

UNIVERSIDADE FEDERAL DE MINAS GERAIS

INSTITUTO DE CIÊNCIAS EXATAS

DEPARTAMENTO DE QUÍMICA

Miguel Daniel Souza Monte

**CONSTRUÇÃO DE MODELO QUIMIOMÉTRICO PARA
QUANTIFICAÇÃO DE MELAMINA EM *WHEY PROTEIN* ATRAVÉS
DE ESPECTROSCOPIA NO INFRAVERMELHO PRÓXIMO**

Belo Horizonte

2026

Miguel Daniel Souza Monte

**CONSTRUÇÃO DE MODELO QUIMIOMÉTRICO PARA
QUANTIFICAÇÃO DE MELAMINA EM *WHEY PROTEIN* ATRAVÉS
DE ESPECTROSCOPIA NO INFRAVERMELHO PRÓXIMO**

Trabalho de Conclusão de Curso
apresentado ao Departamento de
Química do Instituto de Ciências
Exatas, da Universidade Federal de
Minas Gerais, como requisito para a
obtenção do título de Bacharel em
Química.

Orientador(a): Bruno Gonçalves
Botelho

Belo Horizonte

2026

Ata da sessão de avaliação do Trabalho de Conclusão de Curso de Bacharelado em Química da Universidade Federal de Minas Gerais, do aluno Miguel Daniel Souza Monte, nº de registro 2022037353.

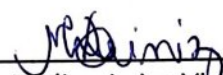
O Trabalho de Conclusão de Curso do aluno **Miguel Daniel Souza Monte**, intitulado **"Construção de modelo quimiométrico para quantificação de melamina em whey protein através de espectroscopia no infravermelho próximo"**, foi avaliado e aprovado pela Banca Examinadora constituída por Profa. Alessandra Schneider Henn (DQ - UFMG), Ma. Melina Luiza Vieira Diniz (DQ - UFMG), e Prof. Bruno Gonçalves Botelho (DQ - UFMG), orientador do trabalho de conclusão de curso, no dia 22 de junho de 2026.



Prof. Bruno Gonçalves Botelho



Profa. Alessandra Schneider Henn



Ma. Melina Luiza Vieira Diniz



Prof. Dario Windmüller / Profa. Dalva Ester da Costa Ferreira
Coordenador dos Cursos de Graduação em Química e Química Tecnológica
ICEx - UFMG

AGRADECIMENTOS

Agradeço, primeiramente, a Deus, pois estar aqui neste momento, realizando esta conquista, foi um pedido recorrente em minhas orações.

Aos meus pais, que me ensinaram que o estudo é a melhor forma de vencer na vida.

À minha namorada, por acreditar em mim mais do que eu mesmo jamais acreditei.

Agradeço ao meu orientador, Bruno, por me ajudar a conceber este projeto e a compreender, com paciência e humildade, os conceitos utilizados neste trabalho. Agradeço também a todos os colegas de laboratório, que tornaram o preparo das amostras mais divertido e descontraído.

Agradeço a todos os professores que contribuíram para minha formação profissional e a todos os amigos que fiz durante os breves quatro anos em que estive no curso de Química da UFMG.

Por fim, agradeço a todos que lerão este trabalho futuramente.

”Somos o que fazemos repetidamente. A excelência,
portanto, não é um ato, mas um hábito” Aristóteles

RESUMO

Este trabalho teve como objetivo desenvolver uma metodologia analítica rápida, não destrutiva e de baixo custo para a quantificação de melamina - composto sintético rico em nitrogênio usado como adulterante em fórmulas infantis na China em 2008 - em amostras de whey protein, concentrado de proteína extraída a partir do soro do leite, utilizando espectroscopia no infravermelho próximo associada à calibração multivariada por mínimos quadrados parciais. Para isso, foram utilizadas 7 amostras comerciais autênticas de whey protein, de diferentes marcas e sabores, adulteradas intencionalmente com melamina em 8 níveis crescentes de concentração, totalizando 56 amostras. O conjunto de dados foi dividido em 42 amostras para calibração e 14 para previsão externa. Os espectros obtidos foram avaliados após diferentes estratégias de pré-processamento, incluindo MSC, SNV e derivadas e alisamento de Savitzky-Golay, sendo o MSC selecionado por apresentar desempenho adequado com menor número de variáveis latentes. O modelo PLS otimizado, construído com 9 variáveis latentes, apresentou coeficiente de determinação de 0,989 para calibração, 0,971 para validação cruzada e 0,975 para previsão externa, além de RMSEC de 1,1% m/m, RMSECV de 2,0% m/m e RMSEP de 2,2% m/m. Também foram estimados o Limite de Decisão ($CC\alpha$) e a Capacidade de Detecção ($CC\beta$), com valores de 2,6% m/m e 5,0% m/m, respectivamente. Os resultados demonstram que a espectroscopia no infravermelho próximo, combinada à quimiometria, apresenta potencial para a quantificação de melamina em whey protein e pode ser empregada como ferramenta preliminar no controle de qualidade de suplementos proteicos.

Palavras-chave: whey protein; melamina; adulteração; espectroscopia no infravermelho próximo; quimiometria.

SUMÁRIO

1	INTRODUÇÃO	7
2	OBJETIVO	12
3	REFERENCIAIS TEÓRICOS.....	13
4	PARTE EXPERIMENTAL	25
5	RESULTADOS E DISCUSSÃO	30
6	CONCLUSÃO	47
	REFERÊNCIAS.....	49
	APÊNDICE 1.....	52

1 INTRODUÇÃO

O mercado de suplementos alimentares tem apresentado grande crescimento nos últimos anos, impulsionado pela busca por qualidade de vida e estética. Dentre esses produtos, o *whey protein* destaca-se como um dos mais consumidos globalmente. Dados da plataforma Google Trends demonstram não apenas o crescimento contínuo no interesse por esse termo ao longo dos últimos anos (Figura 1), mas também destacam o Brasil como o país com o maior volume relativo de buscas no mundo (Figura 2)(GOOGLE, 2026). Esse cenário de alta demanda, associado ao alto valor agregado do produto, torna o mercado brasileiro particularmente suscetível a práticas fraudulentas visando ao lucro econômico.¹

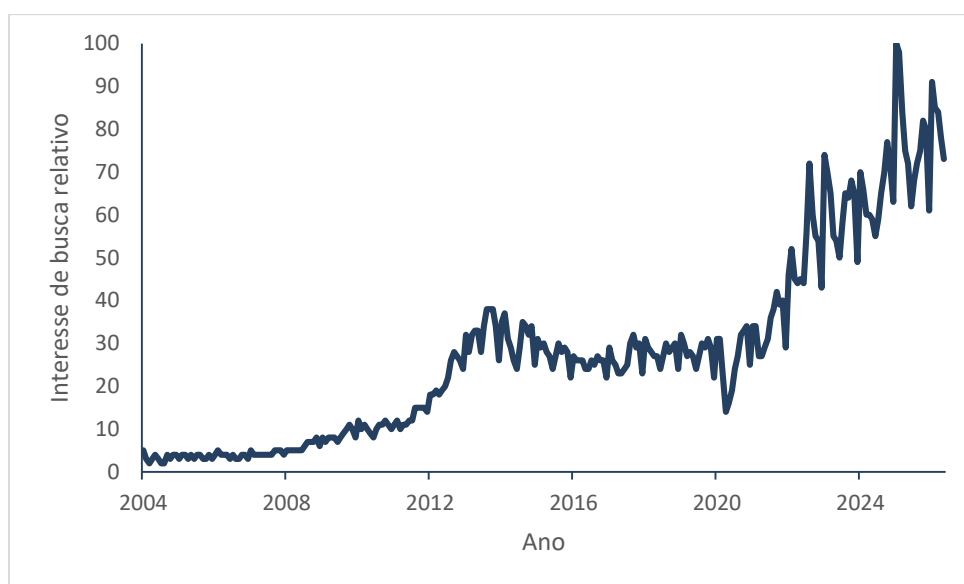


Figura 1: Interesse de busca, no Brasil, pelo termo "Whey" na plataforma Google.

Evolução do interesse de busca global entre 2004 e 2025. **Fonte:** GOOGLE TRENDS

¹ Em atendimento aos princípios recomendados pela UFMG para o uso de Inteligência Artificial (IA) em pesquisas, o autor declara, para fins de transparência, que a ferramenta GEMINI (GOOGLE) foi utilizada na etapa de redação deste documento exclusivamente como apoio para correção ortográfica e aprimoramento da coesão e coerência textual. Além disso, a ferramenta NOTEBOOK LM (GOOGLE) foi usada para direcionamento e esclarecimento de dúvidas em artigos previamente selecionados pelo autor. Todos os resultados foram devidamente revisados pelo autor.



Figura 2: Interesse de busca por região, com o Brasil na primeira posição. Os dados são apresentados em uma escala relativa (0 a 100) normalizada pelo volume máximo, permitindo a comparação entre os países. **Fonte:** GOOGLE TRENDS

O *whey protein* é um suplemento proteico extraído do soro do leite durante a fabricação do queijo. O produto é vendido no mercado como mistura concentrada (WPC), que pode conter de 50% a 85% de proteína (JIAO et al., 2019), ou como produto isolado (WPI), com teor proteico superior a 90% (ANDRADE et al., 2019). Por ser um produto caro e em alta demanda, torna-se um alvo atrativo para fraudadores que buscam lucro fácil (JIAO et al., 2019). Um exemplo recente dessa realidade ocorreu em setembro de 2025, quando uma operação policial no interior de São Paulo apreendeu 4 toneladas de suplementos adulterados, incluindo *whey protein*, que eram comercializados em grandes plataformas digitais (G1, 2025). Incidentes como esse evidenciam que a adulteração não é apenas um risco teórico, mas um problema de segurança alimentar ativo no país, reforçando a urgência de métodos de fiscalização mais ágeis e eficientes. A forma mais comum de adulteração envolve a adição intencional de substâncias de baixo custo e ricas em nitrogênio, como ureia, melamina e aminoácidos mais baratos, misturadas à matriz proteica original (JIAO et al., 2019). A melamina, em particular, é um adulterante estratégico para os fraudadores devido à sua estrutura química, que contém cerca de 66% de nitrogênio em massa, uma proporção significativamente alta capaz de elevar artificialmente os níveis de nitrogênio detectados nas análises (MECKER et al., 2012).

A persistência e o sucesso inicial dessas fraudes devem-se, em grande parte, às limitações dos métodos analíticos clássicos utilizados no controle de qualidade, com destaque para o método de Kjeldahl. Embora seja o padrão oficial para a determinação de proteínas,

esse método não quantifica as moléculas de proteína diretamente, em vez disso, ele determina o teor de nitrogênio total da amostra e aplica um fator de conversão genérico para estimar a concentração proteica. A grande falha dessa abordagem reside na sua incapacidade de distinguir entre o nitrogênio proveniente de fontes proteicas verdadeiras e o nitrogênio não proteico oriundo de adulterantes sintéticos. Consequentemente, amostras adulteradas com compostos nitrogenados apresentam resultados falsamente positivos de alto teor proteico, enganando a análise e chegando ao consumidor como produtos aparentemente conformes (AGUIRRE, 2023; JIAO *et al.*, 2019; MECKER *et al.*, 2012).

Além disso, é importante entender que a melamina é uma substância química sintética, comumente utilizada na indústria para a produção de resinas com propriedades retardantes de chama (MECKER *et al.*, 2012). Sua ingestão apresenta riscos severos e agudos à saúde (CHENG *et al.*, 2010). O mecanismo de toxicidade está associado à sua interação com o ácido cianúrico no organismo, resultando na formação de cristais insolúveis nos túbulos renais. A deposição desses cristais pode levar rapidamente a quadros clínicos graves, incluindo a formação de cálculos urinários, insuficiência renal aguda e, em casos extremos, ao óbito do consumidor (MECKER *et al.*, 2012).

A gravidade relacionada a adulterações por meio do uso desse contaminante ganhou atenção global com o escândalo do leite em pó contaminado na China, ocorrido em setembro de 2008 (ESTADOS UNIDOS, 2008). Nesse episódio, que se tornou um marco na história da segurança alimentar, a adulteração intencional de fórmulas infantis resultou em dezenas de milhares de crianças hospitalizadas e algumas mortes confirmadas, evidenciando o impacto devastador que a fraude econômica pode ter na saúde pública (MECKER *et al.*, 2012). Como consequência, avaliações de risco realizadas pela agência americana *Food and Drug Administration* (FDA) estabeleceram níveis de ingestão diária tolerável extremamente baixos para garantir a segurança dos consumidores (ESTADOS UNIDOS, 2008).

Acompanhando as diretrizes globais, a Agência Nacional de Vigilância Sanitária (ANVISA) atualizou as regulamentações nacionais por meio da Instrução Normativa nº 351, de 18 de março de 2025, estabelecendo o limite máximo tolerado de 2,5 mg/kg para a presença de melamina em alimentos em geral, com limites ainda mais rigorosos para fórmulas infantis – 1,0 e 0,20 mg/kg para fórmulas em pó e líquidas respectivamente – (BRASIL, 2025). A existência de um limite legal rigoroso reforça a necessidade de vigilância sanitária constante e, consequentemente, de metodologias que permitam a fiscalização eficiente desses teores no mercado nacional.

Em resposta aos riscos associados à adulteração, diversas técnicas analíticas de referência foram estabelecidas para a detecção de melamina e outros adulterantes nitrogenados. Entre as metodologias mais empregadas, destacam-se a Cromatografia Líquida de Alta Eficiência (HPLC), a Cromatografia Líquida acoplada à Espectrometria de Massas (LC-MS) e a Cromatografia Gasosa acoplada à Espectrometria de Massas (GC-MS) (CHENG *et al.*, 2010). Embora essas técnicas instrumentais ofereçam excelentes limites de detecção e alta precisão analítica, sua aplicação rotineira enfrenta barreiras logísticas significativas (LU *et al.*, 2009; CHENG *et al.*, 2010).

A principal limitação desses métodos convencionais reside na sua complexidade operacional. São análises geralmente onerosas e demoradas, que demandam etapas trabalhosas de preparo de amostra (CHENG *et al.*, 2010). Além disso, a necessidade de equipamentos robustos de laboratório torna métodos como o HPLC pouco práticos para demandas de resposta rápida, sendo frequentemente descritos como lentos e trabalhosos para o fluxo intenso da indústria (RODRIGUES JÚNIOR *et al.*, 2016). Tais características tornam essas técnicas inviáveis para inspeções rápidas *in situ* ou para a triagem de produtos no mercado (CHENG *et al.*, 2010).

Diante desse cenário, nota-se uma lacuna tecnológica e uma necessidade de métodos alternativos de triagem para supri-la. A indústria e os órgãos fiscalizadores carecem de ferramentas analíticas que sejam simultaneamente rápidas, de baixo custo e capazes de fornecer resultados qualitativos confiáveis sem a necessidade de tratamentos prévios exaustivos da amostra.

Para superar as limitações das metodologias analíticas clássicas e atender à crescente demanda por fiscalizações mais ágeis, as técnicas espectroscópicas têm se destacado como alternativas promissoras (POONIA *et al.*, 2017; YANG *et al.*, 2022). Nesse contexto, a espectroscopia no infravermelho próximo (NIR) vem como uma ferramenta analítica de grande relevância. Trata-se de uma técnica rápida, não destrutiva e que se adequa à química verde (SILVA *et al.*, 2024) devido ao fato da técnica exigir pouco ou nenhum preparo prévio de amostra (POONIA *et al.*, 2017), dispensando o uso de solventes orgânicos, reagentes químicos e a geração de resíduos tóxicos (SILVA *et al.*, 2024).

Graças a essas características, a espectroscopia NIR apresenta um elevado potencial para aplicação em análises de triagem rotineiras na indústria e em órgãos de fiscalização para a garantia da segurança alimentar (YANG *et al.*, 2022). A capacidade de realizar leituras diretas, associada à celeridade da técnica, permite a avaliação de um grande volume de amostras num curto período de tempo, otimizando o fluxo de trabalho. Dessa forma,

torna-se viável a identificação de lotes suspeitos de adulteração que, caso seja necessário, podem ser posteriormente confirmados através dos métodos analíticos de referência (SILVA *et al.*, 2024). Além disso, neste trabalho foi usado um equipamento NIR portátil, o que traz mais uma vantagem ao permitir análises *in situ*.

Contudo, a aplicação do NIR na análise de matrizes alimentares complexas, como os derivados lácteos e o *whey protein*, apresenta desafios. Os espectros obtidos nessa região são caracterizados por bandas largas e frequentemente sobrepostas, resultantes de sobretons e bandas de combinação de vibrações moleculares (LU *et al.*, 2009; SILVA *et al.*, 2024). Dessa forma, a extração de informações visuais diretas sobre a presença de adulterantes como a melamina torna-se impraticável apenas pela observação do espectro bruto (POONIA *et al.*, 2017; LU *et al.*, 2009). Para superar essa barreira, é indispensável o uso da técnica instrumental aliada a ferramentas quimiométricas (POONIA *et al.*, 2017; SILVA *et al.*, 2024). A quimiometria, por meio da aplicação de modelos matemáticos e estatísticos multivariados, atua no reconhecimento de padrões e na extração de informações latentes dos espectros (POONIA *et al.*, 2017), viabilizando o desenvolvimento de modelos preditivos capazes de identificar e quantificar adulterantes em baixas concentrações para a garantia da segurança alimentar (SILVA *et al.*, 2024; YANG *et al.*, 2022).

Diante da vulnerabilidade dos métodos clássicos e da necessidade de fiscalização ágil, o presente trabalho se insere na busca por inovações tecnológicas para a garantia da segurança alimentar. Assim, propõe-se explorar o potencial da espectroscopia NIR portátil combinada a ferramentas quimiométricas como uma estratégia eficiente e não destrutiva para o controle e triagem de fraudes por adição de melamina em matrizes proteicas.

2 OBJETIVO

O objetivo deste trabalho foi construir um modelo quimiométrico quantitativo usando Regressão de Mínimos Quadrados Parciais (PLS) e espectroscopia do infravermelho próximo portátil (NIR) que permita análises preliminares rápidas de adulteração de *whey protein* com melamina.

3 REFERENCIAIS TEÓRICOS

3.1. *Whey protein*

O *whey protein*, proteína concentrada extraída do soro do leite, é atualmente um dos suplementos alimentares proteicos mais consumidos e estudados. Historicamente, o soro do leite era considerado um subproduto residual e de baixo valor agregado da indústria de laticínios, gerado em grandes volumes durante o processo de coagulação do leite para a precipitação da caseína na fabricação de queijos (ZADOW, 1992). No entanto, com o avanço tecnológico no processamento de alimentos, especialmente com o desenvolvimento de técnicas de separação por membranas — como a ultrafiltração — tornou-se viável isolar e concentrar as proteínas solúveis presentes nesse líquido de forma eficiente e sem desnaturá-las (HARAGUCHI *et al.*, 2006; ZADOW, 1992).

Comercialmente, o *whey protein* é classificado em duas categorias principais. O *Whey Protein Concentrado* (WPC) contém de 50% a 85% de proteína. Já o *Whey Protein Isolado* (WPI) passa por etapas adicionais de purificação, o que lhe confere um grau de pureza mais elevado, apresentando teor proteico superior a 90% e concentrações mínimas ou quase nulas de gorduras e carboidratos.

A composição química do *whey protein* destaca-se por seu elevado valor nutricional. Essas proteínas são caracterizadas por possuírem uma boa digestibilidade, rápida absorção gastrointestinal e um perfil completo de aminoácidos essenciais. Dentre estes compostos, destacam-se em altas concentrações os aminoácidos de cadeia ramificada (BCAA – leucina, isoleucina e valina), que desempenham um papel crucial na sinalização da síntese proteica muscular e na recuperação de tecidos após stress físico (HARAGUCHI *et al.*, 2006). Ademais, além de seu caráter estrutural, a proteína do soro é rica em frações peptídicas bioativas, como a beta-lactoglobulina, a alfa-lactoalbumina, a lactoferrina e diversas imunoglobulinas. Esses compostos conferem ao suplemento propriedades funcionais significativas, incluindo modulação do sistema imunológico, ação antioxidante e hipotensora (HARAGUCHI *et al.*, 2006).

Do ponto de vista comercial e econômico, o *whey protein* vai além do nicho de atletas de alto rendimento e já é muito consumido por indivíduos que buscam melhora na composição corporal, longevidade e qualidade de vida (HARAGUCHI *et al.*, 2006). Para muitos, trata-se de uma alternativa viável para atingir o consumo mínimo recomendado de proteínas diárias de forma rápida e conveniente. Esse aumento contínuo na demanda global conferiu a esse produto um alto valor agregado e de mercado. No entanto, essa alta

rentabilidade e procura pelo produto também o tornam um alvo preferencial e altamente suscetível a práticas fraudulentas. Visando a maximização dos lucros, tornou-se comum na indústria clandestina a adição de adulterantes ricos em nitrogênio à matriz do produto, mascarando a sua verdadeira constituição nutricional (JIAO et al., 2019;).

3.2. Melamina, o adulterante

3.2.1. O histórico de fraudes e o escândalo de 2008

A adulteração de alimentos, acidentalmente ou intencionalmente, é uma realidade recorrente na indústria dos alimentos (INGELFINGER, 2008). Um dos exemplos mais emblemáticos de fraude alimentar contemporânea ocorreu na China, em 2008, envolvendo a contaminação de fórmulas infantis em pó com melamina (INGELFINGER, 2008; POONIA *et al.*, 2017).

Nesse episódio, o consumo de produtos lácteos adulterados de aproximadamente 22 marcas atingiu severamente a saúde pública. Estima-se que os danos renais tenham atingido quase 300.000 crianças, resultando em dezenas de milhares de internações e seis óbitos atestados (INGELFINGER, 2008). O evento gerou forte repercussão global, fazendo com que diversos países suspendessem a compra de laticínios da China. Consequentemente, agências reguladoras internacionais se viram forçadas a reavaliar com urgência seus padrões de tolerância para contaminantes químicos (ESTADOS UNIDOS, 2008; INGELFINGER, 2008).

3.2.2. Propriedades Químicas da Melamina

A melamina (1,3,5-triazina-2,4,6-triamina) é uma substância química sintética, inodora e insípida, pouco solúvel em água, rica em nitrogênio, utilizada em aplicações industriais, como a fabricação de laminados, resinas, colas e retardantes de chama (INGELFINGER, 2008). Seu uso para práticas fraudulentas é possível graças a sua estrutura química: a molécula de melamina é composta por aproximadamente 66% de nitrogênio em massa (INGELFINGER, 2008; MECKER et al., 2012). Sua estrutura está representada na Figura 3 a seguir:

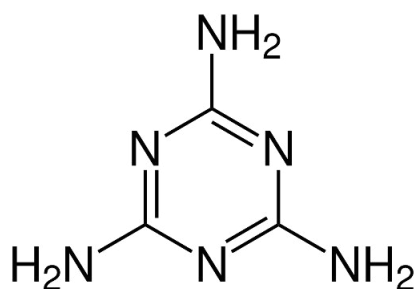


Figura 3: Estrutura química da melamina. **Fonte:** WIKIPEDIA (2026)

Como os métodos convencionais de determinação proteica, como o de Kjeldahl, estimam o teor de proteína indiretamente através da medição do nitrogênio total (AGUIRRE, 2023; INGELFINGER, 2008), a adição de melamina permite elevar artificialmente o valor proteico medido de um produto (POONIA *et al.*, 2017). Dessa forma, a melamina é utilizada por fraudadores para mascarar a diluição de leite ou suplementos, conferindo-lhes uma conformidade nutricional falsa perante os testes laboratoriais clássicos de controle de qualidade (CHENG *et al.*, 2010; INGELFINGER, 2008).

A toxicidade da melamina em seres humanos está diretamente ligada à formação de cristais insolúveis no sistema urinário (INGELFINGER, 2008; ESTADOS UNIDOS, 2008). Embora a melamina isolada possua baixa toxicidade aguda, sua periculosidade aumenta drasticamente quando ingerida simultaneamente com seus análogos estruturais, como o ácido cianúrico, frequentemente presente como um co-contaminante nas fraudes (ESTADOS UNIDOS, 2008).

De acordo com avaliações de risco conduzidas pela Food and Drug Administration (ESTADOS UNIDOS, 2008), a interação entre a melamina e o ácido cianúrico resulta na formação de um complexo estável através de ligações de hidrogênio, conhecido como cianurato de melamina. Esse complexo é insolúvel e precipita-se na forma de cristais nos túbulos renais (ESTADOS UNIDOS, 2008; INGELFINGER, 2008). A deposição contínua desses cristais causa obstrução tubular, provocando a formação de cálculos renais e, em casos graves, progredindo para insuficiência renal aguda e óbito (ESTADOS UNIDOS, 2008; INGELFINGER, 2008).

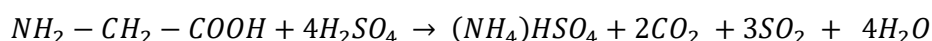
3.3. Métodos Analíticos Oficiais e suas Limitações

3.3.1. O Método de Kjeldahl: Princípios e Mecanismos Químicos

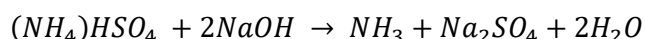
O método de Kjeldahl, desenvolvido há mais de um século, permanece como uma técnica muito utilizada e estudada para a determinação de nitrogênio, baseando-se quimicamente na digestão da amostra em ácido sulfúrico (AGUIRRE, 2023). No contexto da análise de alimentos, como laticínios, essa metodologia continua sendo uma ferramenta de referência para a determinação do teor de proteína. O procedimento experimental é dividido em três etapas consecutivas: conversão (frequentemente chamada de digestão ou mineralização), separação por destilação e determinação quantitativa via titulação (AGUIRRE, 2023).

A primeira etapa consiste em digerir a amostra orgânica utilizando ácido sulfúrico concentrado. Sob essas condições de digestão, o ácido atua degradando completamente a estrutura orgânica da amostra, convertendo-a em dióxido de carbono (CO_2), água (H_2O) e bissulfato de amônio $(\text{NH}_4)\text{HSO}_4$ (AGUIRRE, 2023).

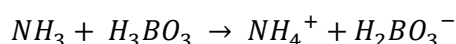
Para aumentar o ponto de ebulição e, conseqüentemente, elevar a temperatura da digestão, realiza-se a adição de um sal, sendo o sulfato de potássio (K_2SO_4) historicamente o sal de escolha, embora o sulfato de sódio também possa ser utilizado com o mesmo fim. Além do sal, utiliza-se um catalisador para otimizar o tempo de oxidação; o sulfato de cobre (CuSO_4) tornou-se amplamente recomendado em substituição ao mercúrio, evitando assim os problemas de alta toxicidade associados a esse (AGUIRRE, 2023). A reação química simplificada da fase de mineralização, utilizando um aminoácido (glicina) como modelo estrutural orgânico, pode ser descrita pela seguinte equação, conforme demonstrado por Aguirre (2023):



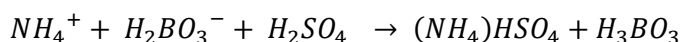
Concluída a etapa de mineralização da amostra, o método avança para a separação do íon amônio. Esse procedimento exige a alcalinização da solução ácida digerida por meio da adição de hidróxido de sódio (NaOH) em excesso, o que promove a volatilização do sal de amônio na forma de gás amônia (NH_3) (AGUIRRE, 2023). A reação de alcalinização ocorre conforme demonstrado por Aguirre (2023):



Atualmente, a rotina laboratorial adota a modificação proposta por Winkler, datada de 1913. Nessa adaptação, a amônia volatizada é conduzida (destilada) até um frasco receptor contendo uma solução de ácido bórico a 3% (AGUIRRE, 2023). Para visualizar o ponto de viragem nessa variante, a literatura reporta o uso de diversos indicadores de pH, sendo bastante comum a aplicação de indicadores mistos, destacando-se a combinação de verde de bromocresol e vermelho de metila (AGUIRRE, 2023). A captura da amônia pelo ácido bórico gera o íon borato, conforme descrito na equação a seguir (AGUIRRE, 2023):



A etapa conclusiva do método de Kjeldahl baseia-se na quantificação do nitrogênio por meio da determinação dos íons amônio (AGUIRRE, 2023). Quando se emprega a modificação do ácido bórico, o destilado é titulado com um ácido forte padronizado, frequentemente o ácido sulfúrico (AGUIRRE, 2023). A neutralização ocorrida nessa etapa final segue a estequiometria demonstrada na equação a seguir (AGUIRRE, 2023):



O volume do ácido titulante gasto permite calcular a concentração de nitrogênio total. Por fim, através de fatores de conversão matemáticos tabelados, o teor de nitrogênio obtido é traduzido na proporção final de proteína bruta presente na matriz analisada (AGUIRRE, 2023).

Apesar da sua importância, esse método apresenta uma vulnerabilidade crítica quando aplicado ao controle de fraudes: ele não é capaz de distinguir entre o nitrogênio proveniente de ligações peptídicas (fontes proteicas verdadeiras) e o nitrogênio não proteico oriundo de compostos sintéticos (INGELFINGER, 2008; JIAO et al., 2019; POONIA et al., 2017).

Para superar essa limitação analítica e detectar de maneira específica a presença da melamina e de outros compostos adulterantes, laboratórios e agências reguladoras recorrem a métodos analíticos mais seletivos (POONIA et al., 2017). As metodologias convencionais e oficiais para essa finalidade incluem a Cromatografia Líquida de Alta Eficiência (HPLC) e a Espectrometria de Massas (LC-MS/MS e GC-MS). Porém, embora sejam tecnologias altamente precisas, elas apresentam limitações operacionais: exigem processos de preparo de amostra complexos, são mais caras e demoradas. Devido a essa lentidão, ao alto custo e ao nível de especialização técnica exigido, tais métodos oficiais não

se mostram adequados para testes rápidos de triagem ou para o controle de qualidade rotineiro, evidenciando a necessidade do desenvolvimento de abordagens analíticas mais eficientes (CHENG et al., 2010; LU et al., 2009; POONIA et al., 2017).

3.4. Espectroscopia no Infravermelho Próximo (NIR)

3.4.1 Fundamentação da técnica espectroscópica

A espectroscopia no infravermelho próximo (NIR) é uma técnica analítica vibracional que opera na região do espectro eletromagnético normalmente situada entre 750 e 2500 nm (PASQUINI, 2018). Diferentemente da região do infravermelho médio, onde ocorrem as transições vibracionais fundamentais de forte intensidade, a interação da radiação eletromagnética na região do NIR baseia-se predominantemente na excitação de sobretons e bandas de combinação desses modos vibracionais (PASQUINI, 2018).

Para que ocorra a absorção da energia NIR, é necessário que ocorra variação no momento dipolar da molécula analisada. Como a intensidade do sinal está atrelada à anarmonicidade dessas vibrações, a técnica capta com mais eficiência ligações envolvendo hidrogênio e átomos de maior massa (como C, N, O e S) (PASQUINI, 2018). Consequentemente, as ligações C-H, N-H e O-H são as responsáveis pelas bandas de absorção de sobretons mais relevantes e intensas na região do espectro NIR (INGLE *et al.*, 2016). Essa resposta a compostos orgânicos torna o NIR uma ferramenta ideal para extrair informações composicionais de amostras complexas, viabilizando a detecção de proteínas e de compostos adulterantes ricos em nitrogênio de forma simultânea (LUKACS *et al.*, 2018; PASQUINI, 2018).

Diferentemente do infravermelho médio, as bandas de sobreton do NIR apresentam menor intensidade de absorção, o que confere à radiação uma maior capacidade de penetrar na matéria. Na prática, essa característica permite que matrizes sólidas e em pó sejam analisadas em seu estado bruto, dispensando etapas destrutivas de preparo de amostra (INGLE *et al.*, 2016; PASQUINI, 2018). Somado a essa facilidade operacional, o desenvolvimento tecnológico recente permitiu uma certa miniaturização dos instrumentos. Hoje, o mercado dispõe de espectrofotômetros portáteis e com custos reduzidos, viabilizando que as análises saiam do laboratório e sejam executadas diretamente no campo ou na linha de produção (PASQUINI, 2018).

A associação da técnica NIR com equipamentos portáteis consolida uma abordagem alinhada aos princípios da química verde, uma vez que o método elimina a necessidade de solventes e reagentes químicos nocivos, apresentando também menor complexidade e gasto energético comparado aos métodos de referência. Além da redução de custos operacionais e de impacto ambiental, a técnica destaca-se pela sua rapidez: a obtenção do espectro e do resultado pode ser concluída em menos de um minuto por amostra, superando as limitações de tempo dos métodos oficiais tradicionais (INGLE *et al.*, 2016).

3.4.2 Instrumentação Portátil

A recente miniaturização dos espectrômetros NIR revolucionou a análise química, transferindo as medições do ambiente laboratorial técnico para o local de interesse de forma rápida e não destrutiva (FRIEDRICH *et al.*, 2014). Diferentemente dos equipamentos de bancada tradicionais, a instrumentação portátil moderna, como a família de espectrômetros MicroNIR – equipamento usado neste trabalho – baseia-se na tecnologia de Filtro Variável Linear (LVF – Linear Variable Filter). O LVF é um elemento dispersivo de filme fino cuja espessura varia fisicamente em formato de cunha; essa característica geométrica faz com que o comprimento de onda transmitido mude de acordo com a posição espacial da radiação sobre o filtro (FRIEDRICH *et al.*, 2014).

Ao acoplar o componente LVF diretamente a um arranjo de detectores de InGaAs não refrigerados, elimina-se completamente a necessidade de peças móveis no interior do equipamento. Essa inovação de design óptico resulta em um instrumento miniaturizado e de robustez física, capaz de suportar variações de temperatura, umidade e choques mecânicos. Além disso, a alta eficiência energética desse arranjo permite que o sistema e suas fontes duplas de iluminação de tungstênio operem alimentados apenas por uma conexão USB convencional, viabilizando análises de controle de qualidade e a identificação de fraudes em tempo real diretamente no campo (FRIEDRICH *et al.*, 2014).

3.5. Quimiometria

A análise de matrizes alimentares complexas, como suplementos proteicos, por técnicas espectroscópicas vibracionais gera um volume expressivo de dados cujos perfis são

frequentemente caracterizados por bandas largas e altamente sobrepostas (PASQUINI, 2018). Para superar esse obstáculo, o emprego de ferramentas quimiométricas e de calibração multivariada é uma boa estratégia, pois permite extrair a informação química analítica oculta na totalidade do espectro (FORINA *et al.*, 2007; JIAO *et al.*, 2019).

Antes da construção dos modelos quimiométricos, os dados espectrais brutos comumente requerem pré-processamentos para minimizar interferências que não agregam informação química relevante, como o espalhamento de luz e as flutuações de linha de base. Técnicas como a Variável Normal Padrão (SNV) e a Correção Multiplicativa de Sinal (MSC) são comumente aplicadas para corrigir as variações de intensidade causadas por diferenças no espalhamento de luz oriundas da granulometria de amostras em pó (PASQUINI, 2018). Além disso, algoritmos de suavização e de derivadas matemáticas (primeira e segunda derivadas), também são usados para atenuar ruídos sistemáticos, corrigir desvios de linha de base e, sobretudo, aumentar a resolução de picos sobrepostos (LU *et al.*, 2009; RODRIGUES JÚNIOR *et al.*, 2016).

Após o tratamento dos espectros, a etapa de quantificação de adulterantes exige um algoritmo robusto, sendo a Regressão de Mínimos Quadrados Parciais (PLS – Partial Least Squares) uma das abordagens mais fundamentais da quimiometria (GELADI; KOWALSKI, 1986). Diferente de métodos de regressão linear múltipla, o PLS opera relacionando simultaneamente a matriz de dados instrumentais (os espectros, comumente chamada de matriz **X**) com a matriz de propriedades de interesse (as concentrações dos analitos, comumente chamada de matriz **Y**). O funcionamento do modelo consiste em realizar a compressão dos dados originais, projetando-os em um novo espaço de dimensão reduzida composto por variáveis latentes (VLs) (GELADI; KOWALSKI, 1986; FORINA *et al.*, 2007).

Essas VLs são calculadas sequencialmente com o objetivo específico de maximizar a covariância entre os dados espectrais e a concentração da substância alvo (GELADI; KOWALSKI, 1986). Dessa forma, o algoritmo PLS lida de maneira eficaz com matrizes de dados onde o número de variáveis (comprimentos de onda) excede o número de amostras (FORINA *et al.*, 2007), viabilizando a quantificação precisa de adulterantes ricos em nitrogênio em matrizes complexas como a do *whey protein* (JIAO *et al.*, 2019; LUKACS *et al.*, 2018).

3.6. Validação analítica e figuras de mérito

A avaliação do desempenho de um modelo de calibração multivariada, como o PLS, requer uma análise de sua capacidade preditiva e de sua confiabilidade analítica. Tradicionalmente, a qualidade do ajuste e a capacidade de previsão do modelo são avaliadas por meio de parâmetros globais de erro, como a Raiz do Erro Quadrático Médio de Calibração (RMSEC), a Raiz do Erro Quadrático Médio de Validação Cruzada (RMSECV) e a Raiz do Erro Quadrático Médio de Previsão (RMSEP) (JIAO et al., 2019; SILVA et al., 2024). A equação fundamental para essas figuras de mérito é dada por:

$$RMSE = \sqrt{\frac{\sum(\hat{y}_p - y_r)^2}{n}}$$

Em que $\hat{y}_p$ é o valor predito pelo modelo, y_r é o valor de referência e n é o número de amostras. Quando essa equação é aplicada às amostras estimadas durante o processo de validação cruzada, obtém-se o RMSECV. Por outro lado, quando aplicada exclusivamente ao conjunto de validação externa, obtém-se o RMSEP, que reflete o erro de predição do modelo para amostras inéditas.

Embora essas métricas sejam fundamentais para quantificar a dispersão dos dados e a exatidão geral do modelo, a validação completa de um método analítico — especialmente quando aplicado à detecção de fraudes e adulterantes alimentares — exige a determinação de seus limites de detecção (CHENG et al., 2010; LU et al., 2009; LUKACS et al., 2018).

Para estabelecer esses limites com rigor estatístico, faz-se o uso de diretrizes internacionais, como a norma ISO 11843-2, que especifica metodologias para a determinação da capacidade de detecção em casos de calibração linear (ISO, 2000). No contexto da avaliação de contaminantes e resíduos, a terminologia estatística descrita na referida norma traduz-se de forma prática nos limites conhecidos como Limite de Decisão ($CC\alpha$) e Capacidade de Detecção ($CC\beta$) (ISO, 2000).

Segundo os preceitos da norma ISO 11843-2 (ISO, 2000), a avaliação da capacidade de detecção deve considerar as probabilidades de ocorrência de falsos positivos e falsos negativos. Dessa forma, definem-se os seguintes parâmetros teóricos para a validação do modelo (VAN LOCO et al., 2007):

Limite de Decisão (CC α): Na norma ISO 11843-2, esse parâmetro é denominado como Valor Crítico. Trata-se do limite a partir do qual se pode afirmar que uma amostra não se encontra no seu estado básico (ou seja, difere de uma amostra em branco, isenta de melamina) (ISO, 2000). O CC α é calculado admitindo-se uma probabilidade de erro α (geralmente fixada em 5%) de se detectar erroneamente a presença do analito quando esse está ausente (ISO, 2000). Em termos práticos, é a menor concentração na qual o método acusa um resultado positivo com confiabilidade estatística estabelecida e é matematicamente expresso por:

$$cc_{\alpha} = \frac{(y_c - b)}{a}$$

Onde:

$$y_c = ax_{MRL} + b + s_{res}t_{\alpha;n-2} \sqrt{1 + \frac{1}{n} + \frac{(x_{MRL} - \bar{x})^2}{\sum(x - \bar{x})^2}}$$

Em que:

b : É o coeficiente linear da reta de calibração construída pelos valores de concentração nominal de melamina em função dos valores preditos pelo modelo PLS.

a : É o coeficiente angular da reta de calibração construída pelos valores de concentração nominal de melamina em função dos valores preditos pelo modelo PLS.

$t_{\alpha;n-2}$: É o valor crítico extraído da distribuição t de Student uni-caudal. O α representa o nível de significância adotado para o erro de falso positivo (usualmente definido em 0,05 ou 5%). O $n - 2$ representa os graus de liberdade associados a curvas de calibração univariadas, para modelos multivariados ele geralmente é calculado como o número de amostras usadas na calibração menos o número de variáveis latentes menos 1 grau de liberdade perdido ao centrar os dados na média.

s_{res} : Corresponde ao desvio padrão residual da regressão.

n : É o número de amostras usadas para construir o modelo.

x_{MRL} : Trata-se da concentração máxima de uma determinada substância legalmente permitida em uma matriz. Para substâncias proibidas considera-se o valor como zero.

$\bar{x}$: Corresponde à média aritmética de todas as concentrações (níveis de melamina) das amostras que compõem o conjunto de calibração.

x : Representa a concentração nominal individual de cada amostra utilizada na construção do modelo de calibração.

Capacidade de Detecção (CC β): Referenciado na norma como Valor Mínimo Detectável, esse limite representa a menor concentração verdadeira do analito na amostra que o método analítico é capaz de detectar de forma consistente (ISO, 2000). O CC β é determinado assumindo-se uma probabilidade de erro β (falso negativo) de não se detectar a substância quando a sua concentração real for igual a esse valor mínimo detectável (ISO, 2000). É matematicamente expresso por:

$$CC_{\beta} = \frac{\varphi(\alpha, \beta, \nu) s_{res}}{a} \sqrt{1 + \frac{1}{n} + \frac{(CC_{\beta} - \bar{x})^2}{\sum (x - \bar{x})^2}}$$

Em que:

a : É o coeficiente angular da reta de calibração construída pelos valores de concentração nominal de melamina em função dos valores preditos pelo modelo PLS.

n : É o número de amostras usadas para construir o modelo.

$\bar{x}$: Corresponde à média aritmética de todas as concentrações (níveis de melamina) das amostras que compõem o conjunto de calibração.

x : Representa a concentração nominal individual de cada amostra utilizada na construção do modelo de calibração.

s_{res} : Corresponde ao desvio padrão residual da regressão.

$\varphi(\alpha, \beta, \nu)$: Representa o parâmetro de não-centralidade. Como os níveis de significância adotados para falsos positivos e falsos negativos foram iguais ($\alpha = \beta = 0,05$), esse termo foi aproximado neste estudo pelo dobro do valor crítico de t de Student uni-caudal ($2t$), representando simultaneamente o limite de 5% para os erros do Tipo I e do Tipo II.

A determinação de ambos os parâmetros (CC α e CC β) atesta a sensibilidade da modelagem PLS aliada à espectroscopia NIR, garantindo não apenas que o erro de predição quantitativa seja baixo — demonstrado pelos valores dos RMSE —, mas assegurando estatisticamente os limiares de concentração nos quais o modelo discrimina adulterações por melamina em Whey Protein com segurança e respaldo normativo (ISO, 2000).

Além da determinação dos limites de detecção, a validação de um método analítico exige a comprovação de sua exatidão e precisão. A exatidão é avaliada numericamente por meio da taxa de recuperação (R%), que consiste em determinar o grau

de concordância entre o valor predito pelo modelo quimiométrico e a concentração nominal do analito conhecida na amostra (INMETRO, 2020). Matematicamente, a recuperação é expressa por:

$$R\% = \left(\frac{C_{predita}}{C_{nominal}} \right) \times 100$$

Onde $C_{predita}$ representa a concentração do analito estimada pelo modelo PLS e $C_{nominal}$ corresponde à concentração real calculada no preparo da amostra.

Por outro lado, a precisão expressa a variabilidade matemática dos dados e é comumente calculada pelo Desvio Padrão Relativo, de acordo com a equação a seguir:

$$DPR\% = \left(\frac{S}{\bar{x}} \right) \times 100$$

Onde S é o desvio padrão amostral das determinações analíticas e $\bar{x}$ representa a média aritmética das concentrações preditas para um mesmo nível de contaminação.

Para refletir a realidade operacional de um laboratório, a precisão costuma ser dividida e testada sob diferentes óticas: a repetibilidade, que mensura a dispersão em análises sucessivas em um breve espaço de tempo sob as exatas mesmas condições; e a precisão intermediária, que avalia o comportamento do método no mesmo laboratório durante um prazo maior, introduzindo mudanças nas condições, como a troca de operadores ou dias de análise (INMETRO, 2020).

4 PARTE EXPERIMENTAL

4.1. Materiais, Reagentes e Equipamentos

Para a simulação das fraudes, utilizou-se como padrão o adulterante melamina (padrão analítico - PA) apresentando pureza superior a 98%. As matrizes foram 7 amostras comerciais autênticas de suplemento alimentar do tipo whey protein, contemplando diferentes marcas e os sabores Natural, Morango, Cappuccino, Cookies 'n Cream, Chocolate e Baunilha. O preparo das amostras envolveu o uso de balança analítica para a pesagem dos pós, tubos do tipo Eppendorf para o processo de mistura e homogeneização manual, e frascos de vidro transparentes (vials) com capacidade de 2 mL para o acondicionamento final e leitura espectral

A aquisição dos espectros no infravermelho próximo (NIR) foi conduzida utilizando um espectrofotômetro portátil modelo MicroNIR 1700 (Viavi Solutions). Por fim, o tratamento dos dados instrumentais e a construção dos modelos de calibração multivariada foram realizados em ambiente computacional, utilizando o software MATLAB (versão 2022a) operando em conjunto com o pacote PLS_Toolbox (Eigenvector Research, Inc.).

4.2. Preparo das amostras

Para a simulação do processo de fraude e posterior construção dos modelos quimiométricos, amostras comerciais de *whey protein* de diferentes marcas e diferentes sabores foram intencionalmente adulteradas com o padrão de melamina PA (padrão analítico) com pureza superior a 98%. A inclusão de sabores e marcas distintas teve como propósito incorporar a variabilidade da matriz alimentícia ao conjunto de dados, englobando as diferenças naturais de constituintes como corantes e aromatizantes presentes em cada produto. A relação completa das 7 amostras autênticas selecionadas para o estudo, contemplando suas respectivas marcas, sabores e códigos de identificação, encontra-se detalhada na Tabela 1.

Código	Sabor
WP0	Natural
WP1	Morango
WP2	Cappuccino
WP3	Cookies 'n Cream
WP4	Chocolate
WP5	Baunilha
WP6	Chocolate

Tabela 1: Códigos e sabores das amostras de *whey protein* utilizadas

Essa estratégia buscou garantir a construção de um modelo de calibração mais abrangente e com alta capacidade de generalização, permitindo que a quantificação do adulterante ocorresse de forma independente da marca ou do sabor do suplemento analisado. Para tal, em cada amostra autêntica de *whey protein*, foi feita uma curva contendo 8 níveis crescentes de incremento de adulterante. As concentrações finais de adulterante preparadas foram de 0, 5, 10, 15, 20, 25, 30 e 35% (m/m) de melamina na matriz proteica. A escolha dessa faixa de adulteração justifica-se pelo foco do estudo em simular e detectar a fraude com fins econômicos, distinguindo-a de contaminações residuais ou acidentais. A adulteração intencional visa mascarar o teor proteico e enganar metodologias clássicas de controle de qualidade. Para que essa prática ilícita proporcione vantagem financeira real aos fraudadores, exige-se a adição de quantidades macroscópicas do composto nitrogenado à matriz do produto. Dessa forma, o limite superior de 35% m/m foi selecionado para abranger cenários extremos de falsificação em larga escala.

As massas adequadas de *whey protein* e do adulterante foram pesadas em balança analítica e transferidas para Eppendorfs. Visando simular as condições reais da fraude no estado sólido, as misturas foram homogeneizadas vigorosamente de forma manual. Esse procedimento foi adotado para garantir a uniformidade na distribuição das partículas do contaminante em meio à matriz de *whey* antes da análise espectroscópica. Após a completa homogeneização, as amostras em pó foram transferidas e acondicionadas em frascos de vidro (vials) com capacidade de 2mL, ficando prontas para a etapa de aquisição dos espectros.



Figura 4: Preparo de amostras. **Fonte:** O autor (2026)

4.3. Aquisição dos Espectros no Infravermelho Próximo (NIR)

Após a etapa de preparo, as amostras foram submetidas à análise instrumental para a obtenção dos dados espectrais. A aquisição dos espectros na região do infravermelho próximo (NIR) foi realizada utilizando um espectrômetro portátil modelo MicroNIR 1700 da Viavi Solutions.



Figura 5: Espectrofotômetro portátil MicroNIR® 1700 da Viavi Solutions.

Fonte: FRIEDRICH, 2014

As leituras foram feitas diretamente nos frascos de vidro (vials) de 2 mL contendo as misturas no estado sólido. Os espectros foram coletados na faixa espectral de 900 a 1700 nm.

O espectro final gerado para cada ponto da curva resultou da média de 20 varreduras sucessivas (scans). Esse procedimento foi realizado sistematicamente para todos os níveis de adulteração (0 a 35% m/m), gerando o conjunto de dados brutos que foi posteriormente exportado para o tratamento quimiométrico.

4.4. Tratamento de Dados e Análise Quimiométrica

Os dados espectrais brutos obtidos no equipamento NIR foram exportados e submetidos ao tratamento quimiométrico utilizando o software MATLAB versão 2022a, operando em conjunto com o pacote PLS_Toolbox (Eigenvector Research, Inc.).

Para a construção e avaliação do modelo, o conjunto total de espectros foi dividido em dois subconjuntos distintos: um conjunto de calibração (destinado ao treinamento do modelo) e um conjunto de validação externa (destinado a testar a capacidade preditiva do modelo em amostras não vistas). Essa separação foi realizada adotando-se uma proporção de 3:1, ou seja, a cada três amostras consecutivas direcionadas para compor o conjunto de calibração, a amostra seguinte foi separada para compor o conjunto de validação.

Previamente à construção dos modelos de regressão, a Análise de Componentes Principais (PCA) foi utilizada de forma exploratória. Essa etapa teve como objetivo avaliar a distribuição natural das amostras no espaço amostral e verificar o comportamento de agrupamento em função das diferentes marcas e sabores de whey protein utilizados no estudo.

Com o objetivo de minimizar interferências, que não contêm informações químicas relevantes, como o espalhamento de luz causado pela diferença na granulometria dos pós e as variações de linha de base, os espectros brutos foram submetidos a testes de pré-processamento. Para avaliar a melhor correção espectral, foram testadas e comparadas três abordagens matemáticas distintas antes da construção dos modelos: (I) Correção Multiplicativa de Sinal (MSC) seguida de centrar na média; (II) Variável Normal Padrão (SNV) seguida de centrar na média; e (III) Derivadas matemáticas seguidas de centrar na média.

Após os pré-processamentos, os modelos de calibração foram construídos utilizando o algoritmo de Regressão de Mínimos Quadrados Parciais (PLS). O número ideal de VLs foi determinado por meio de validação cruzada. Essa validação foi conduzida utilizando o método de validação cruzada por subconjuntos aleatórios (*random subsets*). Para garantir a máxima robustez estatística do modelo e mitigar qualquer viés decorrente do particionamento das amostras, o conjunto de dados foi dividido em 10 subconjuntos (*data splits*). O procedimento foi submetido a 20 iterações independentes, nas quais as amostras foram aleatoriamente reagrupadas. O erro médio resultante de todas as simulações foi expresso como o RMSECV, parâmetro esse utilizado para determinar o número ótimo de VLs.

Seguidamente, o desempenho preditivo e a robustez do modelo quimiométrico foram avaliados por meio do cálculo de suas Figuras de mérito. A qualidade dos ajustes foi

comprovada através das métricas do RMSEC e do RMSECV. A real capacidade preditiva do modelo foi atestada aplicando-o ao conjunto de testes externo, o que permitiu o cálculo do RMSEP.

Além da avaliação da predição quantitativa global, a confiabilidade analítica do modelo para a identificação de fraudes em baixas concentrações foi investigada por meio da avaliação de sua capacidade de detecção. Para isso, os parâmetros de CC α e Capacidade CC β foram calculados em consonância com as diretrizes fundamentais da norma internacional ISO 11843-2.

Para atestar a exatidão e a precisão do método preditivo, amostras de whey protein foram contaminadas com melamina em três níveis distintos de concentração (5%, 20% e 35% m/m). Para a avaliação da repetibilidade, procedeu-se com a leitura espectral em sextuplicata para cada nível de concentração. Para a precisão intermediária, o mesmo ensaio foi repetido em um segundo dia, totalizando 12 amostras por nível. A exatidão foi expressa em termos percentuais de Recuperação, enquanto a precisão foi avaliada através do Desvio Padrão Relativo.

Por fim, além da avaliação estatística, os modelos PLS foram interpretados do ponto de vista químico. Para correlacionar a resposta matemática do algoritmo com as características estruturais das amostras, realizou-se a avaliação do vetor de regressão e dos escores de Importância da Variável na Projeção (VIP Scores). Essa etapa permitiu identificar os comprimentos de onda de maior relevância para a construção do modelo e atribuí-los aos modos vibracionais específicos dos grupos funcionais característicos da melamina e da matriz de *whey protein*.

5 RESULTADOS E DISCUSSÃO

5.1. Preparo e Divisão das Amostras

Para o desenvolvimento e validação do modelo quimiométrico (PLS), o conjunto total foi ordenado de forma crescente em função da concentração do adulterante e dividido na proporção de 3:1. Seguindo esse critério, 42 amostras (75%) foram alocadas para o conjunto de calibração, responsável pelo "aprendizado" do modelo, e 14 amostras (25%) foram reservadas para o conjunto de validação externa, garantindo que o modelo fosse testado com dados inéditos. A Tabela 2 apresenta a distribuição final das amostras:

Concentração melamina (% m/m)	Amostras Totais (n)	Conjunto de Calibração	Conjunto de Validação
0	7	5	2
5	7	5	2
10	7	6	1
15	7	5	2
20	7	5	2
25	7	5	2
30	7	6	1
35	7	5	2
Total	56	42	14

Tabela 2: Distribuição do conjunto de amostras de *whey protein* por nível de adulteração com melamina.

5.2. Pré-processamentos e espectros

5.2.1 Apresentação do Espectro Bruto

A Figura 6 apresenta os espectros originais brutos das amostras analisadas. Observa-se uma variação significativa no deslocamento da linha base. Em amostras sólidas em pó, como o *whey protein*, esse fenômeno ocorre devido ao espalhamento da luz, causado pela heterogeneidade no tamanho das partículas e compactação da amostra, fenômeno que pode mascarar a informação química dos constituintes da amostra.

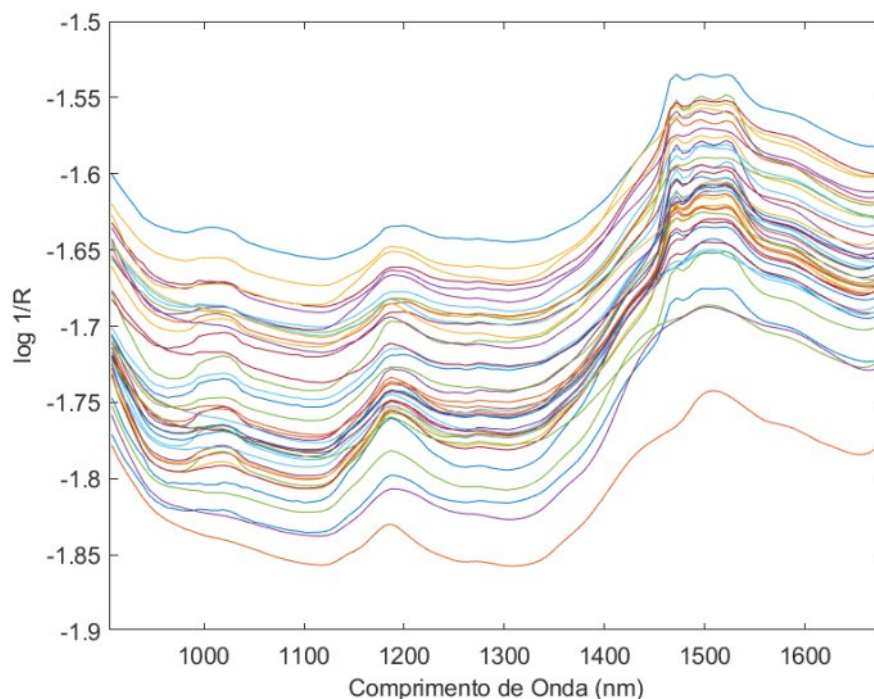


Figura 6 – Espectros NIR brutos das amostras analisadas **Fonte:** O autor

(2026)

5.2.2 Pré-processamentos

Para corrigir essas variações foram testados 3 tipos de pré-processamentos diferentes – MSC, SNV e Derivadas de Savitzky-Golay, seguidas, nos três casos, de uma centralização na média – a fim de avaliar qual ofereceria os melhores resultados.

A seguir, a Tabela 3 mostra os resultados das Figuras de Mérito dos modelos resultantes dos 3 pré-processamentos:

Pré-processamentos	VLs	R^2_{cal}	R^2_{cv}	R^2_{pred}	RMSEC	RMSECV	RMSEP
MSC	9	0,989	0,971	0,975	1,1	2,0	2,2
SNV	10	0,990	0,971	0,977	1,1	2,0	2,1
Derivada+alisamento	12	0,988	0,941	0,970	1,3	3,0	2,2

Tabela 3: Figuras de Mérito dos modelos gerados pelos três pré-processamentos testados. **Fonte:** O autor

Nota-se que o pior desempenho entre os modelos foi no pré-processado com as Derivadas de Savitzky-Golay, em que, mesmo com o número de VLs escolhidas maior, mostrou valores inferiores de R^2_{cal} , R^2_{cv} , R^2_{pred} e superiores de RMSEC, RMSECV e RMSEP.

Entre os modelos pré-processados com MSC e SNV, as Figuras de mérito de ambos se mostraram semelhantes, com pequenas variações entre si. No entanto, o MSC manteve valores próximos ao SNV com menos VLs e por isso, para este estudo, esse pré-processamento foi selecionado como o modelo de calibração ótimo.

A semelhança nos resultados de previsão obtidos entre o MSC e o SNV está em consonância com a literatura quimiométrica. De acordo com Fearn *et al.* (2009), as equações para a aplicação desses dois pré-processamentos possuem formatos semelhantes e são consideradas intercambiáveis na prática, pois ambas são desenhadas para diminuir a elevada variabilidade espectral causada por efeitos físicos de espalhamento de luz.

A Figura 7 a seguir mostra o espectro das amostras de whey e de whey + melamina analisadas após aplicação do pré-processamento MSC.

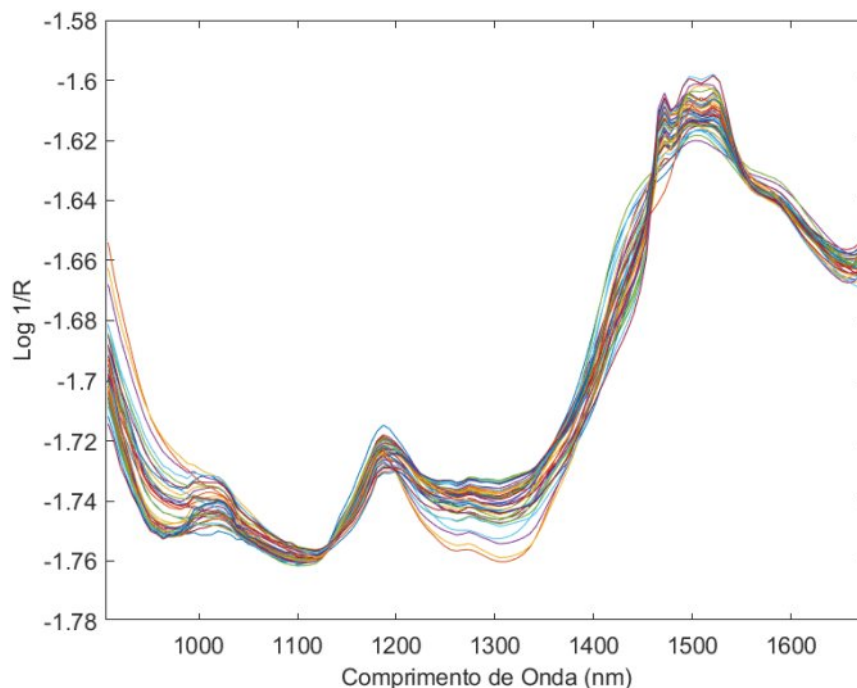


Figura 7 – Espectros NIR das amostras analisadas pré-processados com MSC

Fonte: O autor (2026)

Como pode ser observado na Figura 7, a aplicação do MSC cumpriu com eficácia o seu propósito. Houve uma redução significativa na dispersão do sinal em comparação aos espectros brutos (Figura 6). Com o alinhamento das linhas de base, as bandas de absorção características das amostras tornaram-se visualmente mais definidas, fornecendo uma matriz de dados mais limpa para a etapa de construção do modelo.

5.2.3 Análise Exploratória dos Dados (PCA)

A Análise de Componentes Principais (PCA) foi aplicada aos espectros pré-processados por MSC com o intuito de avaliar a distribuição das amostras e a variabilidade das matrizes de whey protein. A Figura 8 apresenta o gráfico de scores para as duas primeiras componentes principais, que juntas explicam 94,63% da variância total dos dados.

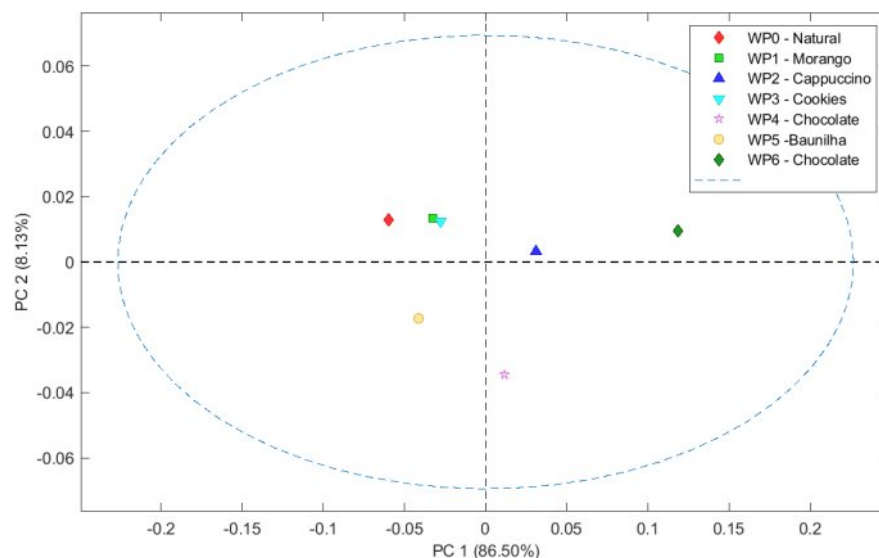


Figura 8: Gráfico de scores das duas primeiras componentes principais (PC1 vs. PC2) para os espectros NIR das amostras autênticas de whey protein pré-processados por MSC. A variância explicada por cada componente está indicada entre parênteses nos respectivos eixos. **Fonte:** O autor (2026)

A primeira componente principal (PC1), responsável pela maior parcela da variância (86,5%), promoveu uma separação das amostras baseada nas características de cor e composição dos ingredientes de sabor. Observou-se que as amostras de coloração clara (Sabores Natural, Morango, Cookies 'n Cream e Baunilha) agruparam-se na região negativa (esquerda) do eixo. Em contrapartida, as amostras de coloração escura, cuja formulação contém adição de ingredientes como cacau em pó (sabores Chocolate e Cappuccino), situaram-se na região positiva (direita). Essa separação indica que a presença desses constituintes altera de forma expressiva o perfil espectral resultante. A Figura 9 apresenta o gráfico de Loadings da PC1:

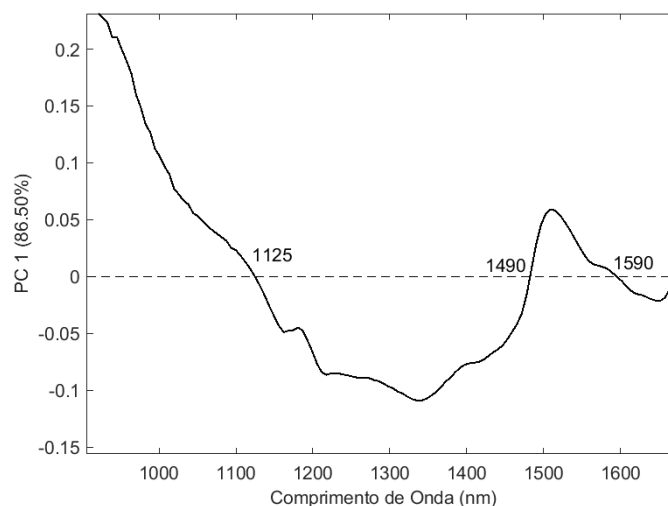


Figura 9: Gráfico de loadings da primeira componente principal (PC1). As regiões positivas (900 a 1125 nm e entre 1490 e 1590 nm) indicam a contribuição espectral característica das amostras escuras. A banda negativa (1125 a 1490 nm) evidencia os comprimentos de onda que caracterizam a matriz proteica base das amostras claras. **Fonte:** O autor (2026)

A análise do gráfico de loadings (pesos) da Componente Principal 1 permite interpretar as variáveis espectrais que conduziram à separação observada entre os dois grupos. Os coeficientes negativos da PC1 demonstraram que a faixa espectral entre 1125 e 1490 nm atuou como o principal vetor para agrupar as amostras claras (sabor baunilha/natural) em um quadrante distinto. Quimicamente, essa região atua como a assinatura da matriz orgânica natural do whey protein. Conforme o atlas de correlação de Workman Jr. e Weyer (2012), esse intervalo engloba as bandas de combinação de estiramento das ligações C-H (1300 a 1500 nm) inerentes às cadeias laterais dos aminoácidos, além de mapear as amidas primárias em 1430 nm e o primeiro sobretom de estiramento N-H das ligações peptídicas em 1471 nm.

Em contrapartida, as áreas com loadings positivos indicam os comprimentos de onda em que as amostras escuras apresentaram maior absorção. A forte correlação das amostras de whey protein sabor chocolate com a região espectral de 900 a 1100 nm na PC1 reflete não apenas o ganho de absorbância aparente devido à coloração mais escura da matriz (efeito físico da variação da linha de base), mas também a detecção de assinaturas químicas intrínsecas ao cacau adicionado como flavorizante. A literatura especializada em espectroscopia NIR atesta que essa janela abriga as bandas da matriz achocolatada. De acordo com Forte et al. (2022), é possível rastrear a acidez titulável do cacau a partir da

resposta espectral entre 930 e 950 nm, uma vez que esse intervalo captura as vibrações correspondentes à segunda região de combinação dos ácidos carboxílicos contidos no cacau, como ácido acético e ácido cítrico. Somado a isso, o comprimento de onda de 1094 nm desempenha uma influência na modelagem e rastreamento do índice de fermentação dos grãos de cacau (CAPORASO et al., 2018). Portanto, a PCA isolou bem as amostras sabor chocolate na PC1 ao capturar a introdução dos compostos ácidos e dos subprodutos de fermentação exclusivos da matriz do cacau, ausentes na proteína láctea de sabor natural.

Além disso, o modelo PCA atribuiu um forte peso positivo à região compreendida entre 1490 e 1590 nm. A coordenada de 1492 nm foi identificada na literatura como uma região relacionada tanto ao índice de fermentação quanto aos polifenóis totais em sementes de cacau (CAPORASO et al., 2018). Entretanto, a região próxima a 1490 nm não se restringe aos polifenóis; ela responde ativamente às vibrações de primeiro sobretom das ligações hidroxílicas (O-H), tornando-se um importante marcador espectral para a detecção de estruturas ricas em celulose (CAPORASO et al., 2018; WORKMAN JR.; WEYER, 2012). Assim, a relevância positiva desse intervalo no modelo corrobora a detecção analítica não apenas das estruturas fenólicas do cacau, mas também do complexo fibroso e de carboidratos característicos do cacau em pó adicionado à formulação.

Complementarmente, a PC2 (8,13%) revelou uma tendência de agrupamento relacionada à identidade do fabricante. As amostras WP4 e WP5 (pertencentes ao mesmo fabricante) apresentaram valores negativos (deslocando-se para a parte inferior do gráfico), o que as diferenciou das amostras de outros fabricantes, que se concentraram na parte superior (valores positivos na PC2). A Figura 10 mostra o gráfico de Loadings da PC2:

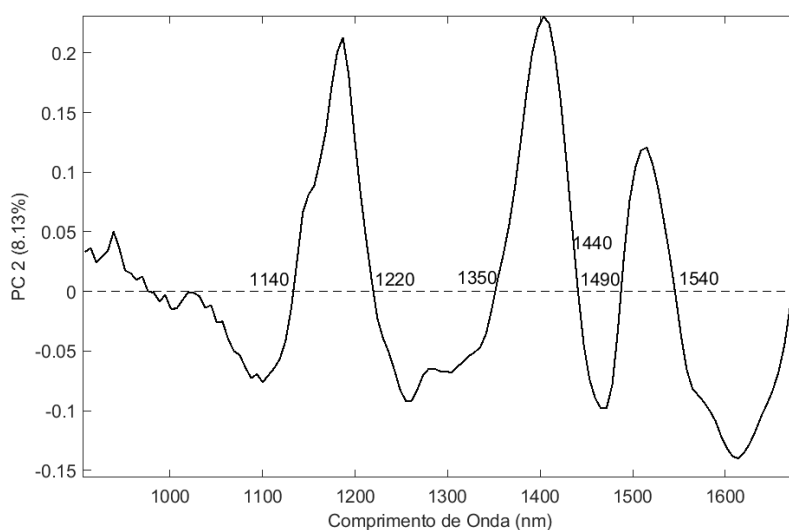


Figura 10: Gráfico de loadings da segunda componente principal (PC2). As regiões positivas (1140 a 1220 nm; entre 1350 e 1440 nm; e entre 1490 a 1540) indicam a

contribuição espectral característica das amostras da região positiva da PC2. As bandas negativas evidenciam os comprimentos de onda que caracterizam as amostras WP4 e WP5 produzidas pelo mesmo fabricante. **Fonte:** O autor (2026)

A separação das amostras WP4 e WP5 no quadrante negativo da Segunda Componente Principal (PC2) pode ser justificada pelas diferenças no balanço de macronutrientes de sua formulação base. Essas amostras caracterizam-se por apresentar um menor teor proteico e uma maior proporção de carboidratos em relação às demais marcas avaliadas. A discriminação desse grupo pela PC2 pode ser explicada pela correlação negativa observada entre proteínas e carboidratos na espectroscopia NIR. Quimicamente, essa separação reflete uma competição direta no espectro. Segundo Workman Jr. e Weyer (2012), as bandas de absorção das proteínas (originadas por vibrações das ligações N-H) localizam-se adjacentes à banda de combinação O-H dos carboidratos. A literatura estabelece que, devido a essa proximidade, o equipamento avalia as proporções desses macronutrientes a partir do perfil contínuo da região espectral, não de forma isolada. Na própria matriz do produto, a adição de maiores frações de carboidrato resulta, obrigatoriamente, em uma redução proporcional na fração proteica total. Burns & Ciurczak (2008) corroboram esse comportamento apontando uma alta intercorrelação negativa na matriz, de modo que o ganho de intensidade nas bandas de carboidrato é acompanhado por uma queda correspondente nas bandas de proteína. Dessa forma, os gráficos de correlação desses compostos costumam exibir um comportamento inversamente proporcional, assemelhando-se a um reflexo invertido.

A análise conjunta das componentes principais e de seus respectivos Loadings demonstrou que a técnica NIR foi sensível o suficiente para captar variações sutis na formulação base entre diferentes marcas comerciais. Consequentemente, a análise exploratória, ao atestar a diversidade do conjunto de amostras, evidencia que o modelo de calibração construído a seguir mostrou eficácia mesmo frente a essas variações de mercado.

5.3 Modelagem quimiométrica (PLS)

Antes de avaliar a capacidade preditiva do modelo PLS, o número ótimo de Variáveis Latentes (VLs) foi determinado utilizando o método de validação cruzada. A Figura 11 apresenta a variação do erro (RMSECV) em função do número de VLs incluídas no modelo.

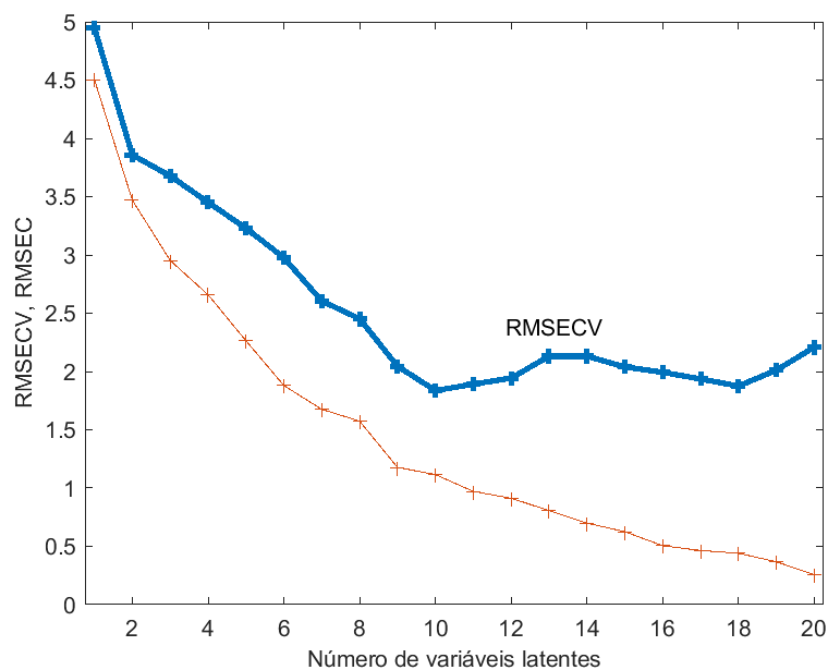


Figura 11: Variação da Raiz do Erro Quadrático Médio de Validação Cruzada (RMSECV) em função do número de variáveis latentes (VLs) para o modelo PLS com pré-processamento MSC. **Fonte:** O autor (2026)

Estabelecido o número de VLs, procedeu-se com a detecção de possíveis outliers no conjunto de dados. A Figura 12 apresenta o gráfico dos Resíduos Q e o T^2 de Hotelling para o modelo.

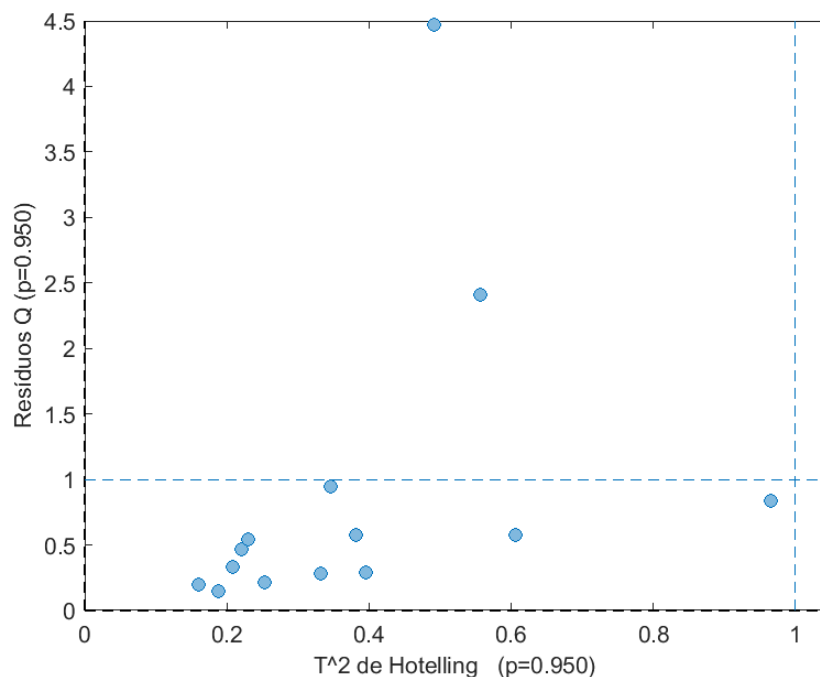


Figura 12: Gráfico de influência mostrando os Resíduos Q em função do T² de Hotelling para as amostras analisadas. As linhas tracejadas indicam os limites de confiança estatística de 95%. **Fonte:** O autor (2026)

Observa-se que todas as amostras se mantiveram dentro do limite de confiança de 95% para o T² de Hotelling. Notou-se, contudo, que algumas amostras ultrapassaram o limite estatístico no eixo dos Resíduos Q, indicando parcelas de variância espectral não capturadas pelas 9 VLs, possivelmente associadas a variações intrínsecas da matriz de diferentes sabores. Como essas amostras apresentaram baixa influência no modelo (baixo T²) e a capacidade preditiva global (RMSEP) se manteve satisfatória, optou-se por mantê-las no conjunto de calibração, preservando assim a variabilidade natural esperada no controle de rotina de formulações comerciais.

Após a definição dos parâmetros ótimos, a Figura 13 apresenta o gráfico de dispersão entre os valores de concentração (%m/m) de melamina medidos (valores de referência) e os valores previstos pelo modelo PLS (MSC com 9 VLs).

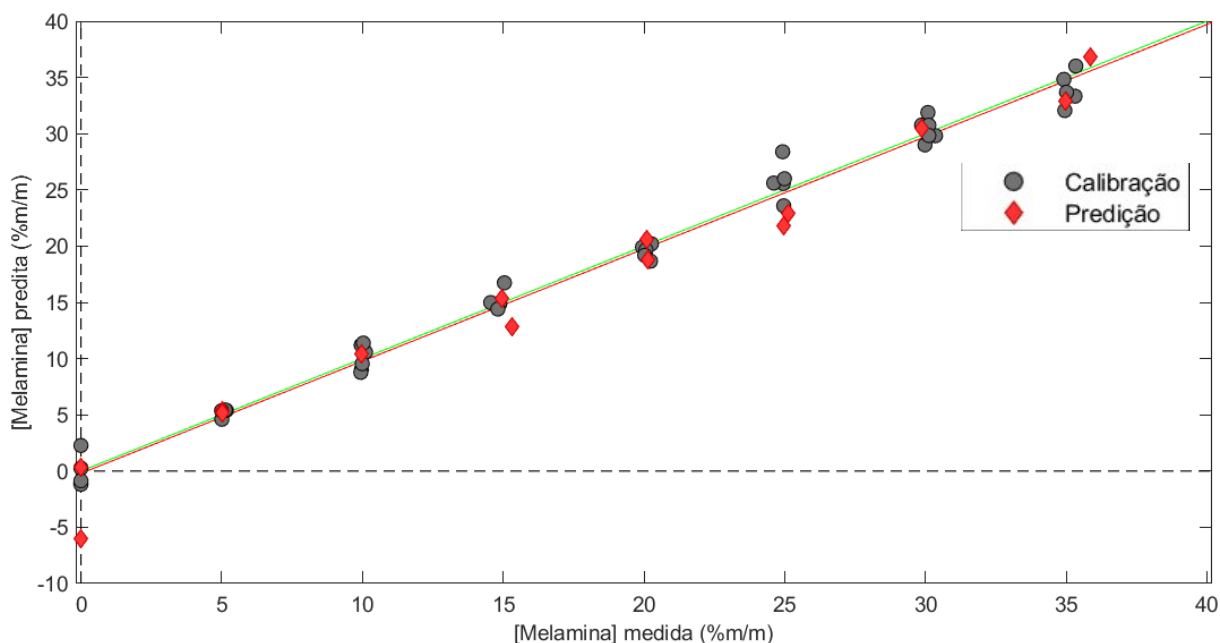


Figura 13 – Gráfico de dispersão entre os valores de concentração (%m/m) de melamina medidos e os valores previstos pelo modelo PLS (MSC com 9 VLs). Os círculos cinzas representam as amostras do conjunto de calibração e os losangos vermelhos representam as amostras do conjunto de previsão (validação externa). **Fonte:** O autor (2026)

A curva obtida apresentou um bom ajuste dos dados em torno da reta ideal, evidenciado pelo elevado coeficiente de determinação em predição ($R^2_{pred} = 0,975$). Esse resultado indica que o modelo conseguiu capturar e explicar 97,5% da variância química relacionada à adulteração por melamina nas amostras independentes, isolando com eficácia as interferências da matriz complexa do *whey protein*. Além disso, o erro médio de predição (RMSEP) obtido foi de 2,2% (m/m). A proximidade entre os valores de RMSEC (1,1% m/m), RMSECV (2,0% m/m) e RMSEP (2,2% m/m) é o principal indicativo da robustez do modelo. A ausência de grande variação entre o erro de calibração e os erros de validação indica que o modelo não apresenta indícios significativos de superajuste ou subajuste.

Por fim, a Tabela 4 apresenta os Limites de Decisão ($CC\alpha$) e Capacidade de Detecção ($CC\beta$) obtidos seguindo o estabelecido pela norma ISO 11843-2:

Parâmetro estatístico	Valor obtido (%m/m)
Limite de Decisão $CC\alpha$	2,6
Capacidade de Detecção $CC\beta$	5,0

Tabela 4: Limites estatísticos de detecção para a quantificação de melamina

Fonte: O autor (2026)

A partir desses parâmetros e aplicando as diretrizes da norma ISO 11843-2, estabeleceu-se o Limite de Decisão ($CC\alpha$) em 2,6% m/m. Esse limiar indica que qualquer leitura que resulte em uma concentração predita acima desse valor permite classificar a amostra de *whey protein* como contendo o analito, com um nível de confiança estatística de 95%.

Por sua vez, a Capacidade de Detecção ($CC\beta$) resultou em 5,0% m/m. Do ponto de vista analítico, isso significa que o método proposto é capaz de identificar, de forma consistente e segura, a presença do contaminante a partir desse teor, operando com uma probabilidade de falsos negativos (erro do tipo II) controlada em 5%.

Do ponto de vista regulamentar e de segurança alimentar, embora a Instrução Normativa nº 351 da ANVISA estipule limites extremamente restritivos para a presença residual de melamina em alimentos (2,5 mg/kg), é fundamental contextualizar que o propósito deste trabalho reside no controle da fraude econômica (adulteração intencional). A adição clandestina de compostos nitrogenados para mascarar o teor proteico real exige a introdução de quantidades maiores do adulterante na matriz do produto (geralmente acima de 5% m/m) para a obtenção de maior vantagem financeira.

5.4 Avaliação da Exatidão e Precisão

Para atestar a exatidão e a precisão do método preditivo em um cenário prático, um novo conjunto de amostras independentes foi preparado do zero. Foram formuladas amostras de *whey protein* contaminadas com melamina em três níveis distintos de concentração (5%, 20% e 35% m/m), as quais foram posteriormente projetadas no modelo PLS previamente validado. Os resultados percentuais médios e os seus respectivos Desvios Padrão Relativos estão apresentados na Tabela 5:

Concentração Nominal (% m/m)	Recuperação Média Dia 1 (%)	DPR% Dia 1 (Repetibilidade)	Recuperação Média Dia 2 (%)	DPR% Dia 2 (Repetibilidade)	Recuperação Média Global (%)	DPR% Global
5	13,0	218,3	18,9	69,2	16,0	133,3
20	89,5	5,1	81,5	6,5	85,5	7,3
35	87,7	6,7	83,9	3,3	85,8	5,6

Tabela 5: Avaliação da exatidão (Recuperação) e precisão (Repetibilidade e Precisão Intermediária) do modelo PLS para a quantificação de melamina em amostras de *whey protein*. **Fonte:** O autor (2026)

A análise dos parâmetros obtidos na Tabela 5 evidencia de forma clara as capacidades e os limites físicos da espectroscopia no infravermelho próximo para essa aplicação. Para o nível de concentração de 5% m/m, o modelo apresentou consideráveis falhas. As recuperações médias encontradas foram extremamente baixas, acompanhadas por um Desvio Padrão Relativo global muito alto. Como a região espectral do NIR atua prioritariamente sobre bandas que possuem menor absortividade molar e, conseqüentemente, geram sinais menos intensos, o ruído instrumental e o espalhamento de luz causado pela matriz complexa em pó do whey protein se sobrepõem à resposta do analito em baixas concentrações. Essa elevada dispersão e perda de exatidão em 5% corroboram o limite de Capacidade de Detecção ($CC\beta$) calculado na seção anterior, provando na prática que teores iguais ou inferiores a esse limiar não podem ser quantificados com confiabilidade por esse método.

Em contrapartida, os dados mostram que para os níveis de 20% e 35% m/m, o modelo apresentou taxas de recuperação globais satisfatórias, de 85,5% e 85,8%, respectivamente. Percebe-se certo erro negativo sistemático observado (previsão abaixo da concentração nominal). Esse viés negativo é comum na análise de sólidos por refletância difusa. Ele é fruto do aumento exponencial do espalhamento de luz causado pelos cristais de melamina em altas concentrações, o que diminui a profundidade de penetração da radiação no vial, criando um efeito de pseudo-caminho óptico (WORKMAN JR.; WEYER, 2012). Além disso, a literatura estabelece que medições por refletância difusa raramente mantêm linearidade em amplas faixas de concentração (PASQUINI, 2018). Ao nos aproximarmos de 35% de concentração no estado sólido, o sistema apresenta desvios de linearidade da Lei de Lambert-Beer devido à interação molecular e ao espalhamento. O modelo PLS, por se tratar de um método de projeção linear, tende a subestimar levemente essas não-linearidades do sistema (LU et al., 2009; PASQUINI, 2018).

Além disso, o modelo obteve boa precisão nessas faixas operacionais. Os valores de repetibilidade mantiveram-se consistentes em ambos os dias de ensaio, e a precisão intermediária alcançou um DPR de 7,3% (para a concentração de 20%) e de 5,6% (para a concentração de 35%). Embora métodos analíticos confirmatórios convencionais prevejam valores de DPR em torno de 2% para essas faixas de concentração, as diretrizes de validação estabelecem que os critérios de aceitação devem se basear na adequação ao uso pretendido pelo método (INMETRO, 2020). A literatura especializada demonstra que, para a espectroscopia NIR portátil operando por refletância difusa em matrizes sólidas em pó, os inerentes erros de amostragem e o forte espalhamento de luz expandem a variabilidade

analítica em relação às análises por via úmida (PASQUINI, 2018). Nesses cenários de triagem exploratória e análise direta, limites de aceitação com coeficientes de variação e tolerâncias de até 10% são considerados aceitáveis e atestam uma robustez adequada para o monitoramento rápido do produto (CASIAN et al., 2021).

5.5 Interpretação do Vetor de regressão e VIP SCORES

Nessa etapa, foi feita a interpretação química dos vetores informativos do modelo PLS. Na interpretação do vetor de regressão, a magnitude dos coeficientes dita a importância da variável, enquanto o seu sinal indica o tipo de correlação. Coeficientes positivos fortes indicam bandas onde o adulterante (melamina) absorve mais energia que a matriz, contribuindo positivamente para a quantificação. Em contrapartida, coeficientes fortemente negativos atuam por correlação inversa, refletindo o efeito de diluição da matriz do *whey protein* à medida que a fraude aumenta.

De forma complementar, os escores VIP (Variable Importance in Projection) avaliam a relevância de cada comprimento de onda para o modelo PLS, independentemente da direção da correlação. Diferentemente do vetor de regressão, o VIP fornece uma medida absoluta de contribuição. Na prática quimiométrica, adota-se o critério de que comprimentos de onda com valores de VIP superiores a 1 são considerados altamente significativos para a predição.

Assim, enquanto o vetor de regressão revela o comportamento químico da banda (direto ou inverso), o VIP confirma quais dessas regiões espectrais efetivamente carregam a maior carga de informação para o modelo.

A seguir é apresentado a **Figura 14** mostrando os gráficos de VIP Escores **(a)** e Vetor de Regressão **(b)**, respectivamente:

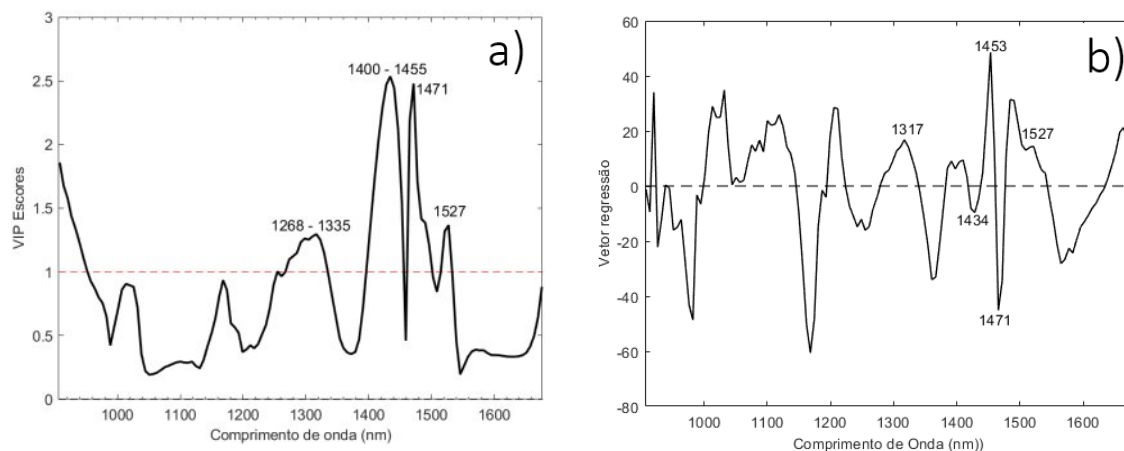


Figura 14 – Vetores informativos do modelo PLS. **a)** VIP Escores. A linha tracejada vermelha indica o valor limite 1. As variáveis acima desse limite são consideradas significativas para o modelo. **b)** Vetor de regressão. **Fonte:** O autor (2026)

A avaliação conjunta do vetor de regressão e dos VIP Escores permitiu identificar as regiões espectrais que mais contribuíram para o modelo preditivo. As variáveis com escores VIP > 1, consideradas de alta relevância para a quantificação, estão nos comprimentos de onda de 1268 a 1335, 1400 a 1455, 1471 e 1527 nm.

O pico de 1434 nm no vetor de regressão está contido na banda larga de 1400 a 1455 com forte influência no VIP escore. Uma vez que, de acordo com Workman Jr. e Weyer (2012), a região do infravermelho próximo é caracterizada por bandas intrinsecamente largas, onde um único pico observado pode ser o resultado de múltiplas vibrações moleculares, interpreta-se que no pico de 1434nm possivelmente ocorre uma coabsorção direta entre a matriz e o adulterante: a literatura estabelece que o comprimento de onda de 1430 nm pertence ao primeiro sobretom simétrico das amidas presentes na proteína, enquanto o valor de 1432 nm pertence ao primeiro sobretom assimétrico de amins aromáticas primárias, grupo funcional encontrado na melamina. O coeficiente de regressão levemente negativo obtido pelo modelo PLS nessa região indica que há um balanço entre duas forças opostas e mostra que o efeito da diluição e consequente perda de sinal das amidas do *whey protein* (contribuição negativa) se destaca frente ao ganho de absorbância provocado pela adição da melamina (contribuição positiva) nesse intervalo espectral.

Já o pico em 1453 nm no vetor de regressão – também contido na banda larga de 1400 a 1455 relevante para o VIP Escore – apresenta um coeficiente de regressão fortemente positivo, indicando uma correlação direta com a adição do adulterante. De acordo com

Workman Jr. e Weyer (2012), os comprimentos de onda de 1452,5 e 1459,5nm são atribuídos ao primeiro sobretom assimétrico do estiramento N-H em aminas aromáticas primárias. Como a melamina possui exatamente esta configuração estrutural — grupos amina primária ligados a um anel aromático nitrogenado —, é provável que esse pico tenha relação direta com o contaminante.

Seguidamente, o comprimento de onda de 1471 nm atua como o marcador da matriz natural do *whey protein*, apresentando forte coeficiente de regressão negativo. A estrutura das proteínas do soro do leite é fundamentada em ligações peptídicas, como pode ser observado na Figura 15, que consistem quimicamente em amidas secundárias. A literatura consolida que o pico em 1471 nm corresponde ao primeiro sobretom de estiramento N-H em grupos N-R de amidas proteicas (WORKMAN JR.; WEYER, 2012). O modelo PLS utiliza o decréscimo relativo da intensidade dessas bandas (correlação inversa) para mensurar o efeito de diluição da matriz verdadeira provocado pela adição da melamina.

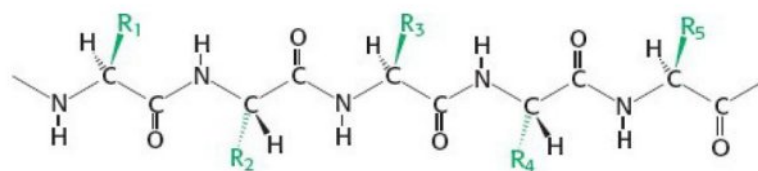


Figura 15 – Cadeia polipeptídica genérica. Nota-se a presença das amidas secundárias responsáveis pela absorção em 1471 nm descrita pela literatura. **Fonte:** BERG; TYMOCZKO; STRYER(2014)

Já a banda positiva em 1527 pode ser interpretada pelo espectro experimental da melamina pura representado na Figura 16. É possível notar que há forte absorbância entre 1460 e 1540 nm, com destaque para um pico acentuado em 1527nm. Segundo Workman Jr. e Weyer (2012) essa região é típica de grupos NH₂ e NH.

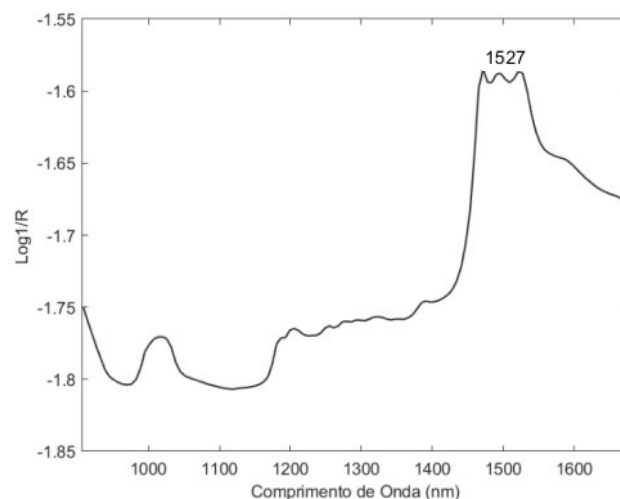


Figura 16: Espectro experimental da melamina (1,3,5-triazina-2,4,6-triamina) pura

Fonte: O autor (2026)

Por fim, a região de 1268 a 1335nm, Segundo Workman Jr. e Weyer (2012), não representa nenhuma região específica de absorção, não tendo, portanto, relevância interpretativa química para o modelo. No entanto, segundo Burns e Ciurczak (2008), na análise de sólidos em pó, o espectro NIR não reflete apenas a composição química, mas também consiste em uma assinatura do estado físico e estrutural do material, capturando fenômenos como o espalhamento de luz. Com isso, é provável que a alta relevância atribuída a essa região pelo modelo PLS não seja devido a uma absorção direta, mas um mapeamento da alteração física da linha de base provocada pelo espalhamento de luz dos cristais de melamina adicionados.

6 CONCLUSÃO

O presente trabalho demonstrou que a espectroscopia no infravermelho próximo (NIR) portátil, aliada à calibração multivariada por Regressão de Mínimos Quadrados Parciais (PLS), é uma ferramenta analítica altamente viável, rápida e não destrutiva para análises de adulterações por adição de melamina em amostras comerciais de *whey protein*.

A avaliação de diferentes metodologias de pré-processamento revelou que a Correção Multiplicativa de Sinal (MSC) apresentou o melhor desempenho para a matriz estudada, corrigindo de forma eficaz o espalhamento de luz decorrente da granulometria dos pós. A partir desse pré-tratamento, foi possível construir um modelo PLS com 9 variáveis latentes, que obteve boa capacidade preditiva, evidenciada por um coeficiente de determinação para predição de 0,975 e um RMSEP de 2,2% m/m em amostras externas.

A determinação das figuras de mérito em conformidade com as diretrizes da norma ISO 11843-2 estabeleceu um Limite de Decisão ($CC\alpha$) de 2,6% m/m e uma Capacidade de Detecção ($CC\beta$) de 5,0% m/m. Os ensaios analíticos confirmaram que o método apresenta limitações em quantificar concentrações exatas em níveis baixos (próximas a 5%), o que é esperado devido à baixa absortividade no NIR e à dispersão óptica das amostras sólidas. Contudo, o método mostrou-se estatisticamente consistente para concentrações mais altas (20% e 35% m/m). Uma vez que fraudes com fins econômicos exigem a adição de quantidades substanciais de melamina (geralmente acima de 5%) para mascarar a diluição proteica, o limite de detecção alcançado atende perfeitamente ao propósito de vigilância no mercado de suplementos.

A análise exploratória (PCA) e a interpretação dos vetores de regressão e dos escores VIP mostraram que a modelagem se baseou em marcadores espectrais quimicamente coesos. O modelo contornou variações de sabores e marcas e identificou eficientemente a fraude através do mapeamento direto das aminas primárias da melamina e do monitoramento inverso (efeito de diluição) das amidas secundárias inerentes à estrutura das proteínas do *whey*.

Por fim, conclui-se que o uso de espectrofotômetros portáteis NIR, atuando sob os preceitos da química verde (sem geração de resíduos tóxicos ou uso de solventes), é uma alternativa tecnológica promissora. Embora não substitua as técnicas cromatográficas confirmatórias na precisão absoluta para baixas concentrações, a metodologia supera as vulnerabilidades do método clássico de Kjeldahl. O modelo preditivo validado neste estudo cumpriu seu papel como ferramenta ágil de análise preliminar, possibilitando avaliações de

alto rendimento *in situ* para a rápida retirada de lotes adulterados de circulação e a proteção da saúde pública.

REFERÊNCIAS

- ANDRADE, J. et al. FTIR-ATR determination of protein content to evaluate whey protein concentrate adulteration. **LWT - Food Science and Technology**, v. 99, p. 166-172, 2019. Disponível em: <https://doi.org/10.1016/j.lwt.2018.09.079>. Acesso em: 30 maio 2026.
- AGUIRRE, J. **The Kjeldahl Method: 140 Years**. Cham: Springer Nature Switzerland, 2023.
- BERG, J. M.; TYMOCZKO, J. L.; STRYER, L. **Bioquímica**. 7. ed. Rio de Janeiro: Guanabara Koogan, 2014.
- BRASIL. Ministério da Saúde. Agência Nacional de Vigilância Sanitária. **Instrução Normativa - IN nº 351, de 18 de março de 2025**. Altera a Instrução Normativa - IN nº 160, de 1º de julho de 2022, que estabelece os limites máximos tolerados (LMT) de contaminantes em alimentos. Diário Oficial da União, Brasília, DF, 2025. Disponível em: <https://www.in.gov.br/en/web/dou/-/instrucao-normativa-in-n-351-de-18-de-marco-de-2025-618776799>. Acesso em: 1 abr. 2026.
- BURNS, D. A.; CIURCZAK, E. W. (Ed.). **Handbook of Near-Infrared Analysis**. 3. ed. Boca Raton: CRC Press, 2008.
- CAPORASO, N. et al. Hyperspectral imaging for non-destructive prediction of fermentation index, polyphenol content and antioxidant activity in single cocoa beans. **Food Chemistry**, v. 258, p. 343-351, mar. 2018. Disponível em: <https://doi.org/10.1016/j.foodchem.2018.03.039>. Acesso em: 29 maio 2026.
- CASIAN, T. et al. Testing the limits of a portable NIR spectrometer: content uniformity of complex powder mixtures followed by calibration transfer for in-line blend monitoring. **Molecules**, v. 26, n. 4, art. 1129, fev. 2021. Disponível em: <https://doi.org/10.3390/molecules26041129>. Acesso em: 28 maio 2026.
- CHENG, Y. et al. Screening melamine adulterant in milk powder with laser Raman spectrometry. **Journal of Food Composition and Analysis**, v. 23, n. 2, p. 199–202, mar. 2010. Disponível em: <https://doi.org/10.1016/j.jfca.2009.08.006>. Acesso em: 20 nov. 2025.
- ESTADOS UNIDOS. Food and Drug Administration. **Interim Safety and Risk Assessment of Melamine and its Analogues in Food for Humans**. Silver Spring: FDA, 2008. Disponível em: <https://wayback.archive-it.org/7993/20170112012209/http://www.fda.gov/Food/FoodbornenessContaminants/ChemicalContaminants/ucm164522.htm>. Acesso em: 1 abr. 2026.
- FEARN, T. et al. On the geometry of SNV and MSC. **Chemometrics and Intelligent Laboratory Systems**, v. 96, n. 1, p. 22-26, 2009. Disponível em: <https://doi.org/10.1016/j.chemolab.2008.11.006>. Acesso em: 6 maio 2026.
- FORINA, M. et al. Multivariate calibration. **Journal of Chromatography A**, v. 1158, p. 61-93, 2007. Disponível em: <https://doi.org/10.1016/j.chroma.2007.03.082>. Acesso em: 6 maio 2026.
- FORTE, M. et al. Quality evaluation of fair-trade cocoa beans from different origins using portable near-infrared spectroscopy (NIRS). **Foods**, v. 12, n. 1, art. 4, dez. 2022. Disponível em: <https://doi.org/10.3390/foods12010004>. Acesso em: 29 maio 2026.
- FRIEDRICH, D. M. et al. Miniature near-infrared spectrometer for point-of-use chemical analysis. In: PHOTONIC INSTRUMENTATION ENGINEERING, 2014. **Proceedings...**

[S.l.]: SPIE, 2014. v. 8992, art. 899203. Disponível em: <https://doi.org/10.1117/12.2040669>. Acesso em: 6 maio 2026.

G1. **Polícia apreende 4 toneladas de creatina e whey protein adulterados em distribuidora no interior de SP.** Disponível em: <https://g1.globo.com/sp/campinas-regiao/noticia/2025/09/01/policia-apreende-4-toneladas-de-creatina-e-whey-protein-adulterados-em-distribuidora-no-interior-de-sp.ghtml>. Acesso em: 20 nov. 2025.

GELADI, P.; KOWALSKI, B. R. Partial least-squares regression: a tutorial. **Analytica Chimica Acta**, v. 185, p. 1-17, 1986. Disponível em: [https://doi.org/10.1016/0003-2670\(86\)80028-9](https://doi.org/10.1016/0003-2670(86)80028-9). Acesso em: 6 maio 2026

GOOGLE. **Google Trends: interesse de pesquisa pelo termo whey.** Disponível em: <https://trends.google.com.br/trends/explore?date=all&q=WHEY&hl=pt-BR>. Acesso em: 28 maio 2026.

HARAGUCHI, F. K.; ABREU, W. C. de; PAULA, H. de. Proteínas do soro do leite: composição, propriedades nutricionais, aplicações no esporte e benefícios para a saúde humana. **Revista de Nutrição**, Campinas, v. 19, n. 4, p. 479–488, ago. 2006. Disponível em: <https://doi.org/10.1590/S1415-52732006000400007>. Acesso em: 1 abr. 2026.

INGELFINGER, J. R. Melamine and the global implications of food web adulteration. **New England Journal of Medicine**, v. 359, n. 26, p. 2745-2748, 2008. Disponível em: <https://doi.org/10.1056/NEJMp0808410>. Acesso em: 6 maio 2026.

INGLE, P. D. et al. Determination of Protein Content by NIR Spectroscopy in Protein Powder Mix Products. **Journal of AOAC International**, v. 99, n. 2, p. 360-363, 2016. Disponível em: <https://doi.org/10.5740/jaoacint.15-0115>. Acesso em: 6 maio 2026.

INMETRO. Coordenação Geral de Acreditação. **Orientação sobre validação de métodos analíticos: documento de caráter orientativo**. DOQ-CGCRE-008. Revisão 09. Duque de Caxias: INMETRO, 2020

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. ISO 11843-2: Capability of detection — Part 2: Methodology in the linear calibration case. Genebra: ISO, 2000.

JIAO, X. et al. Rapid Detection of Adulterants in Whey protein Supplement by Raman Spectroscopy Combined with Multivariate Analysis. **Molecules**, v. 24, n. 10, p. 1889, 16 maio 2019. Disponível em: <https://doi.org/10.3390/molecules24101889>. Acesso em: 20 nov. 2025

LU, C. et al. Rapid detection of melamine in milk powder by near infrared spectroscopy. **Journal of Near Infrared Spectroscopy**, v. 17, p. 59-67, 2009. Disponível em: <https://doi.org/10.1255/jnirs.829>. Acesso em: 6 maio 2026.

LUKACS, M. et al. Near infrared spectroscopy as an alternative quick method for simultaneous detection of multiple adulterants in whey protein-based sports supplement. **Food Control**, v. 94, p. 331-340, 2018. Disponível em: <https://doi.org/10.1016/j.foodcont.2018.07.004>. Acesso em: 6 maio 2026.

MECKER, L. C. et al. Selective melamine detection in multiple sample matrices with a portable Raman instrument using surface enhanced Raman spectroscopy-active gold nanoparticles. **Analytica Chimica Acta**, v. 733, p. 48–55, jul. 2012. Disponível em: <http://dx.doi.org/10.1016/j.aca.2012.05.001>. Acesso em: 20 nov. 2025.

PASQUINI, C. Near infrared spectroscopy: A mature analytical technique with new perspectives - A review. **Analytica Chimica Acta**, v. 1026, p. 8-36, 2018. Disponível em: <https://doi.org/10.1016/j.aca.2018.04.004>. Acesso em: 6 maio 2026.

- POONIA, A. et al. Detection of adulteration in milk: A review. **International Journal of Dairy Technology**, v. 70, n. 1, p. 23–42, fev. 2017. Disponível em: <https://doi.org/10.1111/1471-0307.12274>. Acesso em: 1 abr. 2026
- RODRIGUES JÚNIOR, P. H. et al. FT-Raman and chemometric tools for rapid determination of quality parameters in milk powder: Classification of samples for the presence of lactose and fraud detection by addition of maltodextrin. **Food Chemistry**, v. 196, p. 584–588, abr. 2016. Disponível em: <http://dx.doi.org/10.1016/j.foodchem.2015.09.055>. Acesso em: 1 abr. 2026.
- SILVA, I. C. M. da et al. Fast and low-cost method for direct and simultaneous determination of nitrogen and carbon in soybean leaves using benchtop and portable near-infrared devices. **J Sci Food Agric**, v. 104, p. 1843–1852, 2024. Disponível em: <https://doi.org/10.1002/jsfa.13022>. Acesso em: 6 maio 2026.
- VAN LOCO, J. et al. Calculation of the decision limit ($CC\alpha$) and the detection capability ($CC\beta$) for banned substances: The imperfect marriage between the quantitative and the qualitative criteria. **Analytica Chimica Acta**, v. 586, p. 8–12, 2007. Disponível em: <https://doi.org/10.1016/j.aca.2006.11.058>. Acesso em: 29 jun. 2026.
- WIKIPÉDIA. **Melamina**. Wikipédia: a enciclopédia livre. Disponível em: <https://pt.wikipedia.org/wiki/Melamina>. Acesso em: 6 maio 2026.
- WORKMAN JR., J.; WEYER, L. **Practical Guide and Spectral Atlas for Interpretive Near-Infrared Spectroscopy**. 2. ed. Boca Raton: CRC Press, 2012.
- YANG, F. et al. Guarding food safety with conventional and up-conversion near-infrared fluorescent sensors. **Journal of Advanced Research**, v. 41, p. 129–144, nov. 2022. Disponível em: <https://doi.org/10.1016/j.jare.2022.01.011>. Acesso em: 6 maio 2026.
- ZADOW, J. G. (Ed.). **Whey and lactose processing**. London: Elsevier Applied Science, 1992.

APÊNDICE 1

APÊNDICE 1 – Detalhamento das Concentrações de Melamina

Este apêndice apresenta os valores exatos de concentração de melamina (% m/m) para cada uma das 56 amostras de *whey protein* preparadas. Os valores referem-se às massas efetivamente pesadas na balança analítica.

ID	Nível Nominal (%)	Concentração Real (% m/m)
WP0-1	0	0
WP0-2	5	5.0049
WP0-3	10	9.9649
WP0-4	15	14.8834
WP0-5	20	19.9487
WP0-6	25	25.1164
WP0-7	30	29.9838
WP0-8	35	35.3154
WP1-1	0	0
WP1-2	5	5.0523
WP1-3	10	9.9563
WP1-4	15	14.8759
WP1-5	20	20.1502
WP1-6	25	24.9503

WP1-7	30	30.0871
WP1-8	35	35.3425
WP2-1	0	0
WP2-2	5	5.0059
WP2-3	10	10.1082
WP2-4	15	15.3182
WP2-5	20	20.2651
WP2-6	25	24.9670
WP2-7	30	30.3569
WP2-8	35	35.8686
WP3-1	0	0
WP3-2	5	5.1654
WP3-3	10	9.9577
WP3-4	15	15.0439
WP3-5	20	20.2342
WP3-6	25	24.9274
WP3-7	30	29.8823
WP3-8	35	34.9176
WP4-1	0	0

WP4-2	5	5.0059
WP4-3	10	9.9427
WP4-4	15	14.5577
WP4-5	20	20.0722
WP4-6	25	24.9572
WP4-7	30	29.8657
WP4-8	35	34.9522
WP5-1	0	0
WP5-2	5	5.1271
WP5-3	10	9.9942
WP5-4	15	14.8114
WP5-5	20	20.0849
WP5-6	25	24.6102
WP5-7	30	30.1214
WP5-8	35	35.0124
WP6-1	0	0
WP6-2	5	4.9953
WP6-3	10	10.0300
WP6-4	15	14.9651

WP6-5	20	20.0299
WP6-6	25	24.9905
WP6-7	30	30.1216
WP6-8	35	34.9740