

Abstract

Classification and determination of total protein in mussel *Mytella Falcata* collected in Brazilian river estuaries using NIR and multivariate calibration

G.Q. de Brito, J.H.S. Junior, M.F.V. Moura and K.M.G. Lima*

Department of Chemical, UFRN, Natal, 59072-970, Brazil

*Corresponding author: kassio@ufrnet.br

Introduction

Mussel is the term used to name the various species of bivalve mollusks of the family Mytilidae (*Mytella*). Seafood (mussel) is commonly consumed in coastal regions and is of significant nutritional importance because they constitute sources of dietary protein and minerals. The Kjeldahl method has been used for decades to determine protein levels in sample matrices, such as mussels. This procedure is time-consuming and tedious due to various analytical steps: digestion, distillation and titration. Thus, the aim of this study was classification determination of total protein content of mussel samples collected from five river estuaries on the Brazilian coast using NIR and multivariate calibration methods (PCR, PLS, MLR-SPA and SIMCA).

Materials and Methods

Initially, a portion of mussel powder samples was analysed to determine protein content using the Kjeldahl method; performed in triplicate. After, the spectra were recorded in triplicate using an MB-160 D FT-NIR spectrophotometer (Bomem), with spectral resolution of 8 cm^{-1} and 50 scans in the NIR range of 714 – 2500 nm. Calibration and validation samples were used to model the procedures, whereas the predicted samples were only used for final evaluation. Spectrum differentiation, Savitzky-Golay smoothing and derivatives, principal components analysis (PCA), PCR, PLS, SIMCA were carried out in Unscrambler® 9.8 (CAMO S.A). The *F*-test for SIMCA classification was conducted at a 5% significance level.

Results and Discussion

Principal component regression (PCR) and partial least squares (PLS) multivariate calibrations were used to predict total protein. For comparison purposes, the successive projection algorithm for variable selection was applied to the original set. In the prediction stage the MLR-SPA model obtained an R^2 of 0.90 and root mean square error (RMSEP) of 1.56 g/100 g. Soft independent modeling of class analogy (SIMCA) was also used for full-spectrum classification, resulting in almost 100% classification accuracy based on estuary localisation, regardless of the significance level adopted for the *F*-test.

Conclusion

NIR spectroscopy combined with chemometric data evaluation is a good tool for determining total protein in mussel powder found in Brazilian river estuaries and coupled with SIMCA models resulted in 100% classification accuracy.