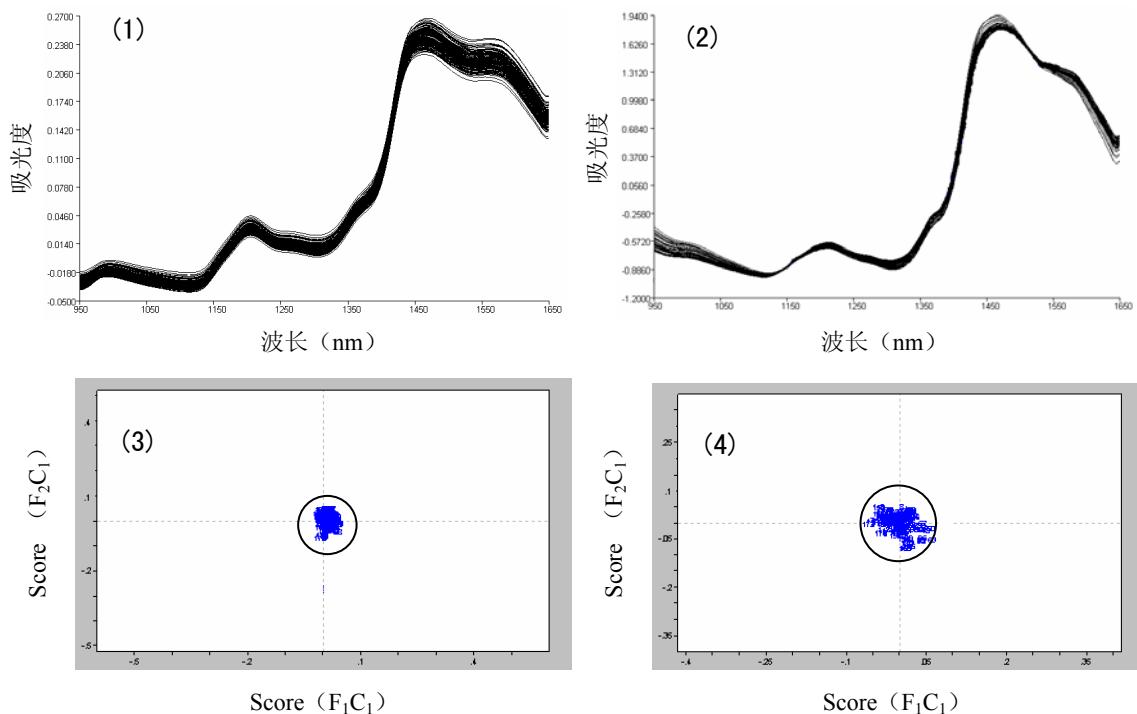


图 3-8 A 植酸酶近红外指纹图谱库

Figure 3-8 NIR fingerprinting database of phytase from manufacture A

针对 A 生产企业的喷雾干燥和吸附干燥植酸酶这一特定产品，采用化学计量学的方法剔除异常样本后，选择 MSC 做散射校正处理、一阶导数和平滑组合为光谱预处理方法，在最优的波数范围内，采用偏最小二乘法回归的方法（PLS-1）结合马氏距离判别法从而建立了植酸酶产品的指纹图谱库。根据产品的特定性质，设定一个马氏距离判定接受的阈值，如果马氏距离大于设定的阈值，则判为不合格，小于设定的阈值就可判为合格产品。本研究中设定马氏距离判定拒绝的阈值为 3，马氏距离判定接受的阈值为 1，而当  $1 < M\text{-Dist.} < 3$  时，就会被判定为 POSSIBLE，这些样品就需进一步的试验室验证。所建立的指纹图谱库可以进一步开发来建立该产品的定性判别模型，用来鉴别喷雾植酸酶和吸附干燥植酸酶产品质量的好坏和真伪（参见第 5 章）。

如果植酸酶产品指纹图谱库比较稳定，样本选择具有很好的代表性，并且能够反映同一加工工艺条件下不同批次产品的特征信息，则可以进一步开发建立植酸酶酶活的定量分析模型，从而用于生产企业终端产品的质量监控（参见第 4 章）。

### 3.3 小结

通过广泛收集具有代表性的试验样本，主要建立了喷雾干燥和吸附干燥植酸酶酶活的图谱库；采用化学计量学的方法剔除异常样本后，选择 MSC 做散射校正处理、一阶导数和平滑组合为光谱预处理方法，在最优的波数范围内，采用偏最小二乘法回归的方法（PLS-1）结合马氏距离判别法，对获得的图谱库进行比较分析，寻找分辨率高，重现性好的图谱；进行图谱辨认和评价，找出具有规律性的“共性”并加以量化；建立了喷雾干燥和吸附干燥植酸酶产品的指纹图谱库。

## 第四章 植酸酶酶活定量分析模型的建立与优化

### 4.1 实验材料、仪器与实验方法

#### 4.1.1 实验样品

本实验中用于建立定量分析模型的样品均来自于构建A厂家指纹图谱库中的样品。

#### 4.1.2 参考分析方法及仪器

植酸酶酶活和水分的测定方法均参照国家标准<sup>[74]</sup>进行。

所用仪器：

550A型pH计，美国Orion公司；GO1340C-1型烘箱，美国Lindberg/Blue公司；SS-Y不锈钢恒温水浴锅，上海三申医疗器械厂；90-3恒温双向磁力搅拌器，上海亚荣生化仪器厂；日立CR22G高速冷冻离心机，天美科技有限公司；梅特勒-托利多AL104-IC系列电子分析天平，瑞士；TU-1901型紫外可见分光光度计，北京普析通用仪器有限责任公司。

#### 4.1.3 近红外光谱分析方法

实验仪器为瑞典波通公司生产的DA7200型近红外光谱分析仪。光谱采集软件为仪器配置的Simplicity，建立校正模型所用软件为Grams32PLS+IQ软件（Thermo Galactic Corp.）。仪器工作参数：波长范围950-1650nm，采样间隔2.0nm，扫描次数32次（以内部陶瓷为参比），分辨率16cm<sup>-1</sup>。每个样品扫描2次，重复装样2次，取其平均光谱值作近红外光谱分析。

#### 4.1.4 数据处理

将样品分为校正集和检验集，对于校正样品集，应用偏最小二乘法（分为PLS-1和PLS-2）、主成分分析法（PCA）、主成分回归法（PCR）等回归方法，来分别建立喷雾干燥植酸酶和吸附干燥植酸酶酶活、水分的定量分析校正模型，再用交互验证的方法对检验集进行外部验证。借助不同的数学处理方法预处理原始光谱数据，主要包括标准正态变量转换（SNVT）、多元散射校正（MSC）、一阶导数和二阶导数（First Derivative & Second Derivative）等，提高和优化校正模型的准确性；然后，通过比较校正决定系数（R<sup>2</sup>）、交互验证标准差（SECV）考核校正模型的实用性，最后根据R<sup>2</sup>、SECV和SEP确定最优模型。

### 4.2 结果与分析

#### 4.2.1 参考方法分析结果

NIR分析技术是一种间接分析技术，要通过校正模型的建立来实现对未知样本的定性或

定量分析。校正模型是在一组已知待测组成或性质且具代表性样品（称为校正集样品）的基础上经化学计量学方法建立的通过现行的标准方法或常规方法来得到建立模型所必需的基础数据。因此，基础数据的准确性是影响近红外光谱分析结果准确性最重要的一个因素<sup>[75]</sup>。植酸酶样本化学成分含量化学分析统计结果见表4-1。从表4-1可以看出各个成分含量的覆盖范围较大，因此所建立的NIRS定标模型的适用范围也较宽。

表 4-1 植酸酶中成分含量的化学分析结果

Table 4-1 Chemical analysis results of phytase compositions

样品	成分	样品数	均值	最小值	最大值	相对标准偏差 (%)
喷雾干燥	酶活(IU)	134	4869	4422	5964	0.105
吸附干燥	水分(%)	134	6.027	4.476	8.386	0.059
喷雾干燥	酶活(IU)	121	2581	1827	3168	0.130
吸附干燥	水分(%)	121	5.457	3.796	6.861	0.118

所选择的用于建立定标模型的植酸酶样本的化学成分含量分布图见图 4-1。从图 4-1 可以看出定标集中喷雾干燥和吸附干燥植酸酶酶活和水分含量分布服从正态分布，所以用于建立定标模型的定标集中的样品具有很好的代表性<sup>[76]</sup>。

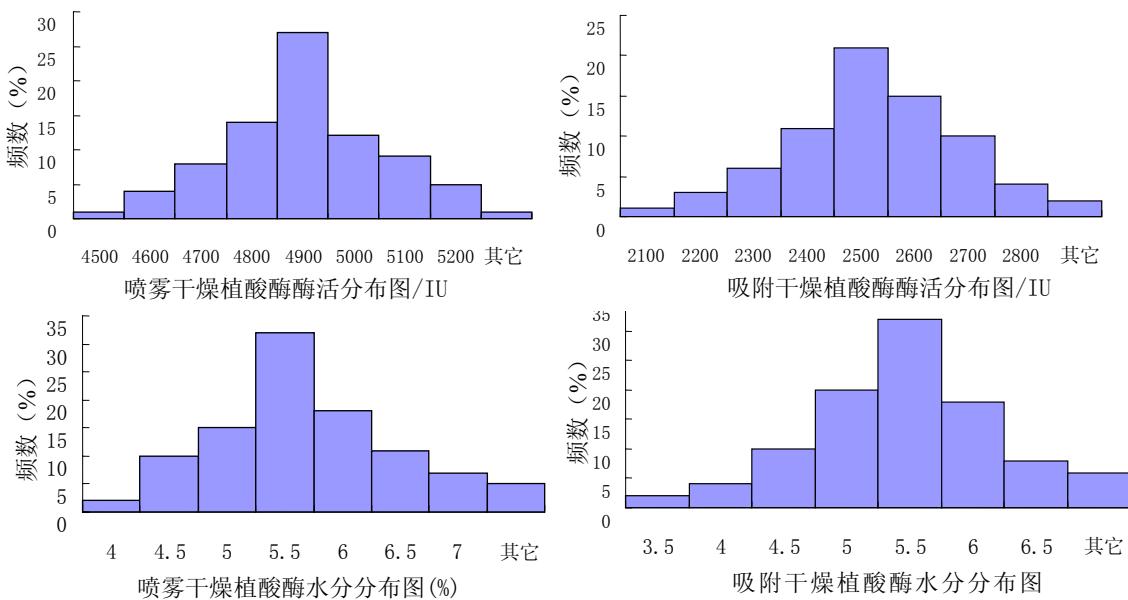


图 4-1 喷雾干燥和吸附干燥植酸酶酶活、水分含量分布图

Figure 4-1 Enzyme activity, moisture content distribution of spray drying and adsorption drying phytase

#### 4.2.2 近红外光谱采集

扫描条件：扫描次数为32，分辨率为 $8\text{cm}^{-1}$ ，光谱范围950–1650nm。每个样品扫描三次谱图后取平均，得到植酸酶样品近红外光谱图（如图3-10所示）。

图4-2 (a) 为A生产企业的喷雾干燥和吸附干燥植酸酶产品原始光谱图。图4-2 (b) 是经过基线校正和一阶导数处理后的A生产企业植酸酶近红外光谱图，从图上我们可以看出，

植酸酶近红外光谱在经SNVT处理和一阶导数变换后，不但较好的去除了基线漂移，而且更为细致的反映了不同样品之间的信息差异。

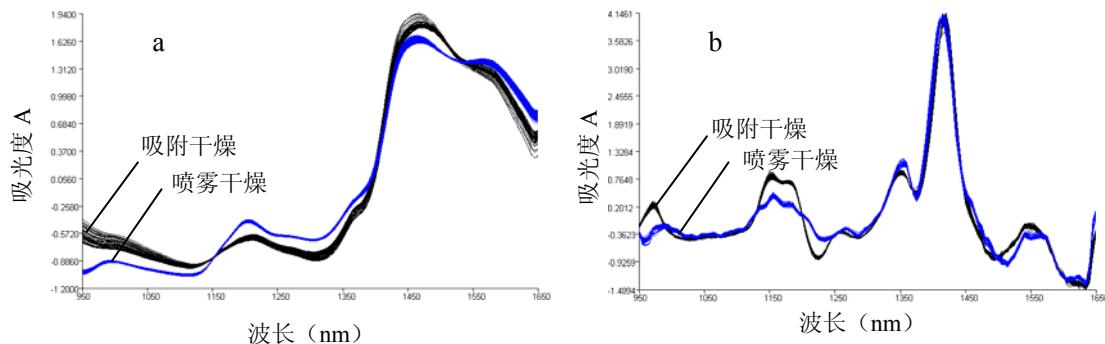


图 4-2 喷雾干燥和吸附干燥植酸酶近红外光谱图

Figure 4-2 NIR spectra of spray drying and adsorption drying phytase

#### 4.2.3 建模回归方法的确定

本研究的主要目的是在选定的散射校正和数学预处理方法的基础上，考查不同的回归分析方法（PLS-1, PLS-2, PCA 和 PCR）对NIRS定标效果的影响，以选择最优的回归分析方法，优化定标模型。利用智能化定量分析软件（Grams32PLS+IQ），在最优的波数范围内，选择四种不同的回归方法分别建立喷雾干燥和吸附干燥植酸酶产品酶活和水分的定量分析模型，来选择最优的回归分析方法，定标结果见表 4-2。

表 4-2 不同回归分析方法定标模型统计值

Table 4-2 Calibration mode statistics results from different regression analysis methods

样品	成分	回归方法	定标集				检验集	
			决定系数 $R^2$	标准差 RMSEC	交互决定系数 1-VR	交互标准差 RMSECV	决定系数 $r^2$	标准差 RMSEP
喷雾干燥	酶活	PLS-1	0.935	286.4	0.936	267.8	0.941	284.7
		PLS-2	0.928	325.7	0.919	352.7	0.926	307.6
		PCA	0.914	402.2	0.917	398.3	0.905	337.5
		PCR	0.907	428.4	0.889	355.0	0.894	390.6
水分	酶活	PLS-1	0.942	0.71	0.934	0.68	0.944	0.67
		PLS-2	0.931	0.76	0.927	0.70	0.927	0.71
		PCA	0.899	0.68	0.911	0.74	0.914	0.76
		PCR	0.917	0.75	0.925	0.66	0.901	0.73
吸附干燥	酶活	PLS-1	0.926	256.3	0.935	283.3	0.933	237.6
		PLS-2	0.899	309.5	0.901	335.8	0.929	277.8
		PCA	0.910	364.8	0.878	414.3	0.895	331.7
		PCR	0.915	378.9	0.923	385.1	0.924	386.3
水分	PLS-1	PLS-1	0.936	0.59	0.928	0.63	0.948	0.62
		PLS-2	0.931	0.61	0.916	0.59	0.926	0.61
		PCA	0.911	0.77	0.904	0.68	0.915	0.74
		PCR	0.923	0.76	0.915	0.72	0.920	0.71

从表 4-2 中可以看出，采用回归分析方法，所建立的喷雾干燥和吸附干燥植酸酶的定标模型的决定系数 ( $R^2$ , 1-VR 和  $r^2$ ) 和标准差 (RMSEC, RMSECV 和 RMSEP) 有一定的差异，说明光谱进行不同的回归方法分析处理，对植酸酶中酶活、水分的NIRS定标效果有一定的影响。

比较四种回归分析方法的定标结果，对植酸酶样本中酶活、水分进行NIRS定量分析，除个别处理例外，回归分析方法采用PLS-1 处理时，与其它三种校正方法 (PLS-2, PCA 和 PCR) 比较，其决定系数 ( $R^2$ , 1-VR 和  $r^2$ ) 最大，标准分析误差 (RMSEC, RMSECV 和 RMSEP) 最小。所以选择回归分析方法PLS-1，作为喷雾干燥和吸附干燥植酸酶酶活、水分含量的NIRS定量分析的处理方法。

#### 4.2.4 光谱预处理方法的确定

首先考察不同光谱散射校正 (MSC, SNV 和 Normalize) 对 NIRS 定标效果的影响，在此基础上，进一步探讨不同光谱数学处理 (平滑、导数及其与平滑组合) 对定标效果的影响，以选择最优的光谱预处理方法，优化定标模型。

##### 1. 散射校正方法的选择

利用智能化定量分析软件 (Grams32PLS+IQ)，在最优的波数范围内，选择偏最小二乘法 (PLS-1) 为回归分析方法，对光谱分别进行附加散射校正 (MSC)、变量标准化 (SNV) 和归一化 (Normalize) 预处理，来选择最优的散射校正方法，定标结果见表 4-3。

表 4-3 不同散射校正定标模型统计值

Table 4-3 Calibration mode statistics results from different scatter correction methods

样品	成分	光程校正	定标集				检验集	
			决定系数 $R^2$	标准差 RMSEC	交互决定系数 1-VR	交互标准差 RMSECV	决定系数 $r^2$	标准差 RMSEP
喷雾干燥	酶活	MSC	0.912	235.6	0.936	246.3	0.940	286.5
		SNV	0.895	336.2	0.914	382.5	0.917	315.7
		Normalize	0.856	412.7	0.897	402.1	0.886	299.4
	水分	MSC	0.916	0.68	0.921	0.71	0.935	0.65
		SNV	0.925	0.73	0.917	0.76	0.928	0.84
		Normalize	0.899	0.82	0.904	0.79	0.934	0.88
吸附干燥	酶活	MSC	0.923	278.3	0.955	282.6	0.923	236.7
		SNV	0.806	334.5	0.886	336.7	0.914	377.2
		Normalize	0.921	425.7	0.937	367.9	0.916	343.3
	水分	MSC	0.940	0.74	0.977	0.62	0.958	0.71
		SNV	0.931	0.81	0.941	0.76	0.935	0.82
		Normalize	0.833	0.86	0.906	0.84	0.924	0.80

从表 4-3 中可以看出，采用不同的散射校正处理方法，所建立的喷雾干燥和吸附干燥植酸酶的定标模型的决定系数 ( $R^2$ , 1-VR 和  $r^2$ ) 和标准差 (RMSEC, RMSECV 和 RMSEP) 有一定的差异，说明光谱进行不同的散射校正处理，对植酸酶中酶活、水分的NIRS定标效果有

一定的影响。

比较三种散射校正方法的定标结果，对植酸酶样本中酶活、水分进行NIRS定量分析，除个别处理例外，当散射校正采用MSC处理时，与其它两种校正方法（SNV, Normalize）比较，其决定系数（ $R^2$ , 1-VR和 $r^2$ ）最大，标准分析误差（RMSEC, RMSECV和RMSEP）最小。所以散射校正采用方法为MSC，作为植酸酶酶活、水分含量的NIRS定量分析的处理方法。

## 2. 数学处理方法的选择

在最优的波数范围内，选择偏最小二乘法（PLS-1）作为回归分析方法，选取多元散射校正（MSC）作为光谱散射校正处理的方法，对光谱分别进行平滑、导数（一阶导数和二阶导数）及其与平滑组合处理，以选择最优的数学处理方法，喷雾干燥和吸附干燥植酸酶产品酶活、水分含量的定标模型结果统计值见表 4-4。从表 4-4 可以看出，对光谱分别进行一阶导数、一阶导数和平滑组合、二阶导数、二阶导数和平滑组合 4 种不同的数学处理，显示出不同的定标结果。

表 4-4 不同数学处理方法建立定标模型统计值

Table 4-4 Calibration mode statistics results from different mathematical treatment methods

样品	成分	校正方法	定标集				检验集	
			决定系数 $R^2$	标准差 RMSEC	交互决定系数 1-VR	交互标准差 RMSECV	决定系数 $r^2$	标准差 RMSEP
喷雾干燥	酶活	1st	0.924	324.7	0.932	296.3	0.931	307.8
		1st & 5-SG	0.950	211.4	0.949	231.6	0.942	214.7
		2nd	0.883	413.5	0.932	420.5	0.904	308.2
		2nd & 5-SG	0.922	365.1	0.924	356.8	0.936	266.6
	水分	1st	0.916	0.87	0.925	0.78	0.933	0.80
		1st & 5-SG	0.931	0.85	0.934	0.82	0.950	0.75
		2nd	0.924	0.82	0.929	0.76	0.941	0.74
		2nd & 5-SG	0.947	0.71	0.967	0.66	0.951	0.59
	酶活	1st	0.896	342.3	0.878	320.6	0.922	286.3
		1st & 5-SG	0.935	246.4	0.942	254.7	0.961	201.7
		2nd	0.875	296.9	0.901	312.0	0.914	298.2
		2nd & 5-SG	0.894	273.6	0.923	260.9	0.947	236.9
吸附干燥	酶活	1st	0.917	0.86	0.918	0.79	0.920	0.83
		1st & 5-SG	0.925	0.75	0.931	0.72	0.937	0.69
	水分	2nd	0.913	0.82	0.924	0.85	0.929	0.77
		2nd & 5-SG	0.952	0.64	0.968	0.58	0.953	0.62

(注：1st 为一阶导数处理，2nd 为二阶导数处理，5-SG 为 5 点平滑处理)

(Note: 1st for first derivative, 2nd for second derivative, 5-SG for 5 point segment smooth treatment)

比较 4 种不同数学处理方法的定标效果，选择决定系数（ $R^2$ , 1-VR 和  $r^2$ ）最大，标准分析误差（RMSEC, RMSECV 和 RMSEP）最小的数学处理方法，作为建立酶活、水分含量的定标模型的最优数学处理方法。由表中的数据可以看出，选择一阶导数和五点平滑相结合的数学处理方法，定标效果最为理想。因此，本研究中建立喷雾干燥和吸附干燥植酸酶产品的

定标模型过程中所选择的数学处理方法是一阶导数结合五点平滑处理的方法。

#### 4.2.5 定标模型的建立与验证

##### 1. 最佳主因子数的确定

采用内部交互验证法（Internal Cross-Validation）对所建立的定标模型进行检验，根据预测残差平方和（PRESS）、交互验证标准差（RMSECV）进行判断，PRESS、RMSECV 值越小，说明定标模型的预测能力越强。通过选取不同的因子数，以最小的 PRESS、RMSECV 值对应的因子数为最佳主因子数，图 4-3，图 4-4 为 A 生产企业喷雾干燥植酸酶产品酶活在最优波数范围内，对光谱进行最优的预处理，交互验证 PRESS、RMSECV 值与主因子数的关系曲线图。

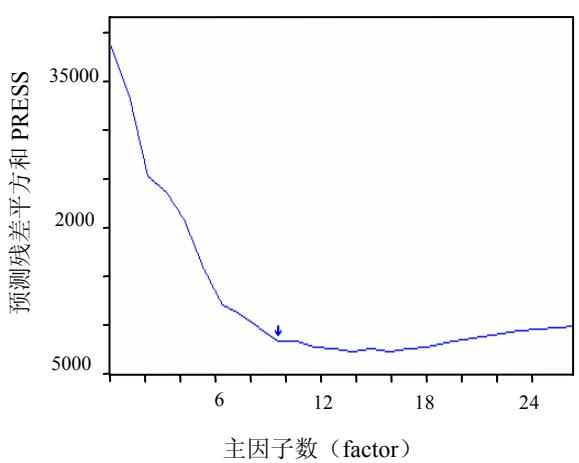


图 4-3 喷雾干燥植酸酶预测残差平方和

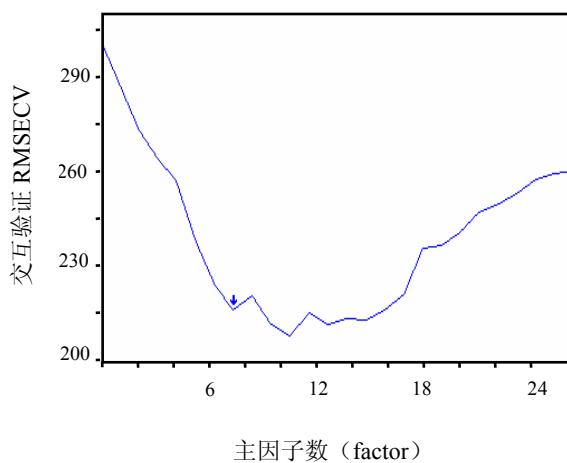


图 4-4 喷雾干燥植酸酶交互验证检验图

Figure 4-3 Prediction Residual Error Sum of Squares from spray drying phytase

Figure 4-4 Root Mean Square Error of Cross Validation from adsorption drying phytase

本研究中所采用的 PRESS、RMSECV 显著性检验法确定的植酸酶样品酶活、水分的最佳主因子数见定标模型建立中的表 4-5。

##### 2. 定标模型的建立

根据以上分析，选择 MSC 做散射校正处理、一阶导数和五点平滑相组合（1st & 5-SG）为光谱预处理方法，在最优的波数范围内，采用内部交互验证法验证，根据定标集交互验证标准差（RMSECV）、预测残差平方和值（PRESS），确定了最佳主因子数，采用偏最小二乘法回归的方法（PLS-1）从而建立了喷雾干燥和吸附干燥植酸酶样本中酶活、水分含量的 NIRS 定标模型（见表 4-5）。

应用相对标准差 RSD 和相对分析误差 RPD 这两个统计参数，评价所建立的植酸酶化学成分含量 NIRS 定标模型的预测精度和可靠性。从表 4-5 可以看出，喷雾干燥植酸酶和吸附干燥植酸酶产品的酶活、水分的决定系数  $R^2$  都在 0.90 以上，说明化学分析值与 NIRS 预测值之间具有很好的线性相关性，相对标准差 RSD 均小于 10%，相对分析误差 RPD 大于 3，取得了良好的定标效果，可以用于实际检测。

表 4-5 植酸酶定量分析定标模型统计值

Table 4-5 Calibration mode statistics results of phytase samples

样品	成分	样品数 (No.)	含量范围 Range	平均值 Mean	处理方法 Pretreatment	波长范围 (nm)	主因子数 Factor
喷雾干燥	酶活	81	4422–5964	4869	MSC & 1st & 5-SG	1050–1650	7
	水分	81	4.776–8.386	6.027	2nd & 5-SG	1050–1650	5
吸附干燥	酶活	80	1827–3168	2581	MSC & 1st & 5-SG	1050–1650	5
	水分	80	3.796–6.861	5.457	2nd & 5-SG	1050–1650	4

表 4-5 植酸酶定量分析定标模型统计值（续）

样品	成分	决定系数 $R^2$	标准差 RMSEC	交互验证决定 系数 1-VR	交互验证标 准差 RMSECV	相对标准差 RSD (%)	相对分析 误差 RPD
喷雾干燥	酶活	0.937	253.5	0.942	262.4	5.23	3.64
	水分	0.964	0.634	0.959	0.549	3.87	4.06
吸附干燥	酶活	0.926	178.9	0.934	169.5	6.21	3.45
	水分	0.951	0.554	0.962	0.479	4.22	4.93

(注: 1st 为一阶导数处理, 2nd 为二阶导数处理, 5-SG 为 5 点平滑处理)

(Note: 1st for first derivative, 2nd for second derivative, 5-SG for 5 point segment smooth treatment)

### 3. 定标模型的验证

利用建立的 NIRS 定标模型, 对验证集 40 个喷雾干燥植酸酶和 36 个吸附干燥植酸酶样本化学成分进行预测, 考察了所建植酸酶化学成分含量 NIRS 定标模型的预测精度和可靠性, 验证结果见表 4-6。

表 4-6 植酸酶样本化学成分含量 NIRS 定标模型的验证结果

Table 4-6 NIRS validation mode statistics results of phytase samples

样品	成分	样品数 (No.)	含量范围 Range	平均值 Mean	决定系数 $r^2$	标准差 RMSEP	相对标准差 RSD (%)	相对分析 误差 RPD
喷雾干燥	酶活	40	4356–5843	4816	0.924	230.7	5.38	3.56
	水分	40	4.653–8.421	6.027	0.954	0.676	4.05	4.48
吸附干燥	酶活	36	1769–3247	2623	0.916	163.5	5.67	4.12
	水分	36	3.654–7.278	5.568	0.958	0.499	4.45	4.37

从表 4-6 中可以看出, 酶活、水分的化学分析值与 NIRS 预测值之间的决定系数  $r^2$ , 达到 0.91 以上, 相对标准差 RSD 均小于 10%, 相对分析误差 RPD 均大于 3 (如图 4-5 所示), 预测效果良好, 说明所建立的 NIRS 定标模型可用于预测植酸酶中酶活、水分的含量, 且具有较好的预测精度。因此, 所建立的定标模型可用于定量检测植酸酶酶活、水分含量。

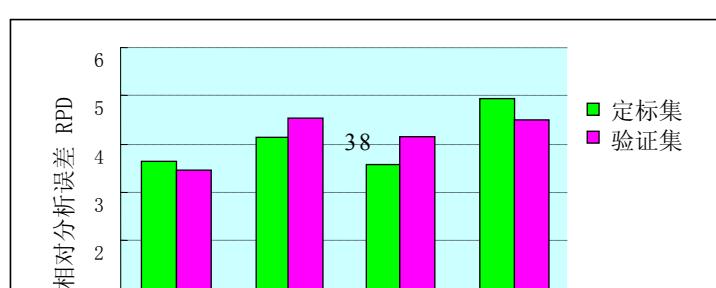


图 4-5 定标集和验证集植酸酶样品相对分析误差 RPD

Figure 4-5 Relative predictive determination of calibration and validation sets

采用一元线性回归的方法<sup>[77]</sup>，得到验证集中植酸酶样品成分含量的化学分析值与NIRS定标模型预测值之间的相关关系，如图 4-6 所示。从图中可直观地看出，喷雾干燥和吸附干燥植酸酶酶活的化学分析值和NIRS预测值之间具有很好的相关性。这说明植酸酶酶活和水分的NIRS定标模型具有较好的稳定性和预测精度。

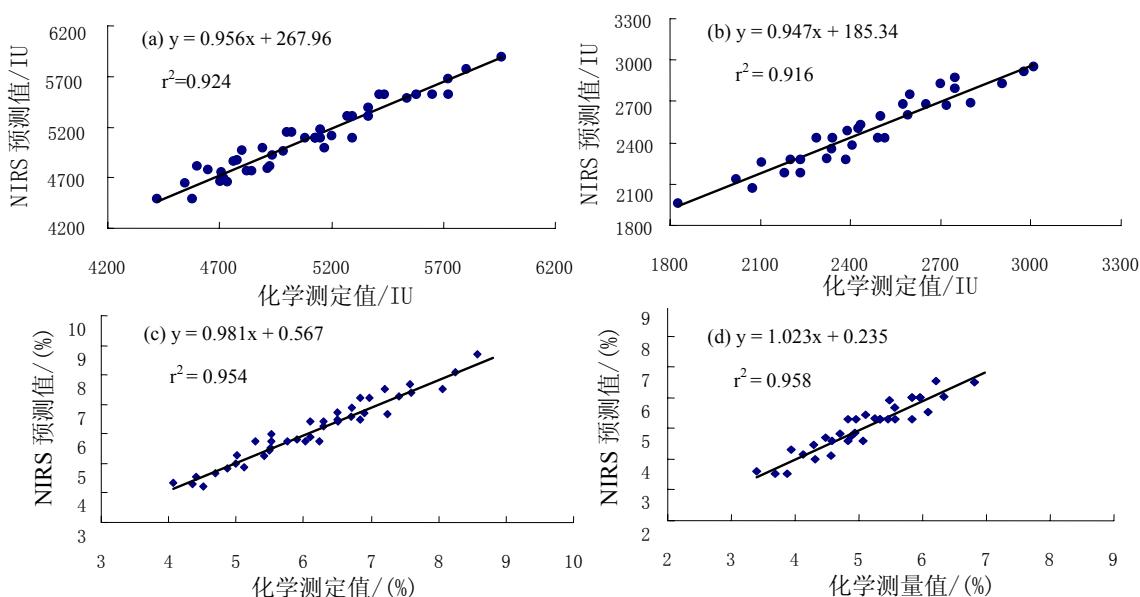


图 4-6 植酸酶校正模型化学测定值 VS 预测值

(a: 喷雾干燥酶活, b: 吸附干燥酶活, c: 喷雾干燥水分, d: 吸附干燥水分)

Figure 4-6 Actual Vs. predicted value of phytase calibration model

(a: Activity of spray drying phytase, b: Activity of adsorption drying phytase, c: Moisture content of spray drying phytase,

d: Moisture content of adsorption drying phytase)

#### 4.2.6 在校准曲线中加入温度补偿样品

温度影响近红外分析结果的主要原因有以下几方面：一方面是由仪器本身引起的，近红外分析仪的检测器对温度特别敏感。另一方面是源于样品本身，而样品温度的差异对近红外检测结果有直接的影响，这也是近红外仪器需要经常调整截距的原因。解决温度对近红外分析结果的影响主要有以下几个办法：①通过计算机选择对温度不敏感的特征波长点来建立样品的定量分析模型，这样的模型对温度比较稳定；②将样品温度调整至与定标时样品温度相

等后再进行上机测试；③将同一样品在不同的温度下进行扫描，通过特定的软件计算出温度补偿系数校正；④在建立定量分析模型时，最好在不同的温度下扫描样品光谱，使建模温度范围包含日常使用时的测量范围，把温度作为建立模型的因素之一，这样建立的模型就具有温度的自动矫正功能。

由于DA7200仪器自身有很好的温控系统，因此这种误差主要是样品温度引起的。因此，需要采取有效方法来消除这种影响，或能够补偿温度变化带来的检测结果误差。具体过程如下：①选择样品：样品应该有代表性，对于其中要检测的最关键指标要求有一定的梯度分布，但不要取极端值样品，正常分布即可，一般情况下5个左右。②样品的化学值：所有样品均须有实时的化学法检测值，双试样偏差均需控制在国标范围以内。对于有化学检测结果的参数都可以进行温度补偿。③改变样品温度并扫描：使用冰箱和烘箱以保证温度均匀。将样品温度从5℃—40℃，等间隔改变，温度不一定十分精确，但所用样品要同时进行。试验过程中样品要注意密封好，检测和扫描过程要迅速，以防止水分散失。样品要逐个从烘箱或冰箱中取出检测，以免间隔时间过长引起温度变化。④每个扫描的样品在不同温度下的化学参考值是相同的，将这些样品完全收集加入收集数据库中，重新建立模型即可。

### 4.3 小结

本试验共收集了134个喷雾干燥和121个吸附干燥植酸酶样本，对植酸酶中酶活、水分含量进行了近红外反射光谱（NIRS）分析。采用光谱杠杆值（Sample Leverage）和学生氏残差（Student Residual）等化学计量学的方法剔除了异常样本，在最优的光谱波数范围内，考察了不同校正模型（PLS-1, PLS-2, PCA, PCR），不同光谱散射校正（MSC, SNV 和 Normalize）和数学处理（平滑、导数及其与平滑组合）对NIRS定标效果的影响。选择最优的光谱预处理方法以及数学处理方法，结合PLS-1，建立了喷雾干燥和吸附干燥植酸酶中酶活、水分含量的NIRS定量分析模型。

1. 定标集化学分析值与NIRS定标模型预测值之间的决定系数 $R^2$ 和标准差RMSEC分别为：喷雾干燥0.964和0.634（水分），0.937和253.5（酶活）；吸附干燥0.951和0.554（水分），0.926和178.9（酶活），其中所建立酶活、水分定标模型的决定系数都在0.90以上，相对标准差RSD均小于10%，定标效果良好。

2. 用建立的植酸酶各个化学成分NIRS定标模型对验证集样本进行预测，预测值与化学分析值之间的决定系数 $r^2$ 和标准差RMSEP分别为：喷雾干燥0.954和0.676（水分），0.924和230.7（酶活）；吸附干燥0.958和0.499（水分），0.916和163.5（酶活），其中所建立酶活、水分预测模型的决定系数都在0.90以上，相对标准差RSD均小于10%，取得理想的预测精度。

3. 用近红外反射光谱分析方法能够较准确地检测植酸酶中酶活、水分的含量。由于近红外分析技术是一种快速、简便、有效的检测分析方法，通过对植酸酶样本进行NIRS定量分析，建立定标模型，可以达到定量检测植酸酶中酶活、水分含量的目的。

## 第五章 植酸酶产品定性分析模型的建立与优化

产品品牌保护是维护公平竞争的市场秩序和保护生产企业知识产权的最基本要求。众所周知，新技术和新产品的研发耗资较大，而造假和模仿成本要低得多，如果不能有效保护产品品牌和发明人权益，创新活动将会被严重遏制，最终导致行业整体效益的降低；另一方面，由于新技术和新产品的利润空间较大，一些不法厂商为了片面追求利润回报，常在新产品中掺假混杂，如果不能有效监控新产品的质量稳定性，将严重损害消费的利益。总之，研究建立快速、准确的植酸酶产品定性分析方法是高效开展植酸酶产品质量安全管理的内在需求，是解决市场植酸酶产品品牌混乱和加快淘汰伪劣产品的重要措施。

近红外光谱技术的定性分析是一大特色，依据同类样品在不同波长下具有相同的光谱，借助多元分析方法，使各种样品能够进行聚类识别，结合某一类定量模型进而对化合物的各定量参数做出判断。定性分析中采用的识别分析方法一般称为模式识别，模式识别方法主要有主成分分析结合马氏距离法、SIMCA法、聚类分析法（作图法）、神经网络法等。在近红外光谱图谱中不同组分所含相同基团产生的谱带、同一组分不同基团产生的谱带及同一基团不同振动形式的倍频、合频谱带会发生严重的重叠，所得图谱复杂，信息多，所以用近红外光谱进行鉴别，得到的信息是复杂体系中各个组分吸收的总和，无法对单个成分进行结构分析。这一点是近红外光谱分析技术的弱点但同时也是其优势。例如：有时候不同产品的人工识别比较容易，但对相同产品不同厂家的识别的准确界定是不容易的，这时，采用其它光谱技术进行精细的结构分析无助于解决问题，而采用近红外光谱与计量学方法结合进行数学判别，可给出准确答案。因此，定性分析方法是近红外光谱分析方法中非常重要的组成部分。

定性分析的实现有以下几种方法与步骤：

### 1. 近红外光谱峰位鉴别法

在近红外光谱图谱上，依据不同种产品所含化学成分的不同，含氢基团倍频与合频振动频率的不同，反映的图谱峰位、峰数及峰强是不同的，成分差异越大，图谱的特征性越强。采用简易的峰位鉴别可对其品种进行鉴别。在采用近红外光谱峰位直接鉴别时，为保证同种产品光谱的同一性，应选用多个不同厂家的同种产品、不同批次的产品同时进行考察，共同比较光谱的原始光谱和一、二阶导数光谱图。根据试验分析，同种产品的原谱在不同波段可能有较小的差异，但其一、二阶导数光谱应当接近一致。而不同形态产品因所含化学成分的不同使得光谱图谱曲线完全不同。如喷雾干燥植酸酶和吸附干燥植酸酶的光谱图谱，虽然这两种植酸酶中含有部分共有成分，但不同组分的差异掩盖了他们共性，通过对光谱进行预处理，可以更清晰地鉴别出喷雾干燥植酸酶和吸附干燥植酸酶（如图5-2所示）。采用峰位鉴别法主要是分析组分相差较远的不同产品，这种方法直观、简便，但对于相近或同种产品之间进行鉴别却爱莫能助。因为相近或同种产品之间相似信息很多，表现在原谱甚至一、二阶导数光谱图谱上均有一定的相似性，因此需要其它方法如：模式识别法来判别，近红外吸收的基团都是C-H、N-H、O-H，当这些基团的周围环境发生变化时，会导致峰位的变化，但变化幅度不大；决定峰位和峰强的是样品中的不同组分，而这些组分在同种产品中都存在，如甲基、亚甲基基团等。

### 2. 模式识别法

在实际工作中，经常遇到只需要知道样品的类别或质量等级，并不需要知道样品中含有的组分数与其含量的问题，这时需要应用到模式识别法，模式识别法主要用于光谱的定性分析。近红外分析技术的定性与定量分析不同之处在于定量分析必须测定训练集的相应基础数据，精度取决于参考方法本身的精度；定性分析依据主要是光谱图谱本身，因为光谱反映了真实的样品组成和结构信息，相同或近似的样品有着相同或接近的光谱。通过应用含糊的数学光谱匹配方法进行实际定性分析就是模式识别法，在近红外光谱定性分析中常用的模式识别方法有很多，如聚类分析法、主成分分析结合马氏距离法和人工神经网络方法等。

### (1) 聚类分析法

聚类分析是采用对光谱图谱进行降维处理，将光谱数据矩阵（光谱图形）分解得到主成分载荷矩阵和得分矩阵，然后使用样本集的主成分对应其相应的得分作图。一般三个主成分对应其得分作图得到的三维空间视图比较直观，这在研究与教学领域应用十分有效。由于近红外光谱的图谱特征不像中红外光谱那样明显，用于表征产品的精细结构并不常用。利用红外光谱特征中的差别，应用化学计量学方法进行分类是比较常用的方法。一般步骤是利用近红外光谱进行主成分分析判别，通过训练集样品光谱确定主因子数，来计算未知样品光谱的因子得分，并在多维空间确定其所属的类别，图谱分解的算法可以是主成分分析法，也可以是偏最小二乘法。

### (2) 主成分分析结合马氏距离法

由于一般作图分析直观但步骤繁琐，经常在试验报告中给出的是统计结果，实际操作中采用的方法是主成分分析法结合马氏距离法，马氏距离概念可以使校正集数据的内部变化表现出来，这样样品的分类结果较好而且也能给出保证测量结果优劣的统计结果。目前聚类分析、马氏距离等判别方法在近红外定性分析方法中已广泛使用。本文采用主成分分析结合马氏距离法，对真伪植酸酶产品进行建模判别，结果发现模型判别准确，性能稳定。

## 5.1 实验部分

### 5.1.1 实验仪器和样品

傅立叶变换近红外光谱仪（瑞典Perten公司）。附件配置：积分球漫反射检测器，29mm内径样品杯。本实验中共选用60个喷雾干燥植酸酶和60个吸附干燥植酸酶样本，其中定标集的45个样本均来自建立定量分析模型中的样本，而检验集中的样本分别只有10个来源于建立定量分析模型中的样品。另外5个为伪劣产品（喷雾干燥检验集中的#11-#15为另外一个生产企业的喷雾干燥植酸酶产品；吸附干燥检验集中的#11-#13为另外一个生产企业的吸附干燥植酸酶产品，而#14-#15为另外一个生产企业的喷雾干燥植酸酶产品），见表5-1。

表 5-1 植酸酶定性判别模型样品集

Table 5-1 Sample collection sets for phytase discrimination model

样品	收集样品数	定标集样品数	检验集样品数
喷雾干燥	60	45	15
吸附干燥	60	45	15

### 5.1.2 光谱采集

扫描条件：扫描次数为32，分辨率为 $8\text{cm}^{-1}$ ，光谱范围950–1650nm。每个样品扫描三次谱图后取平均，得到植酸酶样品近红外光谱图如图5-1所示：其中产品A和产品C为厂家A的喷雾干燥和吸附植酸酶产品，产品B和产品D为厂家B的喷雾干燥和吸附植酸酶产品。

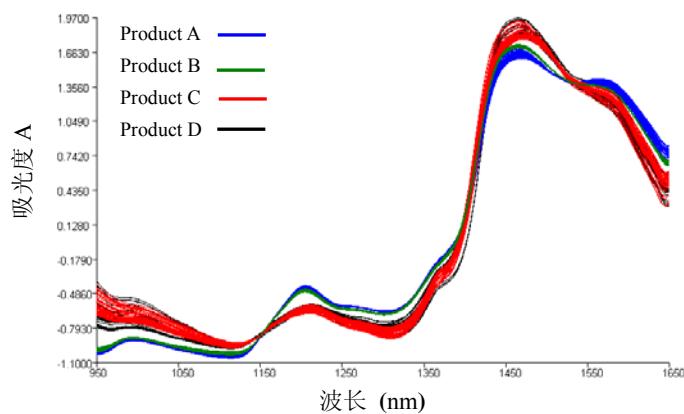


图 5-1 四种植酸酶产品的原始光谱图

Figure 5-1 Original spectra of four phytase products

图 5-1 为 60 个喷雾干燥和吸附干燥植酸酶样品的近红外原始光谱图。虽然不同厂家及真假植酸酶在化学成分组成上可能会存在较大差别，但它们的原始近红外光谱十分相似。这是由于近红外光谱带自身严重重叠所致，因此从原始近红外光谱中很难找出特定的吸收谱带对其加以区分。图中近红外光谱基线的漂移主要是由于样品的颗粒尺寸不同所致，一般随颗粒尺寸的不同，吸收强度也会不一样，从而导致基线的上下漂移。为减小样品颗粒尺寸、均匀性对近红外光谱的影响，常采用多元散射校正技术对原始近红外光谱进行处理。光谱基线的漂移采用一阶导数变换加以消除。

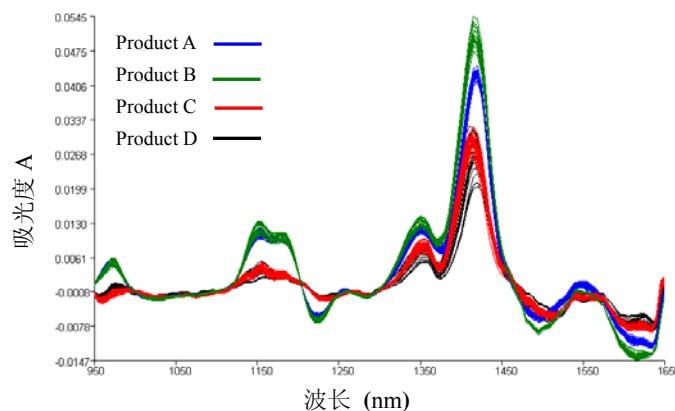


图 5-2 四种植酸酶产品的一阶导数光谱图

Figure 5-2 The first derivative spectra of four phytase products

图 5-2 是经过基线校正后的植酸酶近红外光谱图，从图上我们可以看出，植酸酶近红外光谱在经 SNVT 处理和一阶导数变换后，不但较好的去除了基线漂移，而且更为细致的反映了不同样品之间的信息差异。

## 5.2 结果与分析

### 5.2.1 主成分分析

主成分分析的中心目的是将数据降维，以排除众多化学信息共存中相互重叠的信息，它将原变量进行转换，使少数几个新变量成为原变量的线性组合。同时，这些新变量要尽可能多地表征原变量的数据结构特征而不丢失信息。调用主成分分析程序，在 950–1650nm 范围内提取来自于不同厂家、不同形态植酸酶产品的漫反射近红外光谱的主成分，取主成分1和主成分2作二维显示。结果表明，在不同的模式中，前两个主成分的对信息量的累积贡献率均在99%以上；相同厂家、相同形态的植酸酶产品均能较好的聚为一类。对一阶导数变换后的植酸酶近红外光谱图进行主成分分析，并以第一主成分为X轴，第二主成分为Y轴作图，得到两个生产企业（企业A和企业B）两种形态的植酸酶产品（喷雾干燥和吸附干燥）的得分图（如图5-3所示，图中Product A，和Product C分别是生产企业A的喷雾干燥和吸附干燥植酸酶，Product B，和Product D分别是生产企业B的喷雾干燥和吸附干燥植酸酶）。

由图5-4 我们可以看出，55个真品喷雾干燥植酸酶由于其所用原料和制作工艺的稳定，其产品质量也相对稳定，反映到近红外谱图上，即其谱图在经过主成分变换后投影到PC1和PC2二维空间上的点，聚集在如图标识的圆圈范围内。而伪品植酸酶（不是同一生产企业的产品）由于所用材料和工艺的不确定，导致其相互间质量上存在着差异，因此，近红外谱图之间也有着相当大的差异，反映到主成分投影空间上，即其投影点散落在四周，并不聚集在一处，并且没有一个伪品是投影在真品聚集的范围内，因此，可以看出，真伪品的植酸酶质量存在着显著性的差异，因此，我们可以通过模式识别的方法利用近红外谱图对植酸酶真伪进行鉴别。

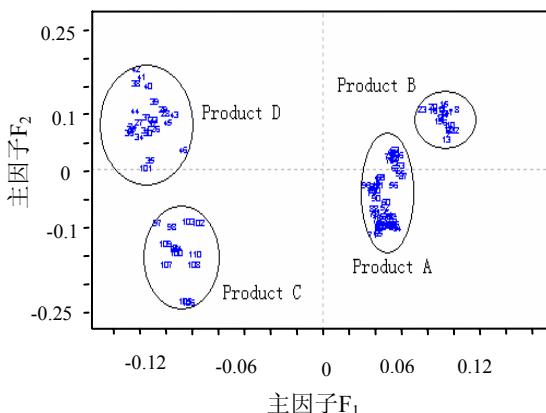


图 5-3 植酸酶近红外光谱得分图

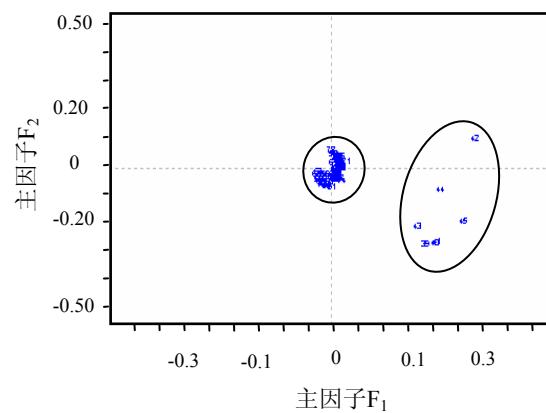


图 5-4 近红外光谱图的主成分分析结果

Figure 5-3 NIR spectra score plot of phytase

Figure 5-4 Principal component analysis results of NIR spectra

### 5.2.2 主成分得分结合马氏距离法建模预测

#### 1. 马氏距离法原理

首先对近红外光谱矩阵进行主成分降维处理，再利用光谱的主成分得分计算马氏距离。

计算步骤如下：

- (1) 首先计算出校正集样本的主成分得分矩阵  $t$ ;
- (2) 利用校正集样本得分矩阵  $T$ , 计算出校正集样本协方差矩阵  $M$ ;

$\bar{T} = \frac{\sum_{i=1}^m t_i}{m}$ $T_{cen} = T - \bar{T}$ $M = \frac{T'_{cen} T_{cen}}{m-1}$	(5-1)
---	-------

- (3) 计算校正集样本的马氏距离;

$D_i = [(t_i - \bar{T}) \bullet M^{-1} \bullet (t_i - \bar{T})']^{1/2}$	(5-2)
---	-------

根据定量校正允许的误差和对应的  $D$ , 确定离群点的  $D_i$ 。

- (4) 由未知样本的光谱  $A_{un}$ , 利用校正集样品求出其载荷, 计算光谱得分  $t_{un}$ , 由  $t_{un}$  根据第 3 步方法, 计算出其马氏距离  $D_m$ , 如果  $D_m < D_i$ , 则说明未知样品与校正集是同一类, 如果  $D_m > D_i$ , 则说明未知样品应归于新类。

## 2. 光谱预处理方法与建模

首先考察不同的回归方法 (PCA、PCR、Discriminate) 对定标模型的影响, 其次考察光谱散射校正 (MSC, SNV 和 Normalize) 对 NIRS 定标效果的影响, 在此基础上, 进一步探讨不同光谱数学处理 (平滑、导数及其与平滑组合) 对定标效果的影响, 以选择最优的光谱预处理方法, 优化定标模型, 得到回归方法和光谱预处理方法见表 5-2。

表 5-2 植酸酶定性判别校正模型统计值

Figure 5-2 Calibration statistical results of phytase qualitative discrimination model

样品	回归方法	预处理方法
喷雾干燥	PCA	多元散射校正 & 一阶导数 & 5 点平滑
吸附干燥	PCR	标准正态变换 & 一阶导数 & 5 点平滑

## 3. 模型预测结果

根据前面的分析结果, 选择多元散射校正 (MSC) 做散射校正处理、一阶导数和 5 点平滑组合为光谱预处理方法, 在最优的波数范围内, 采用内部交互验证法验证, 根据定标集交互验证的标准差 RMSECV 值, 确定了最佳主因子数, 采用主成分得分结合马氏距离法从而建立了喷雾干燥和吸附植酸酶样本的 NIRS 定性判别模型。具体分析结果见表 5-3 和表 5-4。

主成分分析结合马氏距离法判别主要是采用四项指标 (Limit Test) 来判别产品的归属类别。这四项指标分别为主成分得分 (Score), 光谱残差 (Resid.), 马氏距离 (M Dist.) 和浓度残差 (Conc.)。判别结果中的 P=PASS 表示此项指标通过检验, F=FAIL 表示此项指标未能通过检验, ?= POSSIBLE 表示不能确定此项指标是否通过。设定马氏距离判定拒绝的

阈值为 3, 马氏距离判定接受的阈值为 1, 这样当  $1 < M\text{-Dist.} < 3$  时, 就会被判定为 POSSIBLE。从表 2-11 可以看出, Actual 栏表示的是样品的实际情况, Actual 栏中的 YES 表示该样品为真品, Actual 栏中的 NO 表示该样品为伪品。Match 栏表示的是定性判别模型的判别结果, 检验集 15 个样本中, 前 10 个为真品 (来自 A 生产企业), 后 5 个为伪劣产品 (来自于另外一生产企业)。从预测结果可以看出除 7# 和 8# 被判为 POSSIBLE 外, 其余都匹配准确。若设定马氏距离判定接受的阈值为 1.5, 则判断准确率为 100%。因此, 所建立的模型能准确无误的鉴别喷雾植酸酶的真伪。

如果用于建立定标模型的样品质量比较稳定, 植酸酶酶活含量比较一致, 则利用所建立的定性判别模型的马氏距离法, 还可用于生产企业终端产品的质量监控, 根据产品的特定性质, 设定一个马氏距离判定接受的阈值, 如果马氏距离大于设定的阈值, 则判为不合格, 小于设定的阈值就可判为合格产品。

表 5-3 喷雾干燥植酸酶马氏距离建模交互验证预测结果

Table 5-3 Prediction results with Mahalanobis distance of spray drying phytase

Sample	Actual	Match	M-Distance	Limit Tests	F-Ratio	F-Test
sample #1	YES	YES	0.9140	PASS (PPPP)	0.8040	0.6224
sample #2	YES	YES	0.9211	PASS (PPPP)	0.3488	0.4378
sample #3	YES	YES	0.7340	PASS (PPPP)	1.2477	0.7316
sample #4	YES	YES	0.8613	PASS (PPPP)	0.7304	0.5988
sample #5	YES	YES	0.8625	PASS (PPPP)	0.6688	0.5775
sample #6	YES	YES	0.7235	PASS (PPPP)	0.5128	0.5160
sample #7	YES	POSSIBLE	1.1071	PASS (PP?P)	0.6467	0.5695
sample #8	YES	POSSIBLE	1.1948	PASS (PP?P)	0.2979	0.4106
sample #9	YES	YES	0.8418	PASS (PPPP)	0.5645	0.5377
sample #10	YES	YES	0.6124	PASS (PPPP)	1.5238	0.7796
sample #11	NO	NO	3.5553	FAIL (FP?P)	1.8977	0.8223
sample #12	NO	NO	4.3475	FAIL (FP?F)	2.1641	0.8495
sample #13	NO	NO	6.1847	FAIL (F?FP)	3.2022	0.9530
sample #14	NO	NO	4.2086	FAIL (F?FP)	2.2006	0.8528
sample #15	NO	NO	5.7943	FAIL (F?FP)	3.6838	0.9366

注: Limit Test (1234): 1=Scores, 2=Resid, 3=M Dist, 4=Conc / P=Pass, F=Fail, ?=Possible

Calibration: msc & 1st & 5-SG

具体分析结果同表 5-3, 从所建立的吸附干燥植酸酶定性判别分析模型的预测结果可以看出 (见表 5-4), 交互验证的结果正确率为 100% (根据吸附干燥植酸酶的实际情况, 设定马氏距离判定拒绝的阈值为 2.5, 马氏距离判定接受的阈值为 1.5)。这说明应用主成分得分分析结合马氏距离方法建立的吸附干燥植酸酶的 NIRS 定性判别模型具有较好的稳定性和预测精度。

表 5-4 吸附干燥植酸酶马氏距离建模交互验证预测结果  
Table 5-4 Prediction results Mahalanobis distance of adsorption drying phytase

Sample	Actual	Match	M-Distance	Limit Tests	F-Ratio	F-Test
sample #1	YES	YES	0.8205	PASS (PPPP)	0.8919	0.6482
sample #2	YES	YES	0.9373	PASS (PPPP)	0.4358	0.4811
sample #3	YES	YES	0.7421	PASS (PPPP)	1.0124	0.6798
sample #4	YES	YES	0.6958	PASS (PPPP)	0.9006	0.6506
sample #5	YES	YES	0.5978	PASS (PPPP)	1.2469	0.7315
sample #6	YES	YES	0.6314	PASS (PPPP)	0.6489	0.5703
sample #7	YES	POSSIBLE	1.1241	PASS (PP?P)	0.5690	0.5395
sample #8	YES	POSSIBLE	1.0042	PASS (PP?P)	0.5403	0.5277
sample #9	YES	YES	0.8714	PASS (PPPP)	0.5593	0.5356
sample #10	YES	YES	0.4422	PASS (PPPP)	1.2711	0.7362
sample #11	NO	POSSIBLE	2.8199	FAIL (FP?P)	1.6655	0.8936
sample #12	NO	NO	11.2359	FAIL (FP?F)	0.7922	0.8143
sample #13	NO	NO	40.59110	FAIL (F?FP)	1.9285	0.9258
sample #14	NO	NO	361.2165	FAIL (FF?F)	1.8406	0.9058
sample #15	NO	NO	258.6354	FAIL (F?FP)	3.5084	0.9303

注: Limit Test (1234): 1=Scores, 2=Resid, 3=M Dist, 4=Conc / P=Pass, F=Fail, ?=Possible

Calibration: SNVT & 1st & 5-SG

### 5.3 小结

本文建立了一种基于近红外光谱技术的喷雾干燥植酸酶和吸附干燥植酸酶的质量判别方法, 以植酸酶的近红外漫反射光谱结合主成分分析马氏距离法, 所建立的模型都能准确无误的鉴别真伪植酸酶。该方法和以往其它植酸酶真伪鉴别方法相比, 具有快速、准确、客观等优点, 克服了人工判别方法易受人为主观因素的影响以及其它检测方法过于复杂和耗时的缺点, 因此可以推广到其它植酸酶产品的质量判别, 具有良好的应用前景。

### 5.4 液态植酸酶近红外光谱分析

#### 5.4.1 仪器设备

光谱采集使用的是瑞典波通公司的 DA7200 型近红外光谱仪。该仪器的具体配制如下: 近红外碘钨灯, 石英分束器, InGaAs 检测器。样品池采用的是 1mm 石英池, 兰格 YZ515 蠕动泵自动进样系统。

### 5.4.2 样品制备

将液态植酸酶产品 A 和 B, 按表 5-5 的方案分别配制成等浓度梯度的溶液, 其中产品 A 和 B 的浓度都是从 0–100%, 以 10% 的浓度梯度为间隔, 一共 22 个样品。

表 5-5 组分配制方案列表

Table 5-5 The list of components preparation of phytase

序列	1	2	3	...	9	10	11
A (%)	0	10	20	...	80	90	100
H <sub>2</sub> O (%)	100	90	80	...	20	10	0
序列	12	13	14	...	20	21	22
B (%)	0	10	20	...	80	90	100
H <sub>2</sub> O (%)	100	90	80	...	20	10	0

### 5.4.3 光谱采集及数据处理

利用积分球透射采样系统采集液态植酸酶NIR光谱, 以仪器内部陶瓷板为参比, 光谱采集范围 950–1650nm, 扫描分辨率为 4cm<sup>-1</sup>, 每个样品扫描 16 次取平均, 光谱数据以 2cm<sup>-1</sup> 为间隔采集。每个样品重复装样 2 次, 重复扫描 2 次, 取其平均光谱用做分析。测量得到的近红外光谱图见图 5-5。图中最上面的那一条是去离子水的近红外光谱图 (a), 中间的是植酸酶产品A的近红外光谱吸收图 (b)、最下面的是植酸酶产品B的近红外光谱吸收图 (c)。

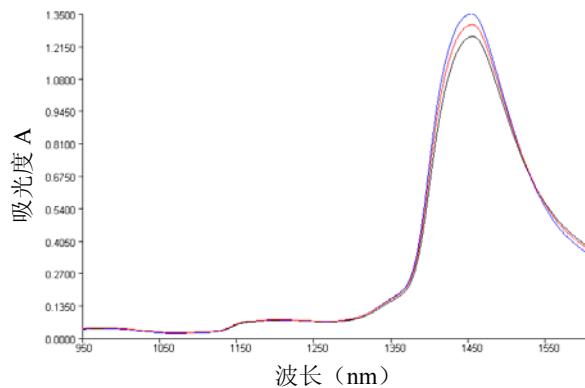


图 5-5 去离子水、液态植酸酶产品 A、B 的近红外光谱图

Figure 5-5 NIR spectra of deionized water, liquid phytase products A and B

将所得到的光谱以去离子水为背景, 处理成吸收光谱, 以便 PCA 分析用。数据处理采用近红外光谱分析仪自带的 Grams32PLS+IQ 软件 (Thermo Galactic Corp.)。由于液态植酸酶产品中存在着大量的水, 而水在近红外区域有强烈的吸收, 所以水的强烈吸收是近红外光谱法测定植酸酶含量的主要干扰。图 3-4 给出了在 950–1650nm 波段区域不同浓度的植酸酶样品的吸收光谱。单独测定两种植酸酶纯组分其近红外光谱图显示在该区域都有吸收, 并且光谱的峰位及峰强均不同。这是因为植酸酶产品中各组分的分子结构中含有 C–H、O–H 及 N–H 键, 而近红外光谱就是反映了 C–H、N–H 和 O–H 的倍频和合频吸收。因此, 近红外光谱法可以作为测定植酸酶产品中各组分含量的分析方法。

#### 5.4.4 液态植酸酶近红外光谱分析

根据光谱实验的具体情况，在近红外光谱分析过程中我们一般选用偏最小二乘法分析，并采用完全交互验证方法（Full Cross Validation）对模型进行评价，主因子数则根据样品的情况和建模过程中的残差计算值反馈加以确定。由图 5-6 可以看到，建模的结果以四幅图形显示，它们分别是：①得分图（Scores）、②载荷图（Loadings）、③光谱残差图（Spectral Residual）、④预测值与参考值对比图（Predicted Vs Actual）<sup>[78]</sup>：

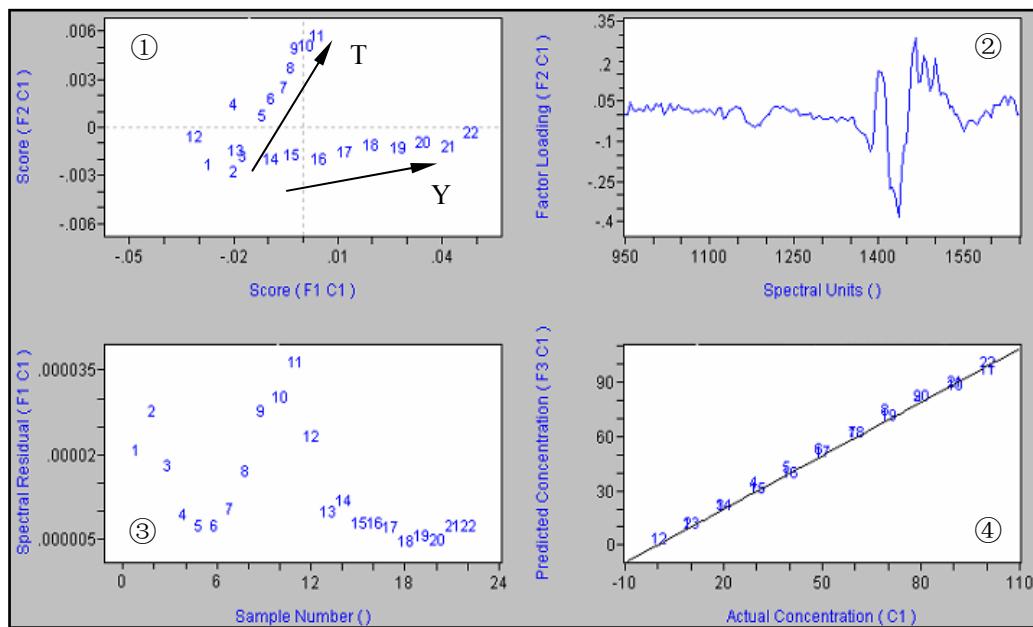


图 5-6 液态植酸酶近红外光谱分析图

Figure 5-6 NIR spectra analysis of liquid phytase

①得分图：此图为样本在以两个得分为坐标轴的坐标系中的分布图，得分（Score）是 PLS-1，PLS-2 做因子载荷分解得到的，它给出了一些样本信息，其中主成分 1 和主成分 2 的分布图是非常重要的，因为主成分 1 和主成分 2 在一般情况下将综合了 X 变量的大部分信息。得分图可以用来辨识样本是否同类，如果在分布图上样本分成几簇，则说明有几类样品，在一簇内则说明它们是同一类型的样本。另外得分图还可以用来识别奇异点，假如在图中有几个点与大多数点相距较远，则可以认为这几个点为奇异点。如图 5-6 中图①所示，随着液态植酸酶浓度的增加，在得分图呈现明显的变化趋势，A 生产企业的液态植酸酶主要是沿着第二主因子方向增加，而 B 生产企业的液态植酸酶主要是沿着第一主因子方向增加。这说明，第一主因子数包含了更多的 B 生产企业植酸酶的结构信息，第二主因子数包含了更多的 A 生产企业植酸酶的结构信息。

②载荷图：此图为 PLS-1 矩阵分解得到的 X 载荷图，从载荷图中可以看出哪个变量（波长）对吸光度值的贡献较大，并且它可与其对应的得分图结合起来进行分析，得分图右边的样本将在载荷图右边的变量有一个大的回归系数，相反在得分图左边的样本将在载荷图左边的变量有一个大的回归系数。如果两个主成分能综合 X 变量的大部份信息，则在载荷图上分布较接近的变量具有很高的正相关，假如两个变量在同一个象限倾斜 45 度的直线上离的较近，那么这两个变量具有同样的作用，如果在同一个象限倾斜 45 度的直线上两个变量方向

相反，则它们具有最大的负相关。从图 3-5 中图②可以看出，第一主成分的第二主因子(Factor Loading  $F_2C_1$ , Factor 表示为 F, Composition 表示为 C) 在波长 1450nm 附近有最大的负相关。

③光谱残差图：光谱残差图可以反应样品之间的差异，光谱残差值约大，则说明样品的差异程度越大，根据光谱残差图可以得出各个样本与样本集总体的相似程度，依据光谱残差值还可以剔除特异样本。图 3-5 中图③即为 A 生产企业和 B 生产企业不同浓度梯度液态植酸酶对应的光谱残差值。

④预测值与参考值对比图：此图为预测值  $Y'$  和参考值 Y 的回归直线图，它包括了建模校正和交互验证得到的两个预测值  $Y_c'$  及  $Y_p'$  和参考值 Y 的回归直线图，校正得到的预测值  $Y_c'$  用来评价模型的拟合能力，交互验证得到的预测值  $Y_p'$  用来评价模型的预测能力。常用标准偏差和相关系数来评价模型的好坏，图中出现的 Correlation 代表校正集的标准偏差和预测值  $Y_c'$  与参考值 Y 的相关系数，是用来评价模型的拟合能力的。相关系数越接近于 1，标准偏差越接近于 0，则此模型的拟合能力和预测能力越好。从图 3-5 中图④可以看出，A 生产企业和 B 生产企业的液态植酸酶都具有很好的相关性。因此，如果进一步建立针对每一个液态植酸酶产品的定量分析模型，就可以应用于植酸酶发酵过程的全程监控和液态植酸酶产品的质量监管。

#### 5.4.5 小结

本试验仅仅是探讨近红外光谱技术在液态植酸酶发酵过程和产品质量检测应用中的可行性。试验结果表明，近红外光谱技术可以更加全面、准确地反映液态植酸酶产品的个性特征信息（成分组成、稳定性）等信息，且该图谱信息比常规化学检测数据体现的产品信息更全面、更直观。如果建立相应的液态植酸酶酶活快速预测模型，则可以实现对液态植酸酶酶活的快速检测，利于生产的在线监控以及确保产品质量的稳定性。

## 第六章 植酸酶近红外指纹图谱库的应用研究

### 6.1 植酸酶喷雾干燥过程与近红外指纹图谱库

#### 6.1.1 概述

喷雾干燥植酸酶产品是市场植酸酶产品最主要的产品形式之一，喷雾干燥产品的使用优势在于瞬时干燥保护酶活、使用方便等。喷雾干燥温度对植酸酶活性的影响是决定植酸酶产品质量的一个重要因素，因此需要对喷雾干燥产品品质进行快速的检测和跟踪，最终确保喷雾干燥产品品质。喷雾制粒是一种较复杂的生产过程，实现生产过程中产品指标的实时监测（酶活、水分含量），对提高产品质量、实现生产装置自动化意义重大。而近红外（NIR）光谱分析技术是一种适宜于进行在线检控的手段，在药物生产工艺研究中已有一定应用。基于近红外光谱与模式识别技术，对喷雾干燥过程建立在线实时过程监控方法，使喷雾干燥过程实现高效与质量可控。此外，近红外光谱技术还可用于监控生产工艺变化，在喷雾干燥植酸酶产品在未包装之前，即可得到相应检测数据。本文探索利用近红外测试分析技术，实现对喷雾制粒生产过程产品指标的实时监测。该方法具有较高的测量精度和抗干扰性，为实现产品指标自动测控提供了一种新的方法。

本实验研究的主要目的是利用进红外光谱图解析原始光谱数据中所蕴藏的喷雾干燥条件信息，从而研究工艺特性对喷雾干燥植酸酶近红外光谱图的影响。

#### 6.1.2 喷雾干燥实验

##### 1. 仪器设备

分析天平：梅特勒-托利多 AL104-IC 系列；

小型喷雾干燥仪器：Büchi B-191, Switzerland；

高压均质机：NS 2006, L PONP, NIRO-SOAVI, Italy。

##### 2. 试剂及样品

可溶性淀粉，上海化学试剂公司；

麦芽糊精（DE=18），河南孟州市长鑫麦芽糊精厂；

液态植酸酶，分别来自 A 生产企业和 B 生产企业。

##### 3. 样品制备

将液态植酸酶样品与固形物（可溶性淀粉 w%/麦芽糊精 w%=2: 1）按 1: 1 进行配置，然后加入蒸馏水，使固形物含量控制在 25%~30% 之间（固形物含量过低，喷雾干燥产品不易干燥而导致产品含水量偏高；固形物含量过高，则易产品沉淀，从而堵塞喷嘴）。之后用高压均质机在 12000r/min 的转速下，高速剪切 5min，使液态植酸酶和载体成为乳化状态，以备喷雾干燥使用。在喷雾干燥的过程中，为防止固形物产生沉淀，应随时进行搅拌。

##### 4. 喷雾干燥条件

本研究中植酸酶所使用的喷雾干燥条件如表 6-1 所示<sup>[79]</sup>：

表 6-1 植酸酶喷雾干燥条件

Table 6-1 Spray drying conditions of phytase

样品	进口温度 (℃)	出口温度 (℃)
Phytase A	140, 150, 155, 160, 170	70, 75, 80, 85, 90
Phytase B	140, 150, 155, 160, 170	70, 75, 80, 85, 90

## 5. 结果与分析

在喷雾干燥工艺参数中，进出口温度是主要的影响因素，它直接影响到喷雾干燥过程本身和最终产品的品质（主要是植酸酶的酶活）。本实验中通过改变喷雾干燥器进出口的温度，研究其对干燥工艺和最终产品质量的影响。通过对改变喷雾干燥器进风口温度和出风口温度进行实验。表6-2为喷雾干燥植酸酶产品的水分含量和植酸酶酶活损失结果。以进口温度为140℃时喷雾干燥所得酶活含量为基准（定为100%），从表6-2可以看出，喷雾干燥过程中，进出口温度的不同，对植酸酶产品品质有很大影响。随着进出口温度的升高，产品的含水率会随之减少。研究表明，进风口的温度越高，干燥过程进入降速干燥阶段时间就越早，对应的干燥产品的含水量就越少，植酸酶损失也就越大。

表 6-2 植酸酶喷雾干燥植酸酶产品水分及酶活含量

Table 6-2 Enzyme activity and moisture content of spray drying phytase products

样品	进口温度 (℃)	出口温度 (℃)	水分含量 (%)	酶活含量 (%)
Phytase A	140	70	5.146	100
	150	75	4.929	97
	155	80	4.753	90
	160	85	4.472	84
	170	90	4.440	85
Phytase B	140	70	4.461	100
	150	75	4.210	94
	155	80	4.133	88
	160	85	4.069	83
	170	90	3.671	81

### 6.1.3 喷雾干燥过程中近红外光谱信息研究

#### 1. 光谱采集及数据处理

利用积分球漫反射采样系统采集喷雾干燥植酸酶近红外光谱，以仪器内部陶瓷板为参考，光谱采集范围为950–1650nm，扫描分辨率为 $4\text{cm}^{-1}$ ，光谱数据以 $2\text{cm}^{-1}$ 为间隔采集，每个样品重复装样2次，重复扫描2次，取其平均光谱用做分析。测量得到的近红外原始光谱图如图6-1所示。将其原始光谱图进行平均与喷雾干燥植酸酶的指纹图谱库平均光谱图进行比较如图6-2所示。

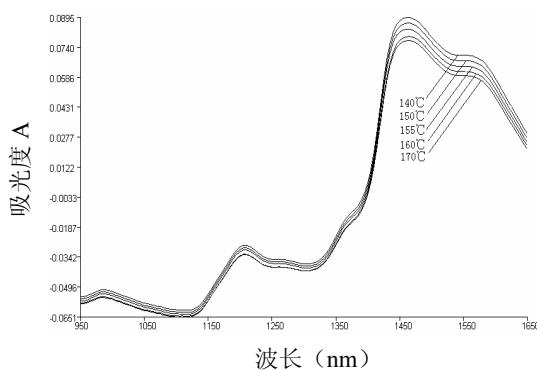


图 6-1 喷雾干燥植酸酶原始光谱图

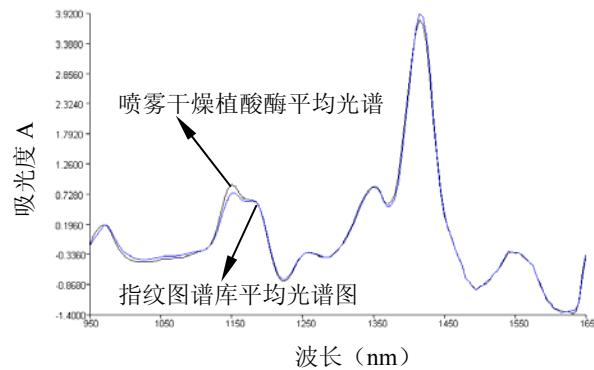


图 6-2 近红外光谱比较图

Figure 6-1 Original spectra of spray drying phytase

Figure 6-2 The comparison of NIR spectra

## 2. 结果与分析

测量得到的植酸酶 NIR 光谱经一阶导数处理，消除了背景的影响。为了减少光谱的变量，从而提高主成分分析（PCA）的速度，我们利用小波变换方法对二阶导数 NIR 光谱进行压缩。利用小波变换数据压缩技术既能高效地减少数据的变量数目，又能保持原光谱的特征。以便 PCA 分析用。数据处理采用近红外光谱分析仪自带的 Grams32PLS+IQ 软件（Thermo Galactic Corp.）。

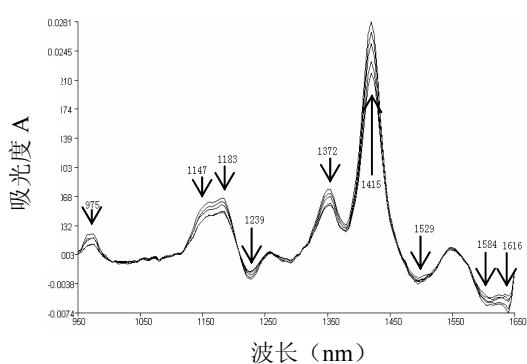


图 6-3 喷雾干燥植酸酶近红外光谱图 (SNVD+1ST)

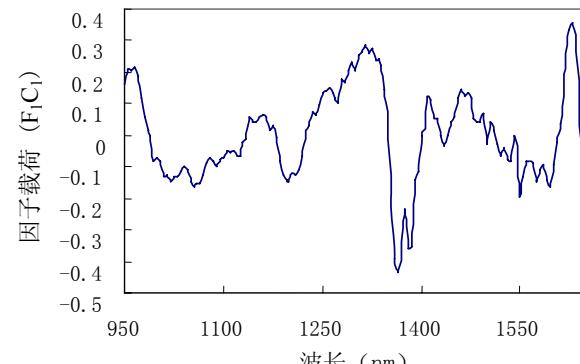
Figure 6-3 NIR spectrum of spray drying phytase  
(SNVD+1ST)

图 6-4 喷雾干燥植酸酶因子载荷图

Figure 6-4 NIR factor loading plot of spray  
drying phytase

图 6-3 为不同喷雾干燥温度条件下的植酸酶近红外光谱图。有研究表明，标准正态变换（Standard Normal Variate and De-Trend Transformation, SNVD）可以消除散射，颗粒大小，基线漂移和共线性所带来的干扰。而导数处理不仅可以减小散射效应，还可以提高光谱峰的分析精度<sup>[80]</sup>。所以为了阐明不同喷雾干燥温度对植酸酶的影响，对其原始光谱谱图进行 SNVD 和一阶导数处理（如图 6-3 所示），从图上我们可以看出，喷雾干燥植酸酶近红外光谱在经 SNVD 处理和一阶导数变换后，不但较好地去除了基线漂移，而且更为细致的反映了不同样品之间的信息差异。最上面一条曲线是为最低喷雾温度下得到的，依次往下，喷雾干燥的温度越来越高。由于氢键的作用以及和植酸酶产品中其它成分的相互作用，在特定的波长段存在着强烈的水分吸收峰，而且通常表现为很宽的吸收谱带。在近红外谱区，960–980 nm 和 1410–1450 nm 是与样品中水分的 O–H 键振动有关的<sup>[81]</sup>。由于蛋白质构象，蛋白质与水的

相互作用，或者其组合变化可能取决于含水量的多少，从而使得干燥样品的近红外反射光谱图受不同加热过程的影响。此试验中的样品在试验前后都放置于同一相对湿度条件下平衡 48 小时，所以植酸酶产品在喷雾后的水分含量保持是一致的，这样就可以减少不同喷雾干燥条件下由于水分含量不同所引起的差异。所以经过喷雾干燥得到的植酸酶近红外光谱图的差异就表现为喷雾干燥进风口温度和出风口温度所引起的差异。第 2.2 节中的蛋白质热变性研究表明，植酸酶产品中蛋白的变形发生在 $T_p$ 附近。此外，由于植酸酶产品中含有大量的官能基团，如 C–H, O–H, N–H, S–H 和 C=O 等，它们都是近红外的选择性吸收基团。因此，如果近红外光谱的变化与喷雾干燥温度之间有任何的线性关系，就可以利用校正模型追踪先前的喷雾干燥过程<sup>[82–83]</sup>。图 6-4 为喷雾干燥植酸酶产品的近红外因子载荷图，此图为偏最小二乘法 (PLS-1) 矩阵分解得到的 X 载荷图，第一主成分的第二主因子 (Factor Loading F<sub>1</sub>C<sub>1</sub>, Factor 表示为 F, Composition 表示为 C)，从载荷图中可以看出波长在 1360nm 附近对吸光度值的贡献较大，即在波长 1360nm 附近有最大的负相关。

利用所建立的 A 生产企业喷雾干燥指纹图谱库的定量分析模型（参见本文第 4 章），对本试验中喷雾干燥试验所得的植酸酶酶活、水分含量进行预测，结果并不理想。这主要是由于实验室条件下模拟的工艺条件与工厂实际生产工艺条件存在较大的差异，导致所建立的定量分析模型并不适应于分析，

#### 6.1.4 小结

借助近红外光谱分析中的一阶导数变换和主因子载荷分析等化学计量学的方法可以很清楚的反应由喷雾干燥不同所引起的植酸酶产品的差异，然后通过与植酸酶产品的指纹图谱库进行比较，就可以发现喷雾干燥条件的变化。因此，通过建立喷雾干燥过程的定量分析模型，近红外光谱分析技术可以用于追踪植酸酶产品的加工特性（喷雾干燥条件）的变化。

## 6.2 植酸酶储藏过程与近红外指纹图谱库

### 6.2.1 概述

如何提高植酸酶在各种环境下的活性，延长植酸酶的保质期，以满足运输、储藏的需要，降低植酸酶对环境的敏感度，使之在各种不同的加工环境中都能保持较高的活性，是植酸酶研究领域的又一重要课题。其中储藏过程中的酶失活是影响产品质量和稳定性的另一重要原因，也是制约产品市场流通的主要原因，及时准确跟踪储藏过程中的酶活特性变化成为确保产品质量和产品品牌的重要任务。高温高湿条件下很容易使植酸酶活性丧失，而植酸酶活性的丧失直接与植酸酶的构象有关，除通过繁琐的植酸酶酶活测定外，如何通过运用主成分分析近红外光谱，用少数几个不相关的综合变量代替原来较多的相关变量，而又能够综合反映原变量提供的大部分信息，从而反映植酸酶的储存条件信息。这对于未知组分光谱的物质极其重要，通过对物质的光谱数据进行主成分分析，可以得出其所含组分的光谱信息。本试验研究主要目的是利用植酸酶指纹图谱库来解析原始光谱数据中所蕴藏的储存条件信息。

### 6.2.2 恒温恒湿储藏试验

#### 1. 仪器设备

双层密封盒（Antibiotic Airtight Container, 27cm×11cm×11cm），CO<sub>2</sub>恒温培养箱（MCO-175, Sanyo），紫外可见分光光度计（UV-2100），恒温水浴（NTT-1100, EYELA），磁力搅拌器，酸度计，离心机。

#### 2. 试剂

植酸钠（C<sub>6</sub>H<sub>6</sub>O<sub>24</sub>Na<sub>12</sub>, Sigma），牛血清白蛋白（BSA），三水乙酸钠，EDTA，钼酸铵，钒酸铵，乙酸，硝酸，氨水，磷酸二氢钾，溴化钠，氯化铜，氯化钠，氯化钾，氯化钡。

#### 3. 样品制备

首先将5种饱和盐分别装入相应的称量皿中，然后加入蒸馏水使其成为饱和盐溶液，这样在一定的温度条件下就能达到一定的恒定的相对湿度（如表6-3所示）。之后将盛有饱和盐溶液的称量皿放入双层密封盒的底部。将同一批次的植酸酶样品分别放入五种不同湿度环境的密封盒中，每个湿度环境两个平行，每个平行两个重复。

表6-3 五种饱和盐溶液所对应的相对湿度（40℃）

Table 6-3 The relative humidity of five saturated salt solution

饱和盐溶液	相对湿度（RH, %）
NaBr	RH <sub>1</sub> =53.17±0.37
CuCl <sub>2</sub>	RH <sub>2</sub> =61.28±0.32
NaCl	RH <sub>3</sub> =74.68±0.13
KCl	RH <sub>4</sub> =82.32±0.25
BaCl <sub>2</sub>	RH <sub>5</sub> =89.90±0.28

#### 4. 结果与分析

每隔一周，在进行近红外光谱扫描以后，取储存在高温高湿条件下的植酸酶样品，按国标测定植酸酶酶活含量，每个湿度条件下取2个样品，求其平均值，得到在40℃条件下，不同相对湿度条件下储存10周植酸酶产品的酶活，如图6-5所示。

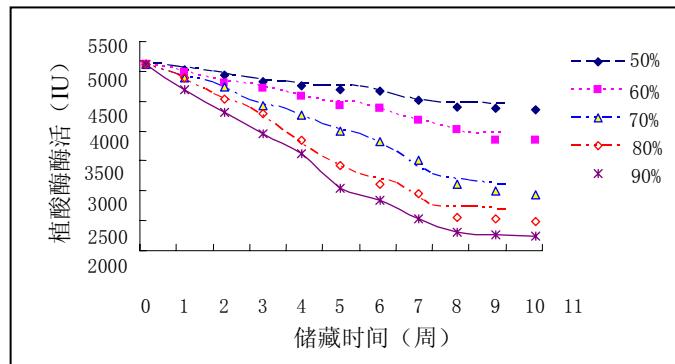


图 6-5 不同相对湿度条件下不同储存时间所对应的酶活 (40℃)

Figure 6-5 The changing of enzyme activity at different relative humidity conditions under different storage time

#### 6.2.3 恒温恒湿储藏过程中近红外光谱信息研究

##### 1. 光谱采集及数据预处理

光谱采集范围950–1650 nm，采样间隔2.0nm，扫描次数32次（以内部陶瓷为参比），分辨率 $16\text{cm}^{-1}$ 。每个样品扫描2次，重复装样2次，取其平均光谱值作近红外光谱分析。测量得到的近红外原始光谱图（如图6-6所示）。

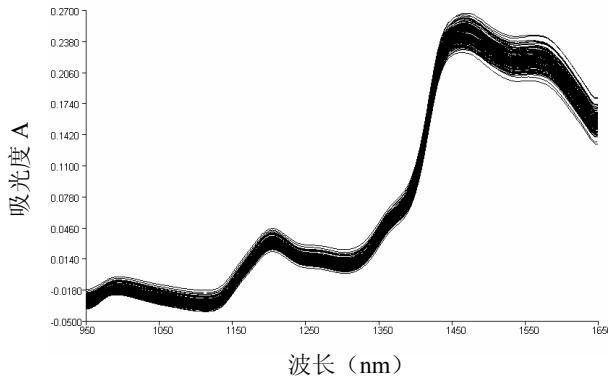


图 6-6 不同储存条件下的近红外原始光谱图

Figure 6-6 NIR original spectra under different storage conditions

##### 2. 结果与分析

NIR光谱采集过程中，由于样品颗粒大小、均匀性、仪器等因素的影响，往往会导致光谱基线产生偏移或漂移，因此建模前原始光谱需经过预理。常用的预处理方法有一阶导数、二阶导数、数据中心化等。按本方法每隔一周对不同相对湿度条件下的植酸酶样品进行近红外光谱扫描，得到植酸酶产品的NIR一阶导数谱图（见图6-7）。从图可以看出，光谱数据经导数处理后，可消除基线偏移，扣除本底吸收，从而更能体现样品的光谱特征。同时也可看出，不同相对湿度储存条件下的植酸酶样品，其一阶导数光谱图存在显著差异。

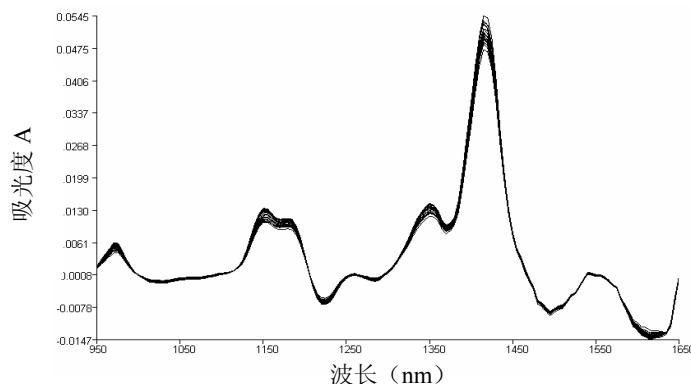


图6-7 植酸酶近红外一阶导数光谱图

Figure 6-7 NIR first derivative spectra of phytase

在利用校正集中的样品建立判别分析模型时，由于NIR光谱数据是一个高维数据矩阵，其中包含有许多无用信息和噪声信息，为充分利用光谱信息建立有效的分析模型，必须对光谱数据进行降维。主成分分析是最为常用和有效的降维方法。前8个主成分的累积贡献率变化见图6-8。对一阶导数光谱数据做主成分分析，前两个主成分的累计贡献率为90.84%。以第一主成分为横坐标，第二主成分为纵坐标绘图，如图6-9所示。从图6-9可以清楚地看出，不同储藏湿度条件下植酸酶聚集成一簇，而且随着储藏湿度的升高，高湿度储藏条件下的植酸酶产品离所建立的该产品指纹图谱库越来越远，而且呈现更加松散的变化趋势。这可能是由于在高湿度条件下，植酸酶结构发生了变化，酶活丧失越来越严重，图6-5的试验结果可以证实这一推测。因此，应用该产品的指纹图谱库可以监测植酸酶产品储藏过程中的水分含量变化（湿度）和品质变化（酶活）。

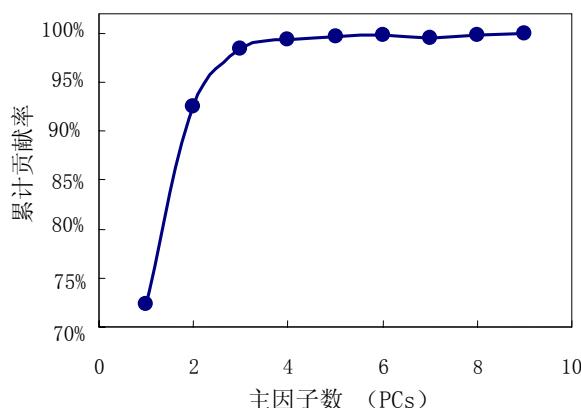


图 6-8 各主成分累计贡献率变化图

Figure 6-8 Cumulative contribution rates of different principal components

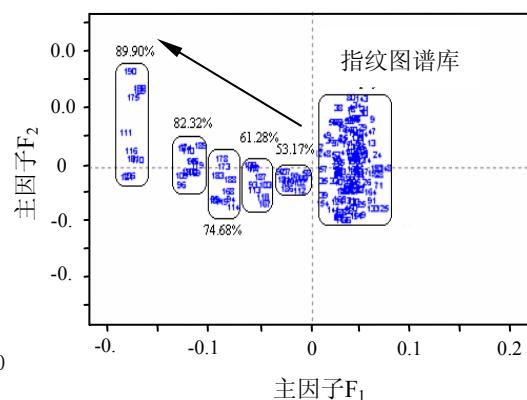


图 6-9 植酸酶近红外光谱得分图

Figure 6-9 NIR score figure of phytase

#### 6.2.4 小结

借助近红外光谱分析中的一阶导数变换和主成分分析等化学计量学的方法，通过与建立的该产品的指纹图谱库进行比较，可以很清楚地反应储藏过程中植酸酶产品水分、品质的变异程度，因此，近红外光谱分析技术可以用于植酸酶产品的储藏监管过程。

## 第七章 结论及建议

### 7.1 主要结论

本研究主要收集了来自 A 生产企业的 134 个喷雾干燥植酸酶和 121 个吸附干燥植酸酶样品，在 950–1650nm 谱区范围内，对试验样本进行了近红外光谱扫描。采用化学计量学的方法剔除异常样本后，选择 MSC 做散射校正处理、一阶导数和平滑组合为光谱预处理方法，在最优的波数范围内，采用偏最小二乘法回归的方法（PLS-1）结合马氏距离判别法从而建立了喷雾干燥和吸附干燥植酸酶产品的指纹图谱库。在所建立的产品指纹图谱库的基础上还进一步建立了喷雾干燥和吸附干燥植酸酶样本中酶活、水分含量的 NIRS 定量分析模型，和判别植酸酶产品质量好坏真伪的定性分析模型，同时，本论文还利用 SDS-PAGE 电泳、反相高效液相色谱（RP-HPLC）、差示扫描量热仪（DSC）研究分析了不同植酸酶产品的组成及特性，具体研究结果如下：

#### 1 植酸酶产品定量分析模型的建立与评价

在所构建的指纹图谱库的基础上，分别建立了 A 生产企业喷雾干燥和吸附干燥植酸酶中酶活、水分含量的 NIRS 定量分析模型。定标集化学分析值与 NIRS 定标模型预测值之间的决定系数  $R^2$  和标准差 RMSEC 分别为：喷雾干燥 0.964 和 0.634（水分），0.937 和 253.5（酶活）；吸附干燥 0.951 和 0.554（水分），0.926 和 178.9（酶活），其中所建酶活、水分定标模型的决定系数都在 0.90 以上，相对标准差 RSD 均小于 10%，定标效果良好。植酸酶酶活的化学分析值与 NIRS 定标模型预测值之间的决定系数  $r^2$  达到 0.90 以上，相对标准差 RSD 均小于 10%，相对分析误差 RPD 大于 3。喷雾干燥植酸酶酶活的决定系数  $r^2$  为 0.924，相对标准差 RSD 为 5.38%，相对分析误差 RPD 为 3.56；吸附干燥植酸酶酶活的决定系数  $r^2$  为 0.916，相对标准差 RSD 为 5.67%，相对分析误差 RPD 为 4.12。结果表明，利用近红外光谱分析技术能够准确检测喷雾干燥植酸酶和吸附干燥植酸酶酶活含量。

2 利用主成分分析定性判别程序，在 950–1650nm 范围内提取来自不同生产企业，不同形态的植酸酶产品的漫反射近红外光谱的主成分，取主成分 1 和主成分 2 作二维图。结果表明，在不同的模式中，前两个主成分对信息量的累积贡献率均在 90% 以上；相同厂家，相同形态的植酸酶产品均能很好的聚为一类；采用主成分分析（PCA）所建立的喷雾干燥和吸附干燥植酸酶定性判别模型分别对 15 个未知样品进行判别，准确率均达 100%，说明应用主成分分析所建立的定性分析模型可用来定性判别植酸酶产品的真伪，通过马氏距离法结合设定相应的阈值，还可以对产品质量的好坏和稳定进行判别。

3 同时借助反相高效液相色谱（RP-HPLC）和 SDS-PAGE 电泳研究其蛋白组成、差示扫描量热仪（DSC）研究其变性温度和热稳定性，探讨植酸酶产品蛋白组成、热稳定性同产品的近红外光谱信息之间的联系，明确了植酸酶产品近红外光谱指纹图谱库涵盖了产品的蛋白组成等相关信息，所构建的指纹图谱库能够为植酸酶等生物制品的质量监控和登记注册提供参考与借鉴。最后，借助所建立的产品指纹图谱库分别建立了快速跟踪喷雾干燥加工、储藏过程中植酸酶质量变化的近红外光谱方法。

## 7.2 建议

近红外光谱指纹图谱库既能满足政府监控饲料生产企业的原料和生产工艺的变化，降低产品质量监测的工作强度，促进生产企业提高产品质量，同时又在一定程度上保护生产企业的知识产权，激发企业创新积极性。应该说，针对某一特定产品建立的近红外光谱指纹图谱库具有广阔的应用前景和社会效益，而如何将近红外光谱指纹图谱库进一步细化，以及引入到产品的安全管理和登记注册将是进一步研究的方向，具体包括以下几个方面：

1. 研究制定建立指纹图谱库的标准产品确证规则；
2. 研究建立产品质量快速指纹鉴别方法标准；
3. 研究建立将指纹图谱库技术引入产品登记注册的关键技术；
4. 研究产品指纹技术网络诊断应用管理模式。

## 参考文献

1. Herschel W. . Experiments on the refrangibility of the invisible rays of the sun. Philosophical Transactions of the Royal Society 1800, 90: 225~289.
2. Abney W., Festing E.R. . The influence of the atomic grouping in the molecules of organic bodies on their absorption in the infrared region of the spectrum. Philosophical Transactions of the Royal Society 1881, 172: 887~918.
3. Wheeler O.H. . A neglected field of spectral study. Journal of Chemical Education 1960, 37: 234~236.
4. Norris K.H, Rowan J.D. . Use of the automatic green-rot detector to improve the quality of liquid egg. Food Technology 1957, 11: 374~377.
5. Birth G.H., Norris K.H. . An instrument using light transmittance for non-destructive measurement of fruit maturity. Food Technology 1958, 12: 592~594.
6. Norris K.H., Hart J.R. . Direct spectrophotometric determination of moisture content of grain and seeds. In Proceeding. International Symposium on Humidity and Moisture 1963, 4: 19~25.
7. Weyer L.G. . Near-Infrared Spectroscopy of organic substances. Applied Spectroscopy Review 1985, 21(1, 2): 1~43.
8. Stark E., Luchter K. . Near-Infrared analysis: A technology for qualitative and quantitative analysis. Applied Spectroscopy Review 1986, 22(4): 335~339.
9. Ciurczak E.W. . Uses of Near-Infrared Spectroscopy in pharmaceutical analysis. Applied Spectroscopy Review 1987, 23(1, 2): 147~163.
10. Martin K.A. . Recent advances in Near-Infrared Reflectance Spectroscopy. Applied Spectroscopy Review 1992, 27(4): 325~383.
11. 陆婉珍, 袁洪福等. 现代近红外光谱分析技术. 北京: 中国石油化工出版社, 2000.
12. 金钦汉. 从匹兹堡会议看分析仪器的一些新动向. 分析仪器, 1999 (2): 1~5.
13. 金钦汉. 从 2000 年匹兹堡会议看分析化学和分析仪器的一些新动向. 现代科学仪器, 2000 (3): 14~16.
14. 陈焕文, 金钦汉. 2003 年匹兹堡会议简介. 现代科学仪器, 2003 (2): 41~45.
15. Davies T. . The history of near infrared spectroscopic analysis: Past, present and future “From sleeping technique to the morning star of spectroscopy”. Analysis 1998, 26(4): 17~19.
16. 吴秀琴. 应用 51A 型近红外测定小麦种子中赖氨酸的含量. 农业测试分析, 1985, 2 (2): 40~42.
17. 王文珍, 蒋自强, 林夕. 近红外光谱分析及其在我国的应用. 国外科学仪器, 1989.
18. 北京农业大学 FT 近红外 DRSA 研究组. 北京农业大学学报, 1991, 16: 1~4.
19. 袁洪福, 龙应成, 陆婉珍. 近红外光谱仪的研制. 分析化学, 1999, 27 (5): 608~614.
20. 刘国林, 陈国广, 相秉仁等. 近红外光谱技术在元胡止痛散定量分析中的初步应用研究. 中国现代应用药学杂志, 2000, 17 (5): 383~385.
21. 赵龙莲, 阎顺耕等. 傅里叶变换近红外光谱法测定烟草中 9 种品质参数. 光谱学与光谱分

- 析, 1998, 18 (4): 89~90.
22. Yan YanLu, Lao Cailan. Quantitative calibration transfer of gasoline in FTNIR analysis. Oral Presentation Abstracts NO. 980 at PITTCOM' 98, 1998.
23. Min ShunGeng, Zhou Xueqiu, Zhao Longlian et al. Can a lower S/N dispersion NIR instrument accurately determine the nitrogen in tobacco. Oral Presentation Abstracts NO. 1241 at PITTCOM' 98, 1998.
24. 袁洪福, 褚小立, 陆婉珍等. 一种新型在线近红外光谱分析仪的研制. 分析化学, 2004, 32 (2): 255~261.
25. 吉海彦, 严衍禄. 在国产近红外光谱仪实验样机上应用偏最小二乘法定量分析大麦成分. 分析化学, 1998, 26 (5): 607~611.
26. 徐广通, 陆婉珍, 袁洪福等. CCD 近红外光谱仪在柴油生产控制分析中的应用. 石油炼制与化工, 1999, 30 (9): 57~61.
27. 蒲登鑫, 王文茂, 李军会等. 近红外在线质量监控技术在中药葛根素生产中的应用. 现代仪器, 2003 (5): 27~29.
28. AACC. Approved methods of the AACC, Methods 39-10, 39-11, 46-30 and 08-01. St. Paul, Minn.: American Association of Cereal Chemists 2002.
29. AOAC. Official Methods of Analysis. (15th Ed.). Association of Official Analytical Chemists, Arlington, VA, 1990.
30. Anonymous. Standard methods of the International Association for Cereal Chemistry (ICC) ICC-standards 1998, Verlag Moritz Schäfer: Detmold.
31. 严衍禄, 赵龙莲, 韩冬海, 杨曙明. 近红外光谱分析基础与应用. 北京: 中国轻工业出版社, 2005.
32. 褚小立, 袁洪福, 陆婉珍. 近年来我国近红外光谱分析技术的研究与应用进展. 分析仪器, 2006, 2: 1~10.
33. 籍保平. 近红外光谱技术在农产品加工中的应用. 粮油加工与食品机械, 2000, 6: 31~33.
34. Blanco M., Villarroya I. . NIR spectroscopy: A rapid response analytical tool. Trends in Analytical Chemistry 2002, 21(4): 240~250.
35. 冯新泸, 史永刚. 近红外光谱及其在石油产品分析中的应用. 北京: 中国石化出版社, 2002.
36. 李峰. 山东道地药材金银花的多元多息指纹图谱鉴别研究 [博士论文]. 山东: 山东中医药大学, 2001.
37. 黄宝康, 朱斌, 郑汉臣等. 不同产地楮实子及伪品的近红外漫反射指纹图谱聚类分析. 中药材, 2002, 25(12): 874~875.
38. 孙素琴, 袁子民, 白雁. 红外指纹图谱与计算机辅助解析技术鉴别山药的道地形. 计算机与应用化学, 2002, 19 (1): 77~80.
39. 荆森, 金兴良, 庄峙夏等. 近红外漫反射光谱技术在甘草指纹图谱中的研究. 现代中药研究与实践, 2004, 18: 6~9.
40. 张龙翔, 张庭芳. 生物化学实验方法和技术, 北京: 人民教育出版社, 1982.

41. 张玉奎, 王杰, 张维冰. 实用高效液相色谱法的建立, 北京: 华文出版社, 2001, 第二版: 426~428.
42. Rivier J., McClintock R. . Reversed-phase high performance liquid chromatography of insulin from different species. *Journal of Chromatography* 1983, 268: 112~119.
43. Hsieh Y.L.F., Wang H., Elicone C. et al. Automated analytical system of the examination of protein primary structure. *Analytical Chemistry* 1993, 68: 455~462.
44. McPhie P. . *Methods Enzymol* 1971, 22: 23~26.
45. Agilent ZORBAX HPLC 柱在生物技术应用中的选择指南, Agilent Technologies.
46. 黄友如, 华欲飞, 裴爱泳. 差示扫描量热技术及其在大豆蛋白分析中的应用. 粮食加工, 2004, 2: 58~61.
47. 汪立君, 李里特, 张晓峰. 利用 DSC 对大豆蛋白质热变性的研究. *中国农业大学学报*, 2001, 6 (6): 93~96.
48. Sinead M. Fitzsimons, Daniel M. Mulvihill, Edwin R. Morris. Denaturation and aggregation processes in thermal gelation of whey proteins resolved by differential scanning calorimetry. *Food Hydrocolloids* 2007, 21: 638~644.
49. 吴瑾光. 近代傅立叶变换近红外光谱技术及应用(上卷), 科学技术文献出版社, 1994.
50. Tormod Næs, & Tomas Isaksson. Selection of samples for calibration using spectral information. *NIR News* 1995, 6: 1~13.
51. Tormod Naes, Tamas Isaksson. Locally weighted regression scatter correction for near infrared reflectance data. *Analytical Chemistry* 1990, 62: 664~673.
52. 严衍禄, 张录达. 傅里叶变换近红外漫反射光谱分析应用基础的研究, 北京农业大学学报, 1990, 16 (增刊): 5.
53. Savitsky A., & Golay Marcel J.E. . Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry* 1964, 36: 1627~1639.
54. Hamming R. W. . *Digital Filters*, Chapters 7 & 9, Prentice-Hall, 1983.
55. Karstang T.V., Kvalheim K. . Multivariate prediction and background correction using local modeling and derivative spectroscopy. *Analytical Chemistry* 1991, 63: 767~772.
56. 赵强, 张工力, 陈星旦. 多元散射校正对近红外光谱分析定标模型的影响光. 学精密工程, 2005, 13 (1): 53~58.
57. Barnes R.J., Dhanoa M.S., Lister S.J. . Standard normal variate transformation and de-trending of near infrared diffuse reflectance spectra. *Applied Spectroscopy* 1989, 43: 772~777.
58. Tahboub Y.R., Pardue H.L. . Evaluation of multiwavelength first and second-derivative spectra for the quantitation of mixtures of polynuclear aromatic hydrocarbons. *Analytical Chemistry* 1985, 57: 38~41.
59. Alsberg B.K., Woodward A.M., Kell D.B. . An introduction to wavelet transforms for chemometrists: A time-frequency approach. *Chemometrics and Intelligent Laboratory Systems* 1997, 37: 215~239.
60. Wold S., Antti H., Lindgren F., Ohman J. . Orthogonal signal correction of near infrared

- spectra. *Chemometrics and Intelligent Laboratory Systems* 1998, 44: 175~185.
61. Burns D.A., Ciurczak E.W. . “Handbook of near infrared analysis”, Marcel Dekker, 1992.
62. Weisberg S. . *Applied linear regression*. John Wiley and Sons, 1985.
63. Maesschalck R. De, Candolfi A. . Decision criteria for soft independent modeling of class analogy applied to near infrared data. *Chemometrics and Intelligent Laboratory Systems* 1999, 47: 65~77.
64. Girosi F., Makhoul J., Manolakos E. & Wilson E. . *Neural networks for signal processing*. IEEE Press, New York, 1995.
65. 祝诗平, 王一鸣, 张小超等. 近红外光谱建模异常样品剔除则与方法. *农业机械学报*, 2004, 35(4): 115~119.
66. 中华人民共和国药典 2005 年版, 附录 XIXK. 近红外分光光度法指导原则.
67. 牛智有. 鱼粉、精料补充料及其中肉骨粉含量的近红外漫反射光谱分析 [博士论文]. 北京: 中国农业大学, 2005.
68. Thomas E.V., Haaland D.M. . Comparison of multivariate calibration methods for quantitative spectral analysis. *Analytical Chemistry* 1990, 62: 1091~099.
69. Thomas E.V. . A primer on multivariate calibration. *Analytical Chemistry* 1994, 66: 795~804.
70. Wu W., Walczak B., Massart D.L. . Spectral transformation and wavelength selection in near infrared spectra classification. *Analytica Chimica Acta* 1995, 315: 243~255.
71. 王艳斌, 袁洪福, 陆婉珍. 近红外分析方法测定润滑油基础油粘度指数. *润滑油*, 2001, 16: 53~56.
72. 王芳, 陈达, 邵学广. 近红外光谱与卷烟样品常规成分的关系模型研究. *烟草科技*, 2002, 5: 23~26.
73. Guo Jun. Development of dispersive and Fourier transform near infrared spectroscopy methodology for food and edible seed analysis [Dissertation]. USA: University of Illinois at Urbana-Champaign, 2004.
74. GB/T 18634-2002. 饲用植酸酶活性的测定分光光度法.
75. 刘蓉, 陈文亮, 徐可欣等. 奇异点快速检测在牛奶成分近红外光谱测量中的应用. *光谱学与光谱分析*, 2005, 25(2): 207~210.
76. 褚小立, 袁洪福, 陆婉珍. 基础数据准确性对近红外光谱分析结果的影响. *光谱学与光谱分析*, 2005, 25(6): 886~889.
77. 吴坚. *应用概率统计*. 北京: 高等教育出版社, 2003.
78. 高荣强, 范世福, 严衍禄等. 近红外光谱的数据预处理方法研究. *光谱与光谱分析*, 2004, 24 (12): 1563~1565.
79. 邹龙贵. 植酸酶预混剂微粒的制备工艺与设备. *饲料工业*, 2007, 28 (1): 8~10.
80. Barnes R.J., Dhanoa M.S., & Lister S.J. . Standard normal variance transformation and de-trending of near infrared diffuse reflectance spectra. *Applied Spectroscopy* 1989, 43: 772~777.
81. Murray I., & Williams P.C. . Chemical principles of near infrared technology. In *Proceeding*.

- Williams & K. Norris (Eds.), near infrared technology in the agricultural and food industries. St. Paul, MN: American Association of Cereal Chemists, 1990.
82. Musleh Uddin, Emiko Okazaki, Moin Uddin Ahmad et al. . NIR spectroscopy: A non-destructive fast technique to verify heat treatment of fish-meat gel Food Control 2005: 1~5.
83. Uddin, M., Ishizaki, S., Okazaki, E., & Tanaka, M. . Near infrared reflectance spectroscopy for determining end-point temperature of heated fish and shellfish meats. Journal of the Science of Food and Agriculture 2002, 82: 286~292.

## 致 谢

在论文即将完成之际，深深地感谢那些曾教导和鼓励我的师长、帮助我的朋友、支持和关怀我的亲人。

本论文是在导师秦玉昌研究员的悉心指导下完成的。从论文选题及科研实践和论文编写的工作中，都得到了秦老师的精心指导和启发，使本人受益匪浅。秦老师严谨求实的科研态度、忘我工作的工作作风、朴实无华的生活作风令我深深敬佩，也必将始终激励我在今后的工作和生活中奋发进取、锐意进取。除了科研上的悉心指导之外，更重要的是从秦老师身上我学会了许多做人、做事的道理，这些都让我受益终生。在此特向秦老师表示崇高的敬意和深深的谢意。

在本课题的实验设计和实验过程中，中心分析试验室的张萍研究员，课题组的李军国研究员、吕小文博士、李俊副研究员、牛力斌老师、董颖超博士、黄火清博士、李辉硕士等都给予了我无私的帮助和花费了大量的心血。张萍研究员在科研过程中对我的启发和诸多指导，李军国研究员和吕小文博士诲人不倦的育人风范和实验全过程的悉心指导，李俊副研究员严谨的治学态度和工作作风，牛力斌老师和董颖超博士在生活上的诸多关照都让我时刻铭记在心；所中心分析实验室的常碧影老师、刘庆生老师、范志影老师和司进老师在实验方面为我创造了诸多便利条件；我的师弟谭永辉硕士、王海东硕士、王兴敏硕士和师妹梁克红硕士在论文的开展中和生活中都给予了无私的帮助；在此一并向他们表示诚挚的谢意。同时也衷心地感谢所有曾帮助过我的老师和同学们！

最后感谢我的家人和朋友多年来对我的关心和支持。

杨海锋

2007年6月

于中国农业科学院饲料研究所

## 作者简历

杨海锋，男，汉族，中共党员。1981年10月6日出生于湖北省赤壁市。

2000年毕业于湖北省赤壁市车埠高级中学，同年以优异的成绩考入湖北省武汉工业大学(动物营养与饲料科学专业)。2004年7月毕业，获工学学士学位，在校四年间，两次荣获“校优秀学生干部”称号、两次荣获“优秀团员”称号、一次荣获“社会实践先进个人”等称号，多次获得校级奖学金和企业奖学金(“唐人神”奖学金)。大学毕业论文“猪肠膜蛋白粉生产工艺研究”于2007年获中国粮油学会第五届优秀论文奖。

2004年9月考入中国农业科学院饲料研究所攻读硕士学位，在秦玉昌研究员的指导下，从事近红外光谱技术应用研究。在此期间两次荣获“中国农业科学院优秀共产党员”称号，一次荣获“社会活动积极分子”称号。

硕士期间发表论文：

1. Near-infrared spectroscopy-based methods for phytase applications in animal feeds, *J. Agric. Food Chem.*, in submission.
2. 杨海锋，吕小文，秦玉昌等. 近红外光谱分析技术在豆粕质量监控中的应用研究，*饲料工业*，2006，27(19)：31~34.
3. 秦玉昌，杨海锋，吕小文. 疯牛病的安全风险评估和监控技术研究动态，*饲料工业*，2006，27(13)：1~4.