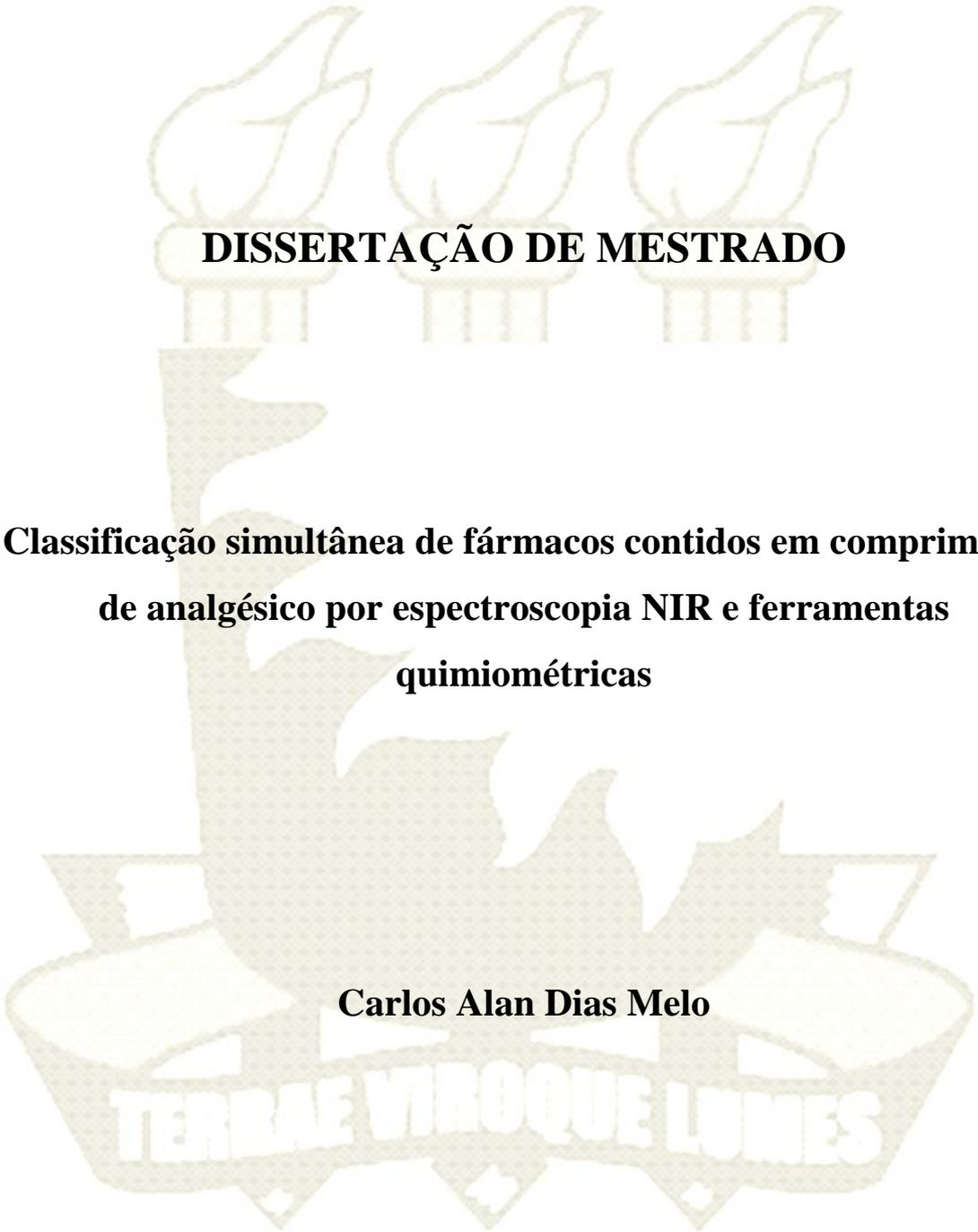


**UNIVERSIDADE ESTADUAL DA PARAÍBA**  
**CENTRO DE CIÊNCIAS BIOLÓGICAS E DA SAÚDE**  
**DEPARTAMENTO DE FARMÁCIA**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS FARMACÊUTICAS**



**DISSERTAÇÃO DE MESTRADO**

**Classificação simultânea de fármacos contidos em comprimidos  
de analgésico por espectroscopia NIR e ferramentas  
quimiométricas**

**Carlos Alan Dias Melo**

Campina Grande – PB – Brasil

Março/2013

**UNIVERSIDADE ESTADUAL DA PARAÍBA  
CENTRO DE CIÊNCIAS BIOLÓGICAS E DA SAÚDE  
DEPARTAMENTO DE FARMÁCIA  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS FARMACÊUTICAS**

## **DISSERTAÇÃO DE MESTRADO**

**Classificação simultânea de fármacos contidos em comprimidos  
de analgésico por espectroscopia NIR e ferramentas  
quimiométricas**

**Carlos Alan Dias Melo**

Dissertação apresentada ao Curso de Pós-Graduação em Ciências Farmacêuticas, nível mestrado, da Universidade Estadual da Paraíba, como parte dos requisitos para obtenção do grau de Mestre em Ciências Farmacêuticas.

**Orientadora: Ana Claudia Dantas de Medeiros**

**Coorientador: José Germano Vêras Neto**

Campina Grande – PB – Brasil

Março/2013

É expressamente proibida a comercialização deste documento, tanto na sua forma impressa como eletrônica. Sua reprodução total ou parcial é permitida exclusivamente para fins acadêmicos e científicos, desde que na reprodução figure a identificação do autor, título, instituição e ano da dissertação

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA CENTRAL – UEPB

M528c      Melo, Carlos Alan Dias.  
Classificação simultânea de fármacos contidos em comprimidos de analgésico por espectroscopia NIR e ferramentas quimiométricas. [manuscrito] / Carlos Alan Dias Melo. – 2013.  
57 f. : il. color.

Digitado

Dissertação (Mestrado em Ciências Farmacêuticas) – Universidade Estadual da Paraíba, Pró-Reitoria de Pós-Graduação e Pesquisa, 2013.

“Orientação: Prof. Dra. Ana Claudia Dantas de Medeiros, Departamento de Farmácia”.

“Co-Orientação: Prof. Dr. José Germano Vêras Neto, Departamento de Química”.

1. Espectroscopia. 2. Quimioterapia. 3. Dipirona. 4. Farmacologia. I. Título.

21. ed. CDD 615.1

Carlos Alan Dias Melo

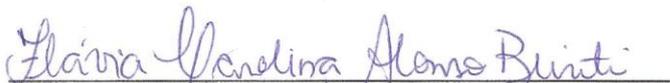
**Classificação simultânea de fármacos contidos em comprimidos  
de analgésico por espectroscopia NIR e ferramentas  
quimiométricas**

Aprovada em 11 / 04 / 13

Banca Examinadora

  
Prof.<sup>a</sup>. Dr.<sup>a</sup>. Ana Claudia Dantas de Medeiros - Orientadora

  
Prof. Dr. José Germano Vêras Neto - Coorientador

  
Prof.<sup>a</sup>. Dr.<sup>a</sup>. Flávia Carolina Alonso Buriti - Examinadora Interna

  
Prof.<sup>a</sup>. Dr.<sup>a</sup>. Marta Maria Conceição - Examinadora Externa (UFCG)

*À minha Mãe, Vanda Regina Dias Melo,  
mulher de valores indescritíveis. Dedico.*

# AGRADECIMENTOS

---

Ao nosso Deus, por tudo.

A minha esposa, Janayna, pelo amor e companheirismo.

A Francisco (Zé tribucha), pelo afeto e auxílio na minha formação.

Ao meu pai, João Barreto.

A Leonardo (Maga Véia), pelo cuidado, apreço e preocupação a mim destinados.

Aos meus irmãos, cunhadas e sobrinhos, pelo carinho.

A toda a família da minha esposa, pelo acolhimento e aceitação.

Aos meus amigos e colegas de trabalho, por me deixarem estudar quando necessário.

A minha amiga Dani, pela amizade, favores e matriculas realizadas.

A todos que compõem o LABDEM, pelo respeito.

Aos amigos do LQAQ, pelo amparo, brincadeiras e auxílios necessários.

A David (Brasil), pelos ensinamentos diversos, a Priscila e Fernando pela amizade e colaboração.

A Ana Claudia, minha orientadora, pelos ensinamentos, contribuições e, principalmente, por cair na besteira de me aceitar como orientando.

A Germano (Titio), meu ORIENTADOR, pela paciência, ensinamentos, amizade e, especialmente, pelos aconselhamentos para a vida, antes, durante e, certamente, após o mestrado.

Ao Departamento de Química da Universidade Estadual da Paraíba, que me liberou e reduziu minha carga horária, possibilitando o desenvolvimento das minhas atividades perante o mestrado.

A EMBRAPA, por disponibilizar os equipamentos e permitir a realização das análises, especialmente ao professor Everaldo Medeiros.

## **EU VOU PRO CRATO**

Eu vou pro Crato  
Vou matar minha saudade  
Ver minha morena  
Reviver nossa amizade

Eu vou pro Crato  
Tomar banho na nascente  
Na subida do Lameiro  
Tomo estrago de aguardente

Eu vou pro Crato  
Comer arroz com pequi  
Feijão com rapadura  
Farinha do Cariri

Eu vou pro Crato  
Vou matar minha saudade  
Ver minha morena  
Reviver nossa amizade

Eu vou pro Crato  
Pois a coisa melhorou  
A luz de Paulo Afonso  
O Cariri valorizou

Eu vou pro Crato  
Já não fico mais aqui  
Cratinho de açúcar  
Coração do Cariri

Eu vou pro Crato  
Vou matar minha saudade  
Ver minha morena  
Reviver nossa amizade

Eu vou pro Crato  
Vou pra casa de seu Pedro  
Seu Felício é velho macho  
Tô com Pedro, tô sem medo

Eu vou pro Crato  
Vou viver no Cariri  
Cratinho de açúcar  
Tijolo de buriti

Luiz Gonzaga

# SUMÁRIO

---

ÍNDICE DE FIGURAS .....	i
LISTA DE ABREVIACÕES E SIGLAS .....	ii
RESUMO .....	iii
ABSTRACT .....	iv
<b>CAPÍTULO I .....</b>	
1.0 INTRODUÇÃO .....	2
<b>CAPÍTULO II .....</b>	
2.0 FUNDAMENTAÇÃO TEÓRICA .....	5
2.1 Boas Práticas de Fabricação.....	5
2.2 Controle de qualidade .....	6
2.2.1 Controle de qualidade na indústria farmacêutica .....	6
2.3 Falsificações na indústria farmacêutica .....	7
2.4 Analgésicos .....	9
2.4.1 Dipirona .....	10
2.4.2 Citrato de orfenadrina .....	11
2.4.3 Cafeína .....	12
2.4.4 Métodos analíticos empregados em análises de dipirona, citrato de orfenadrina e cafeína .....	13
2.5 Classificação .....	14
2.5.1 Análises qualitativas e quantitativas .....	14
2.5.2 Técnicas alternativas .....	14
2.6 Fundamentos da espectroscopia e quimiometria .....	15
2.6.1 Espectroscopia .....	15
2.6.1.1 Espectroscopia na região do Infravermelho .....	16
2.6.1.2 Espectroscopia NIR .....	17
2.6.1.3 Tipos de medidas .....	17
2.6.2 Quimiometria.....	18

# SUMÁRIO

---

2.6.2.1 Métodos de reconhecimento de padrão .....	19
2.6.2.2 Análise das Componentes Principais .....	20
2.6.2.3 Amostras anômalas .....	21
2.6.2.4 Análise de Agrupamentos Hierárquicos .....	21
2.6.2.5 Modelagem Independente e Flexível por Analogia de Classe .....	23
2.5.2.6 K-ésimo Vizinho mais Próximo .....	24
<b>CAPÍTULO III .....</b>	
<b>3.0 MATERIAL E MÉTODOS .....</b>	<b>26</b>
3.1 Obtenção das amostras .....	26
3.2 Espectrofotômetro de Infravermelho Próximo .....	26
3.3 Aquisição dos espectros .....	27
3.4 Tratamento dos dados e softwares empregados .....	27
<b>CAPÍTULO IV .....</b>	
<b>4.0 RESULTADOS E DISCUSSÃO .....</b>	<b>29</b>
4.1 Pré-tratamento dos dados espectrais .....	29
4.2 Análise por PCA .....	30
4.3 Análise por HCA .....	32
4.4 Modelagem SIMCA .....	34
4.5 K-ésimo Vizinho mais Próximo - KNN .....	39
<b>CAPÍTULO V .....</b>	
<b>5.0 CONCLUSÃO .....</b>	<b>41</b>
5.1 Propostas Futuras .....	41
<b>REFERÊNCIAS .....</b>	<b>42</b>
<b>APÊNDICE A .....</b>	<b>50</b>

---

<b>Figura 1.</b> Estrutura molecular da dipirona sódica .....	10
<b>Figura 2.</b> Estrutura molecular do citrato de orfenadrina .....	12
<b>Figura 3.</b> Estrutura molecular da cafeína .....	12
<b>Figura 4.</b> (A) Transmitância, (B) Transflectância e (C) Refletância difusa .....	18
<b>Figura 5.</b> Gráfico de um conjunto de dados bidimensionais ( $\lambda_1$ e $\lambda_2$ ), mostrando os eixos das componentes principais (PC1,PC2) .....	20
<b>Figura 6.</b> Exemplo de dendograma .....	22
<b>Figura 7.</b> Exemplo de modelos SIMCA .....	23
<b>Figura 8.</b> Espectros brutos para faixa de: 800 a 2499,5nm (A) e 1100 a 2499,5nm (B) .....	29
<b>Figura 9.</b> Espectros derivados para faixa de 1104 a 2495,5 nm .....	30
<b>Figura 10.</b> Gráfico de scores do modelo PCA para os dados brutos .....	31
<b>Figura 11.</b> Gráfico de scores do modelo PCA para os dados pré-processados ..	32
<b>Figura 12.</b> Dendograma para todas as amostras de comprimidos .....	33
<b>Figura 13.</b> Gráfico de scores de PC1 <i>versus</i> PC2 <i>versus</i> PC3 para medicamentos de referência (A); medicamentos similares 1 (B); medicamentos similares 2 (C) e medicamentos similares 3 (D).....	35
<b>Figura 14.</b> Gráfico linear de <i>loading</i> para PC1, PC2 e PC3 dos modelos PCA dos medicamentos: Referência 1 (A); Similar 1 (B); Similar 2 (C) e Similar 3 (D) .....	36
<b>Figura 15.</b> Gráficos dos resultados Si/S0 vs Hi do modelo SIMCA para o conjunto de amostras do medicamento: referência (A); similar 1 (B); similar 2 (C) e similar 3 (D) .....	38

---

AINE	Anti-inflamatório não esteroideal
ANVISA	Agência Nacional de Vigilância Sanitária
BPF	Boas Práticas de Fabricação
CLAE	Cromatografia Líquida de Alta Eficiência
HCA	Análise Hierárquica de Classes
ICS	Sociedade Internacional de Química
IR	Infravermelho
KNN	K-ésimo Vizinho mais Próximo
NIR	Infravermelho Próximo
OMS	Organização Mundial da Saúde
PC	Componentes Principais
PCA	Análise de Componentes Principais
PF	Polícia Federal
PRF	Polícia Rodoviária Federal
RDC	Resolução da Diretoria Colegiada
SFFC	Espúrios-Falsamente etiquetado- Falsificados-Contrafeitos
SIMCA	Modelagem Independente e Flexível por Analogia de Classe
UV -VIS	Ultravioleta e Visível

---

**Classificação simultânea de fármacos contidos em comprimidos de analgésico por espectroscopia NIR e ferramentas quimiométricas**

Autor: **Carlos Alan Dias Melo**

Orientadora: **Ana Claudia Dantas de Medeiros**

Coorientador: **José Germano Véras Neto**

O foco deste trabalho foi construir um algoritmo quimiométrico para classificação de amostras de comprimidos, a base de dipirona (300mg), cafeína (50mg) e orfenadrina (35mg), com uso da espectroscopia NIR. O banco de dados conta com 300 espectros das amostras, retirados a partir de três comprimidos, frente e verso, por lote, em 50 lotes distintos e pertencentes a quatro fabricantes distintos. Os dados foram processados a priori, selecionando um intervalo espectral de 1100 a 2499,5 nm, e, em seguida, empregou-se filtro Savitzky-Golay com primeira derivada, janela de 17 pontos e polinômio de segunda ordem. A caracterização dos comprimidos foi feita usando modelos quimiométricos baseados na Análise de Componentes Principais (*PCA*), Análise de Agrupamentos Hierárquicos (*HCA*), Modelagem Independente e Flexível por Analogia de Classe (*SIMCA*) e Método do K-ésimo Vizinho mais Próximo (*KNN*), elaborados para cada grupo de amostras de comprimidos. Para o PCA e HCA, a formação de agrupamentos para cada classe de comprimidos foi observada. Para os modelos SIMCA e KNN, o conjunto de treinamento foi construído utilizando 20 e 40 espectros, enquanto o conjunto de teste fez uso de 10 e 20 espectros, ambos de medicamentos similares e de referência, respectivamente. Todos os modelos obtiveram 100% de classificação correta, sem erros de tipo I ou de tipo II. Usando esta estratégia, foi possível classificar os comprimidos de modo rápido e não destrutivo sem a necessidade de várias determinações analíticas.

**Palavras-chave:** Espectroscopia NIR; Dipirona; Classificação; Falsificações; Quimiometria.

---

**Simultaneous classification of drug contained in tablets of analgesic by NIR spectroscopy and chemometrics tools**

Autor: **Carlos Alan Dias Melo**

Orientadora: **Ana Claudia Dantas de Medeiros**

Coorientador: **José Germano Vêras Neto**

The focus of this work was to build a chemometric algorithm for classifying samples of tablets, the basis of dipyrone (300 mg), caffeine (50mg) and orphenadrine (35mg) using NIR spectroscopy. The database has 300 spectra of samples taken from three tablets, two-sided, per lot, in 50 different lots belonging to four different manufacturers. The data were processed in advance by selecting a spectral range from 1100 to 2499,5 nm, and then was employed the filter Savitzky-Golay with first derivative, window 17 points and second-order polynomial. The characterization of the tablets was done using chemometric models based on Principal Component Analysis (PCA), Hierarchical Cluster Analysis (HCA), and Independent Modeling of Class Analogy for Flexible (SIMCA) method and the K-th Nearest Neighbor (KNN), prepared for each group of samples tablets. For PCA and HCA, the formation of clusters for each class of tablets was observed. For models SIMCA and KNN, the training set was constructed using 20 and 40 spectral measures, while the test set made use of 10 and 20 spectra, both, of similar and reference drugs, respectively. All models achieved 100% correct classification without errors type I or type II. Using this strategy it was possible to classify the tablets rapidly and non-destructively without the need for multiple analytical determinations.

**Keywords:** NIR Spectroscopy; Metamizole; Counterfeit; Chemometrics techniques

# CAPÍTULO I

## 1.0 - INTRODUÇÃO

No século XX, a partir de meados dos anos 50, métodos específicos de pesquisa e avaliação de medicamentos começaram a ser desenvolvidos para atestar propriedades fundamentais, como segurança, eficácia e qualidade. Em busca por melhorias na qualidade dos seus produtos, diminuição dos custos de produção e redução na geração de resíduos, as indústrias farmacêuticas passaram a investir no constante aperfeiçoamento dos métodos de análise, nas mais diversas áreas de produção<sup>1</sup>.

Em decorrência dos inúmeros parâmetros de qualidade avaliados em um processo industrial, normalmente, uma abundante quantidade de análises é requerida. Em sua maioria, estas análises são realizadas empregando métodos instrumentais, sendo a cromatografia líquida de alta eficiência (CLAE) a metodologia mais comumente utilizada no controle da qualidade de fármacos. Apesar de reconhecida exatidão e precisão, o método em questão caracteriza-se também por ser laborioso e oneroso, requerendo geralmente pré-tratamento das amostras, reagentes ultrapuros e operadores especializados, somados negativamente a degradação das amostras e produção de resíduos orgânicos prejudiciais ao meio ambiente<sup>2</sup>. Em função de seu custo de operação e de manutenção, essas análises são realizadas apenas em forma de amostragem, tanto nas etapas de produção (doseamento e produto final), bem como na fiscalização.

Credenciada a solucionar tal problemática, uma nova modalidade de análise vêm sendo avaliada e implementada nas indústrias farmacêuticas, denominada "screening analysis", análise de varredura. Este método remete a análise prévia de uma amostra, fundamentando-se na interação do analista em uma ou mais etapas de um processo analítico para a tomada de decisões (sim ou não, presente ou ausente, positivo ou negativo, continue ou pare, etc)<sup>3,4</sup>, o que permite condicionar a escolha e aplicação de rotas ou decisões posteriores, e optar por métodos analíticos não destrutivos, envolvendo mínima ou nenhuma manipulação da amostra, e que possibilitem obter informações em tempo real, reduzindo ao mínimo possível as análises químicas empregadas para avaliar a composição do medicamento (amostragem do produto acabado).

Para atender estes requisitos uma técnica que possibilita realizar medidas diretamente no sólido, sem necessidade de dissolução da amostra para comprimidos, por exemplo, se faz necessária. A técnica de espectroscopia no infravermelho próximo (NIR, do inglês "near infrared"), associada a quimiometria, vem começando a ser explorada em análises qualitativas de amostras sólidas<sup>5,6</sup>, sendo muito adequada para atender os requisitos anteriormente citados.

Neste contexto, o trabalho proposto teve como principal objetivo construir um algoritmo quimiométrico para classificação de amostras de comprimidos, com uso da espectroscopia NIR, que possibilitasse uma análise rápida, não invasiva, não destrutiva e de baixo custo. Para isto, alguns fármacos comercialmente disponíveis no mercado, contendo dipirona, orfenadrina e cafeína, foram escolhidos para avaliar a possibilidade da realização do controle idealizado.

# **CAPÍTULO II**

## 2.0 - FUNDAMENTAÇÃO TEÓRICA

### 2.1 - Boas Práticas de Fabricação

A qualidade de produtos, bens e serviços está vinculada a satisfação e a proteção do consumidor. No Brasil, a Agência Nacional de Vigilância Sanitária (ANVISA) tem o papel institucional de promover e proteger a saúde da população por intermédio do controle sanitário da produção e da comercialização de produtos e serviços submetidos a vigilância sanitária, incluindo os medicamentos. Assim, a legislação brasileira estabelece padrões de qualidade para produtos farmacêuticos e instituiu, entre outras normas, as Boas Práticas de Fabricação (BPF) <sup>7,8</sup>.

As BPF são vista como um tratado de garantia da qualidade de um produto e estabelecem normas e procedimentos técnicos para produção de medicamentos a fim de garantir a qualidade sanitária e a conformidade dos produtos farmacêuticos seguindo suas especificações de registro. Aplicada a toda cadeia produtiva, as boas práticas ainda fazem referência às exigências mínimas da infraestrutura laboratorial, contemplando áreas e locais específicos a exemplo dos setores de produção, armazenamento e controle de qualidade <sup>7,8</sup>.

Segundo a ANVISA <sup>8</sup>, as boas práticas de fabricação determinam que:

- ✓ todos os processos de fabricação devem ser claramente definidos e sistematicamente revisados, mostrando-se capaz de fabricar medicamentos dentro dos padrões de qualidade exigidos;
- ✓ as etapas críticas dos processos de fabricação devem ser validadas;
- ✓ as áreas de produção devem ser providas de toda infraestrutura necessária (pessoal qualificado e treinado, procedimentos aprovados, espaço, instalações, equipamentos e matérias-primas, bem como armazenamento e transporte adequados);
- ✓ registro de todas as etapas de produção que possibilitem fácil rastreamento e investigação de quaisquer desvios e
- ✓ sistema de atendimento a reclamações pós-vendas, bem como de recolhimento de respectivos lotes.

Ainda de acordo com a ANVISA <sup>8</sup> é obrigatório que todos os estabelecimentos detentores da autorização de funcionamento para fabricar medicamentos tenham um controle da qualidade, e que o mesmo seja independente dos demais departamentos, principalmente da produção.

## 2.2 - Controle da qualidade

O controle da qualidade é definido como o conjunto de procedimentos que asseguram que os ensaios necessários e relevantes sejam efetuados e que os produtos não sejam liberados até que a qualidade dos mesmos seja julgada satisfatória. O controle de qualidade tem por objetivo avaliar as características físicas, químicas e microbiológicas das matérias-primas, embalagens, produtos em processo e produtos acabados, não se limitando apenas às operações laboratoriais, abrangendo todas as decisões relacionadas à qualidade, sendo visto como um requisito necessário para a garantia da qualidade, segurança e eficácia do produto e não somente como uma exigência regulatória<sup>7,8,9</sup>.

### 2.2.1 - Controle de qualidade na indústria farmacêutica

Além das características e especificações gerais, atribuições como: estabelecer, validar e implementar procedimentos, assegurar a correta rotulagem dos recipientes de materiais e produtos, avaliar os produtos acabados considerando todos os fatores relevantes, desde as condições de produção até o exame da embalagem final, são pertinentes ao controle da qualidade no processo industrial<sup>7,8,9</sup>.

Perante a complexidade e abrangência, requisitos mínimos a serem seguidos pelo controle da qualidade nas indústrias farmacêuticas são listados, dentre tais, podem ser citados<sup>9</sup>:

- ✓ equipamentos e instalações adequadas, pessoal capacitado e procedimentos operacionais aprovados disponíveis para todo o processo de produção;
- ✓ procedimentos de amostragem, exame de embalagem e análise de produtos intermediários e acabados realizados por métodos aprovados e validados em farmacopeia;
- ✓ descrição e registro de todas atividades;
- ✓ produtos acabados atendendo as especificações conforme descrição no seu registro, apresentando composição qualitativa/quantitativa e pureza exigidas;
- ✓ registros de todos os resultados obtidos através de inspeções e ensaios de controle de qualidade;
- ✓ liberação de lotes de produtos acabados somente mediante certificação da conformidade com suas especificações e aprovação do pessoal autorizado e
- ✓ armazenamento de matérias-primas e produtos acabados em quantidades suficientes, permitindo a realização de exames futuros.

Por fim, devem estar disponíveis recursos adequados para garantir que as atividades realizadas pelo controle de qualidade sejam efetivas e confiavelmente realizadas.

Embora a maioria das indústrias farmacêuticas operem com padrões de qualidade estabelecidos internacionalmente, muitos dos medicamentos produzidos ainda chegam a alguns mercados fora desses padrões<sup>10</sup>, sendo provenientes de um controle da qualidade ineficaz ou de falsificações, realizadas ilegalmente pelas empresas, na busca de lucros ainda maiores.

### **2.3 - Falsificações na indústria farmacêutica**

A falsificação de medicamentos é parte de um amplo processo envolvendo a distribuição de fármacos que não cumprem os padrões de qualidade, segurança e eficácia. Segundo a Organização Mundial da Saúde (OMS), medicamentos Espúrios-Falsamente etiquetados-Falsificados-Contrafeitos (SFFC, do inglês, "spurious/falsey-labelled/falsified/counterfeit") são aqueles rotulados indevidamente, de forma deliberada ou fraudulenta, no que diz respeito à sua identidade ou fonte. A adulteração abrange produtos de referência, similares ou genéricos, podendo incluir produtos com princípios ativos corretos, incorretos, insuficientes ou ausentes e/ou com embalagem falsa<sup>11,12</sup>.

LEI Nº9.787, DE 10 DE FEVEREIRO DE 1999

Art. 3º

XX – Medicamento Similar – aquele que contém o mesmo ou os mesmos princípios ativos, apresenta a mesma concentração, forma farmacêutica, via de administração, posologia e indicação terapêutica, preventiva ou diagnóstica, do medicamento de referência registrado no órgão federal responsável pela vigilância sanitária, podendo diferir somente em características relativas ao tamanho e forma do produto, prazo de validade, embalagem, rotulagem, excipientes e veículos, devendo sempre ser identificado por nome comercial ou marca;

XXI – Medicamento Genérico – medicamento similar a um produto de referência ou inovador, que se pretende ser com este intercambiável, geralmente produzido após a expiração ou renúncia da proteção patentária ou de outros direitos de exclusividade, comprovada a sua eficácia, segurança e qualidade, e designado pela Denominação Comum Brasileira ou, na sua ausência, pela Denominação Comum Internacional;

XXII – Medicamento de Referência – produto inovador registrado no órgão federal responsável pela vigilância sanitária e comercializado no País, cuja eficácia,

segurança e qualidade foram comprovadas cientificamente junto ao órgão federal competente, por ocasião do registro.

Reconhecido problema mundial, o uso de medicamentos falsificados ou com desvios de qualidade pode resultar em tratamentos ineficazes, perante quantidades insuficientes de princípios ativos prescritos, e/ou provocar danos graves a saúde dos pacientes, contendo substâncias industriais tóxicas de uso e consumo inapropriado<sup>12,13</sup>.

De uma maneira geral, a incidência de medicamentos falsificados em determinadas regiões ou países é um espelho fidedigno da atuação dos sistemas de regulação dos mesmo. Na maioria dos países industrializados com sistemas eficazes de regulação e controle de mercado (ou seja, Austrália, Canadá, Japão, Nova Zelândia, a maior parte dos países da União Europeia e os Estados Unidos da América), a ocorrência de medicamentos SFFC é extremamente baixa, inferior a 1 % do valor do mercado de acordo com estimativas dos países em questão<sup>14,15</sup>. Em contrapartida, apresentando maior incidência em regiões onde os sistemas de regulação e fiscalização são frágeis, como nos países em desenvolvimento, os SFFC alcançam alarmantes 25%, contribuindo fortemente para as estatísticas que apontam como ilegal aproximadamente 10 % do mercado farmacêutico mundial<sup>14,15</sup>.

No Brasil, em decorrência do crescimento exacerbado do número de falsificadores e medicamentos falsificados, a ANVISA, em dezembro de 2008, firmou um termo de cooperação com o Conselho Nacional de Combate à Pirataria para dar maior agilidade às ações desenvolvidas pelas duas instituições. O convênio prevê a assessoria técnica da ANVISA nas ações da Polícia Federal, em especial nas zonas de fronteira, nos crimes cibernéticos, laboratórios clandestinos e na falsificação ou adulteração de medicamentos, alimentos, cosméticos ou saneantes<sup>16</sup>.

Dados estatísticos mostram que durante todo o ano de 2008, as apreensões de medicamentos contabilizaram cerca de 40 toneladas. Fruto do trabalho integrado, de janeiro a março do ano de 2009, a ANVISA já havia apreendido mais de 170 toneladas de medicamentos sem registro, falsificados, contrabandeados ou que estavam com o prazo de validade vencido, alcançando 333 toneladas de apreensões no decorrer de todo o ano, demonstrando o sucesso da parceria com as Polícias Federal (PF) e Rodoviária Federal (PRF)<sup>16,17</sup>.

Apesar da evolução significativa nos números que refletem o combate aos medicamentos falsificados, melhorias nos sistemas de identificação de amostras adulteradas

fazem-se necessárias. Visto que das 170 toneladas de produtos apreendidos no primeiro trimestre, 90 % eram oriundos de empresas de fachada que remetem medicamentos sem a devida licença expedida pelos órgãos competentes, sendo esta modalidade de delito facilmente identificada sem a necessidade de qualquer equipamento e, por isso, largamente autuada. Em contrapartida, processo envolvendo análise de medicamentos falsificados, sob julgamento da qualidade, quantidade ou tipo de princípios ativos presentes nos medicamentos e divergentes dos registros, são minoria nas apreensões<sup>16,17</sup>.

Com apenas 10% das apreensões, divididos entre contrabandos e falsificações, é notória a inviabilidade da utilização de procedimentos analíticos adequados em todos os medicamentos sob consulta, observando-se o tempo e custo dos métodos validados e utilizados pelas agências reguladoras.

Entre os medicamentos apreendidos nos três primeiros meses de 2009 estão principalmente cópias dos medicamentos Cialis<sup>®</sup> e Viagra<sup>®</sup>, usados para disfunção erétil. Dentre os analgésicos, medicamentos com dipirona em sua composição são os principais alvos de falsificações em virtude da sua popularidade e aceitação, refletidas pelos elevados índices de consumo e venda destes fármacos<sup>16,17,18</sup>.

#### **2.4 - Analgésicos**

Responsáveis por diminuir ou interromper as vias de transmissão nervosa, os analgésicos são utilizados para o alívio ou cessação de dores de diversas intensidades. Com mecanismos e propriedades distintas, essa classe medicamentosa se divide em várias subclasses, tendo como principais: os opióides e os agentes anti-inflamatórios não esteroides - AINE's<sup>19</sup>.

Com uma variedade de indicações clínicas: antipirético, analgésico e anti-inflamatórios, os anti-inflamatórios não esteroides, a exemplos dos salicilatos (ácido acetilsalicílico), dos derivados pirazolônicos (dipirona), dos derivados do para-aminofenol (paracetamol) e de compostos correlatos, atuam no alívio de dores brandas e moderadas. Enquanto que, os analgésicos opióides (morfina) possuem mecanismo de ação e ações farmacológicas muito diferentes, sendo capazes de desencadear farmacod dependência, tolerância e aliviar dores de grande intensidade e/ou crônicas<sup>19</sup>.

Contrários as regras impostas a alguns opióides, a exemplo da morfina, que tem sua comercialização restrita e seu uso exclusivo de unidades hospitalares, os medicamentos AINE's, no Brasil, são revendidos, nas farmácias, sem exigência de prescrição médica e, por isso, largamente utilizados, muitas vezes, indiscriminadamente, pela população.

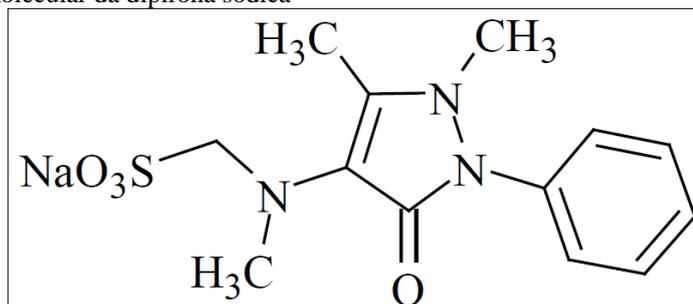
Segundo a Bayer, a dipirona é a substância mais vendida entre os analgésicos. Derivada da pirrazolona, em um grupo que ainda possui a fenilbutazona, a oxifenbutazona, a antipirina e aminopirina, atualmente, apenas a antipirina em gotas oftalmológicas está disponível nos EUA<sup>20</sup>.

#### 2.4.1 - Dipirona

A dipirona sódica, denominada por "sal de sódio do ácido 1-[(2,3-dihidro-1,5-dimetil-3-oxo-2-fenil-1H-pirazol-4-il) metilamino] metanosulfonato"<sup>21</sup> (Figura 1) é classificada farmacologicamente no grupo dos AINEs, apresentando efeitos diversos em função da dose empregada que pode ser<sup>22</sup>:

- ✓ baixa (< 10 mg.kg-1) para ação antipirética;
- ✓ mediana (15-30 mg.kg-1) para ação analgésica;
- ✓ alta (> 50 mg.kg-1) para ações anti-inflamatória e antiespasmódica.

**Figura 1.** Estrutura molecular da dipirona sódica



No Brasil, sob o nome de *novalgina*<sup>®</sup>, foi disponibilizada para venda a partir de 1922. Sua comercialização cresceu mundialmente até a década de 70, quando foram relatados graves casos de agranulocitose, condição clínica de déficit imunológico potencialmente fatal, que exige suspensão do tratamento e atendimento médico de urgência com potentes antibióticos. Com taxa de fatalidade em torno de 6 a 10 %, motivo da retirada do fármaco de muitos países, a associação de dipirona à agranulocitose exibe incidência variável em diferentes estudos<sup>23</sup>, podendo ser parcialmente explicada por diferenças em padrões de uso, doses, duração de tratamento e uso concomitante de outros medicamentos<sup>24</sup>.

Atualmente, a incidência estimada da agranulocitose é aceitável, pois oscila entre 0,2 a 2 casos por milhão de pessoas por dia de uso, níveis bem abaixo dos atribuídos ao antipsicótico clozapina, com incidência de 8 casos para cada 1000 pacientes após 6 meses de tratamento<sup>22,23</sup>.

Frente a tal problemática, em alguns países (Japão, Estados Unidos, Suécia, etc.), mediante a disponibilidade de drogas substitutas, a exemplo dos AINEs para o tratamento de dor (paracetamol, salicilatos, etc) com menor toxicidade, a dipirona foi retirada do mercado e teve sua comercialização proibida. Fundamentado nesta decisão e comprovando a viabilidade da intercambialidade entre estes medicamentos, em estudos recentes, a dipirona 500 mg demonstrou similar eficácia a 400 mg de ibuprofeno, 600 a 650 mg de ácido acetilsalicílico e 500 a 1000 mg de paracetamol no tratamento de dor pós-operatória de moderada a intensa<sup>25</sup>.

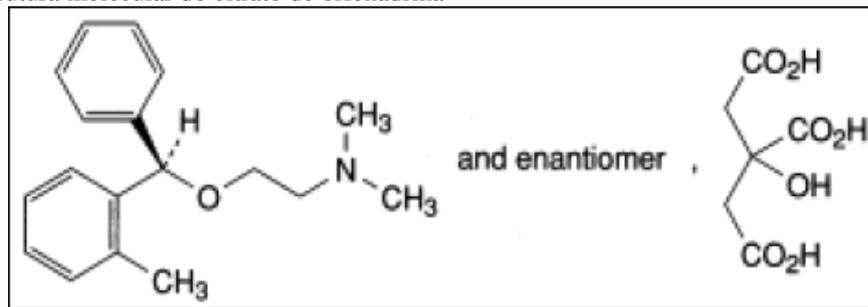
Por outro lado, a dipirona foi reintroduzida na Europa a aproximadamente 15 anos, pois estudos epidemiológicos sugeriram que o risco dos efeitos adversos era similar ao paracetamol e menor do que o ácido acetilsalicílico. Entretanto, seu uso continua limitado<sup>20</sup>.

Entretanto, bem como em diversos outros países, no Brasil não houve suspensão da sua comercialização e atualmente, segundo dados do laboratório Boehringer Ingelheim, a dipirona é um dos principais analgésicos da nossa terapêutica, ocupando, nesta classe, a primeira colocação nas vendas com 39 % do mercado, seguido pelo paracetamol (30 %) e o ácido acetilsalicílico (26 %)<sup>26</sup>.

Comercializada principalmente na forma sódica em diferentes formulações farmacêuticas (solução oral, injetável, comprimidos e supositórios), a dipirona encontra-se comercialmente disponível nas mais diversas associações medicamentosas. Dentre as mais conhecidas e com maiores níveis de vendas estão as associações com 1) Cloridrato de papaverina e extrato fluido de *Atropa beladonna*; 2) Maleato de clorfeniramina e ácido ascórbico; 3) Prometazina e adifenina; 4) Mucato de isometepteno e cafeína e 5) Citrato de orfenadrina e cafeína. Esta última, de acordo com a consultoria IMS Health, através de seu medicamento de referência, se apresenta como o medicamento mais vendido no Brasil<sup>27</sup>.

#### 2.4.2 - *Citrato de orfenadrina*

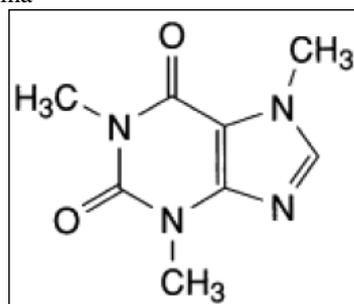
Comumente associado a analgésicos, o citrato de orfenadrina, denominado por (RS)-N, N-Dimetil-2-[(2-metilfenil) fenilmetoxi] etanaminadihidrogeno-2-hidroxiopropano-1,2,3-tricarboxilato<sup>28</sup> (Figura 2) é relaxante da musculatura esquelética, amplamente utilizado para tratar a espasticidade que pode ocorrer na síndrome do neurônio motor superior e a dor ou espasmo muscular decorrente de condições musculoesqueléticas periféricas. Essas condições são comumente encontradas na clínica e em muitos pacientes pode causar incapacidade funcional e dor, reduzindo marcadamente a qualidade de vida<sup>29</sup>.

**Figura 2.** Estrutura molecular do citrato de orfenadrina

Alguns autores afirmam que os benefícios terapêuticos dos relaxantes musculares parecem estar relacionados às propriedades sedativas dessas drogas, proporcionando alívio aos sinais e sintomas do espasmo muscular. Apesar de que nenhuma delas se mostrou mais eficaz do que os analgésicos e anti-inflamatórios na melhora da dor dos pacientes<sup>19</sup>.

#### 2.4.3 - Cafeína

Alcaloide pertencente ao grupo das metilxantinas, a cafeína, denomina quimicamente por 1,3,7-Trimethyl-3,7-dihydro-1*H*-purine-2,6-dione<sup>28</sup> (Figura 3), está presente naturalmente no café, guaraná, erva-mate, chocolate, sendo ainda adicionada em muitas bebidas. Utilizada no combate ao letargo e depressão provocada por analgésicos, a cafeína pode produzir um estado de excitação psíquica que favorece o trabalho intelectual e a aplicação mental, afastando o sono e a sensação de fadiga, bem como potencializando os efeitos dos fármacos associados.

**Figura 3.** Estrutura molecular da cafeína

Um estudo sobre os efeitos da cafeína na saúde humana indicou que seu consumo moderado (máximo de 4,6 mg/kg de peso)<sup>30</sup>, praticado por adultos saudáveis em idade reprodutiva, não está associado a efeitos adversos. Em compensação, a excessiva ingestão de cafeína pode causar efeitos indesejáveis, como tremor, taquicardia e problemas gastrointestinais<sup>30,31</sup>.

#### 2.4.4 - Métodos analíticos empregados em análises de dipirona, citrato de orfenadrina e cafeína

Em casos de medicamentos com um único fármaco, as farmacopeias brasileira<sup>21</sup>, britânica<sup>28</sup> e europeia<sup>32</sup> indicam a titulação iodimétrica para a determinação de dipirona. Apesar disso, este método é limitado em relação ao tempo de análise e a precisão, necessitando ainda que a solução seja resfriada a 10 °C, para evitar a perda de iodo por volatilização.

No caso da cafeína, segundo as mesmas farmacopeias, procedimentos como titulações em meio não aquoso, necessidade de aquecimento, utilização de substâncias como anidrido acético, ácido perclórico e benzeno, encarecem, retardam e dificultam a realização do método<sup>21,28,32</sup>. No entanto, ao contrário da dipirona, a cafeína também está presente na farmacopeia americana<sup>33</sup>, onde, diferentemente das demais, ao invés de métodos titulométricos, é recomendada a cromatografia líquida de alta eficiência como técnica padrão para quantificação.

Por sua vez, ausente na farmacopeia brasileira, conforme indicações das farmacopeias britânica<sup>28</sup> e europeia<sup>32</sup>, o citrato de orfenadrina segue uma metodologia semelhante aos demais princípios ativos da formulação, onde o procedimento para determinação do fármaco envolve o uso de substâncias como anidrido acético e ácido perclórico, além de metodologias titulométricas e potenciométricas. Contudo, assim como para cafeína, na farmacopeia americana<sup>33</sup>, o citrato de orfenadrina também apresenta CLAE como metodologia validada.

Distinções a parte, as técnicas citadas acima são validadas, aceitas e indicadas nas farmacopeias, o que garante a obtenção de resultados satisfatórios na análise dos princípios ativos em separado. Porém, de acordo com a farmacopeia americana<sup>33</sup>, quando dois ou mais ingredientes ativos estão presentes em uma mesma formulação, a quantificação deve ser realizada por CLAE com detecção ultravioleta (UV)<sup>34</sup>. Ciente das desvantagens inerentes aos métodos oficiais para determinações individuais (dipirona, cafeína e orfenadrina) e multicomponentes, pesquisadores têm proposto inúmeros métodos para análise de fármacos, onde diferentes técnicas são empregadas, a exemplo da espectroscopia<sup>35,36</sup>, amperometria<sup>37</sup>, eletroforese capilar<sup>38</sup>, dentre outras.

De fácil percepção e recorrente nas metodologias já validadas, os métodos propostos, em sua grande maioria, fazem uso das análises quantitativas e buscam exaustivamente a determinação e quantificação dos componentes de uma formulação. Contrário a esta realidade, procedimentos para uma simples e eficaz classificação de fármacos inexistem nas

farmacopeias, não dispondo de qualquer tipo de validação, ofertando um leque de oportunidade e promovendo seu estudo neste trabalho.

## 2.5 - Classificação

A classificação tem por finalidade agrupar amostras de acordo com suas características. Para comprimidos, atributos como a quantidade, qualidade, ausência ou presença de adjuvantes e princípios ativos são alguns dos parâmetros que afetam o processo classificatório e distribuem as amostras em grupos semelhante ou diferentes.

### 2.5.1 - Análises qualitativas e quantitativas

Ainda que compoñham a grande maioria das metodologias já validadas, fundamentados em um único princípio: "limites farmacopéicos aceitáveis de não menos que 90 % e não mais do que 110 %",<sup>33</sup> os procedimentos analíticos quantitativos, quando realizados em amostras complexas, com muitas substâncias, apresentam limitações como<sup>39</sup>:

- ✓ necessidade de determinar quantidades ínfimas, as vezes menores do que o limite de quantificação disponível e
- ✓ interferências na quantificação, diante de uma quantidade elevada de matrizes de interesse.

Na presença das circunstâncias citadas, quando submetidos a fiscalização de órgão reguladores ou controle da qualidade de uma empresa, medicamentos com qualidade duvidosa, acrescidos de uma quantidade não detectável de uma substância tóxica podem ser considerados legítimos e apropriados para uso humano, respectivamente, desde que apresentem quantidades de princípios ativos dentro do limite imposto pelas farmacopeias<sup>39</sup>.

Contrário e alheio às limitações e desvantagens dos métodos quantitativos, as técnicas de análises qualitativas são independentes do número e quantidade de ingredientes (ativos/adjuvantes) presentes e do tipo de formulação avaliado, proporcionando execução mais simples, além de resultados rápidos e satisfatórios.

### 2.5.2 - Técnicas Alternativas

Apesar das várias possibilidades para classificação e determinação dos componentes de uma formulação, nas indústrias farmacêuticas e nas "blitz" de órgão fiscalizadores, um grande número de análises precisa ser realizado num curto intervalo de tempo. Por isto, são mais recomendados procedimentos rápidos e simples de serem executados. Deste modo,

técnicas analíticas que não requerem o uso de substâncias químicas e/ou etapas de pré-tratamento são uma boa alternativa para os métodos convencionais.

Com esta visão, algumas técnicas, a exemplo da espectroscopia NIR<sup>40,41</sup>, têm sido utilizadas para a classificação de fármacos no estado sólido, estejam eles na sua forma individual (princípio ativo único) ou em misturas. Pouco explorada em compostos de interesse farmacológico, esta ferramenta analítica apresenta inúmeras vantagens, fornecendo resultados rápidos e de modo não destrutivo, podendo simplificar e agilizar muitas rotinas de análise, permitindo, inclusive, o controle "on-line" e "in-situ"<sup>42</sup>.

Desafortunadamente, a complexidade do sinal proporcionado pela espectroscopia NIR dificulta o estabelecimento de rotinas de análise de medicamentos, principalmente em razão das frequentes interferências observadas entre fármacos ou entre os fármacos e os excipientes<sup>43</sup>. Contudo, felizmente, métodos quimiométricos tem sido utilizados com sucesso na resolução de problemas como os anteriormente descritos, viabilizando o desenvolvimento de rotinas analíticas espectroscópicas, mesmo frente a severa interferência espectral<sup>39</sup>.

## **2.6 - Fundamentos da espectroscopia e quimiometria**

Atualmente, evoluindo de acordo com as necessidades de cada época, as técnicas analíticas progredem acerca da precisão de um monitoramento contínuo dos processos de produção de fármacos e medicamentos. Tratando de produtos ofertados a pessoas já debilitadas, a premência de um controle da qualidade rigoroso é mais do que evidente.

Com o intuito de promover avanços nos controles da qualidade, diminuir o número de medicamentos inaptos ao consumo disponíveis para venda e, dessa maneira, garantir a qualidade, segurança e eficácia dos medicamentos, diversas técnicas tem sido utilizadas e novas metodologias propostas. Dentre elas, a espectroscopia NIR vem demonstrando resultados positivos e angariando atenções<sup>40,41,43</sup>.

### *2.6.1 - Espectroscopia*

A análise espectrométrica é um processo instrumental de medição baseado nas propriedades de absorção e emissão de energia eletromagnética. Sendo uma técnica consolidada, de fácil implementação, ampla aplicabilidade, boa exatidão, baixo custo e sensibilidade compatível com outras técnicas analíticas, tornou-se uma grande aliada no controle da qualidade de produtos farmacêuticos ou nas demais áreas da análise química<sup>39</sup>.

Com aplicação em regiões distintas do espectro eletromagnético, a metodologia em questão se apresenta como espectroscopia na Região do Ultravioleta(UV), do visível (VIS) e

do Infravermelho (IR). Introduzida durante a segunda guerra mundial e em evidência até os dias atuais, a espectroscopia no infravermelho caracteriza-se por ser rápida e confiável, com emprego na indústria<sup>39</sup>, sendo largamente utilizada nos controle da qualidade.

#### 2.6.1.1 - Espectroscopia na região do Infravermelho

A região espectral do infravermelho compreende radiação com números de onda no intervalo de aproximadamente 12.800 a 10  $\text{cm}^{-1}$ , se dividindo em infravermelho próximo, médio e distante, Tabela 1.

**Tabela 1:** Regiões espectrais do Infravermelho<sup>44</sup>

Região Espectral	Abreviação	Comprimento de onda	
		$\text{cm}^{-1}$	Nm
Infravermelho Próximo	NIR	12.800 a 4.000	700 a 2.500
Infravermelho Médio	MID	4.000 a 200	2.500 a 15.000
Infravermelho Distante	FIR	200 a 10	15.000 a 50.000

A aplicação da espectroscopia no infravermelho como ferramenta analítica começou apenas no período da segunda guerra mundial, no setor de controle da qualidade em algumas indústrias químicas alemãs. Apesar de fornecer uma gama de informações espectrais, devido a maior complexidade nos espectros, menor largura das bandas e as limitações instrumentais do aparelho, determinações quantitativas eram praticamente impossíveis de serem extraídas, restringindo-se basicamente a aplicações qualitativas ou para reforçar hipóteses propostas sobre a estrutura química das espécies<sup>45,46</sup>.

Tida como uma das maiores dificuldades do método, grande parte do número de dados não possui informações correlacionadas diretamente com o analito de interesse, sendo utilizadas diversas metodologias de processamento, como escalonamento, cálculos de derivadas e filtragem digital para remoção de ruído afim de evitar distorções no modelo, e consequentemente a conclusões errôneas nas análises<sup>47</sup>.

A região do infravermelho médio ainda é a área que concentra o maior interesse dos químicos, principalmente para interpretação de estruturas químicas orgânicas. Apesar de que nos últimos anos, as demais regiões tem obtido bastante atenção, como é o caso do infravermelho próximo, onde estudos envolvendo análises farmacêuticas vem sendo realizadas frequentemente. Isto é devido a sua boa determinação qualitativa e quantitativa de parâmetros físicos e químicos, habilidade no processo, além da compatibilidade com a fibra ótica, que acaba por facilitar as análises<sup>48</sup>.

### 2.6.1.2 - Espectroscopia NIR

É uma técnica analítica individualizada, devido a utilização da faixa de comprimento de onda interposta entre o infravermelho médio até o visível (780 a 2500 nm), a espectroscopia NIR se relaciona com sobretons e combinações vibracionais fundamentais dos grupos funcionais –CH, –NH e –OH, Tabela 2, apresentando bandas de absorção mais acentuadas nesta região<sup>44</sup>.

Caracterizadas negativamente por serem largas, sobrepostas e até 100 vezes mais fracas quando comparadas as da região do infravermelho médio, as bandas de absorção NIR, entretanto, vantajosamente, apresentam baixa absorvidade e elevado poder de penetração. Aspectos estes que permitem análises diretas de fortes absorventes e nivela amostras com espalhamento, a exemplo dos sólidos nos modos de reflectância ou transmitância sem necessidade de pré-tratamento<sup>43,44,49</sup>.

Apontados como a principal dificuldade para o uso da espectroscopia NIR, a complexidade dos seus espectros ocultam inúmeras informações químicas relevantes, correlacionadas com propriedades físicas, dificultando ainda mais a interpretação direta dos dados obtidos. Entretanto, para extração de informações a partir dos espectros NIR das amostras, é indispensável à utilização de ferramentas quimiométricas<sup>43,44,49</sup>.

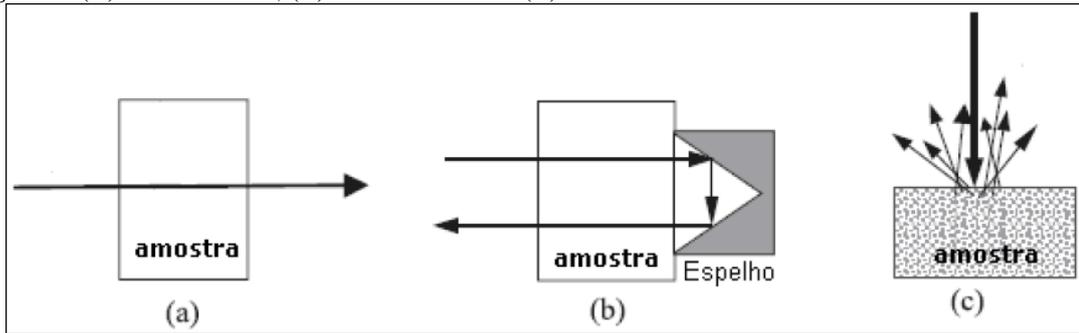
**Tabela 2:** Comprimentos de onda referentes à absorção na região do infravermelho próximo de grupos funcionais<sup>44</sup>.

Grupo Funcional	Comprimento de Onda (nm)
-CH <sub>alifático</sub>	1.212, 1.414, 1.725, 1.760 e 2.308
-CH <sub>aromático</sub>	1.077, 1.111 e a faixa 1.615-1.665
-OH <sub>água</sub>	1.440 e 1.940
-OH <sub>fenóis</sub>	Faixas de 1.500 a 1.800 e 2.000 a 2.200
-NH <sub>aminas</sub>	1.534, 2.000 e 2.100
-NH <sub>amidas</sub>	1980, 2050 e 2180

### 2.6.1.3 – Tipos de medidas

Primordialmente vinculada a informações contidas em uma pequena parte da radiação refletida de maneira difusa por uma amostra sólida. Na atualidade, a espectroscopia NIR apresenta diversas maneiras para realizar medidas analíticas nesta região, onde, além da inicial reflectância difusa, podem ser citadas transmitância e transreflectância (Figura 4)<sup>49,50</sup>.

**Figura 4.** (A) Transmitância, (B) Transflectância e (C) Reflectância difusa<sup>50</sup>



Para medidas por transmitância (Figura 4a) ou transflectância (Figura 4b), cubetas de vidro ou quartzo são utilizadas como suporte para análise das amostras. Com espessuras que podem variar de 1 a 50 mm, o tipo de cubeta, bem como a extensão do caminho óptico, são definidas de acordo com a região espectral estudada<sup>49,50</sup>.

Facilmente visualizadas a partir da Figura 4, as metodologias se diferenciam basicamente pelo caminho óptico duplo adotado pela transflectância, a qual o feixe de radiação percorre a amostras duas vezes.

Com aplicação em amostras sólidas, a refletância difusa ocorre em superfícies não totalmente planas, podendo o substrato ser contínuo ou fragmentado. Neste processo de reflexão, o feixe incidente penetra a superfície da amostra interagindo com a matriz, retornando a superfície da mesma, após absorção parcial e múltiplos espalhamentos, conforme a Figura 4c. Com isto, a radiação que sofre a refletância difusa fornece informações qualitativas e quantitativas, sobre a natureza química da amostra<sup>51</sup>.

### 2.6.2 - Quimiometria

A Sociedade Internacional de Quimiometras (ICS, do inglês "International Chemometrics Society"):

“Quimiometria é a ciência relacionada a medidas realizadas em um sistema ou processo químico, obtendo informações sobre o estado do sistema através da aplicação de métodos matemáticos ou estatísticos”.

Com desenvolvimento inicial datado na década de 1960, a quimiometria foi utilizada na validação dos mais diversos métodos em distintas áreas da ciência. Utilizando-se de métodos instrumentais computadorizados, as análises químicas passaram a dispor de uma enorme quantidade de dados, contrastando com a realidade até ali, onde a tomada de decisões era dependente de uma quantidade escassa de informações. Desde então, grandes quantidades

de dados passaram a ser obtidas com maior rapidez e menor desprendimento de tempo, promovendo acúmulo de informações em pequeno espaço de tempo<sup>43,44,49,51</sup>.

Diante da necessidade em se analisar tamanha quantidade de dados e obter o máximo de informações relevantes, criou-se a quimiometria, área mais recente da química analítica, responsável pelo processamento e interpretação dos dados de maneira rápida e eficaz, possibilitando o emprego destes para fins analíticos<sup>43,44,49,51</sup>.

Na atualidade os sistemas multivariados permitem a medição de muitas variáveis simultaneamente. Nestes sistemas, a conversão da resposta instrumental no dado químico de interesse requer a utilização de técnicas de estatística multivariada, álgebra matricial e análise numérica. A organização dos conjuntos de dados em análises multivariadas é geralmente expressa em uma matriz, onde o conjunto de dados das amostras fica disposto em linhas e as colunas dizem respeito às variáveis medidas<sup>43,44,49,51</sup>.

Na conjuntura da análise de medicamentos (comprimidos), com auxílio da espectroscopia NIR, técnicas quimiométricas são utilizadas para o tratamento dos dados encontrando aplicações distintas, dentre as quais, a cargo das técnicas de reconhecimento de padrão, observa-se a classificação de dados e amostras.

#### 2.6.2.1 - Métodos de reconhecimento de padrão

São técnicas responsáveis por identificar as semelhança e diferenças das amostras, a partir do grau de similaridade entre suas propriedades físicas e químicas, comparando-as entre si. As técnicas quimiométricas de reconhecimento de padrão promovem o agrupamentos destas, possibilitando, inclusive, a identificação de amostras anômalas (do inglês, "outliers").

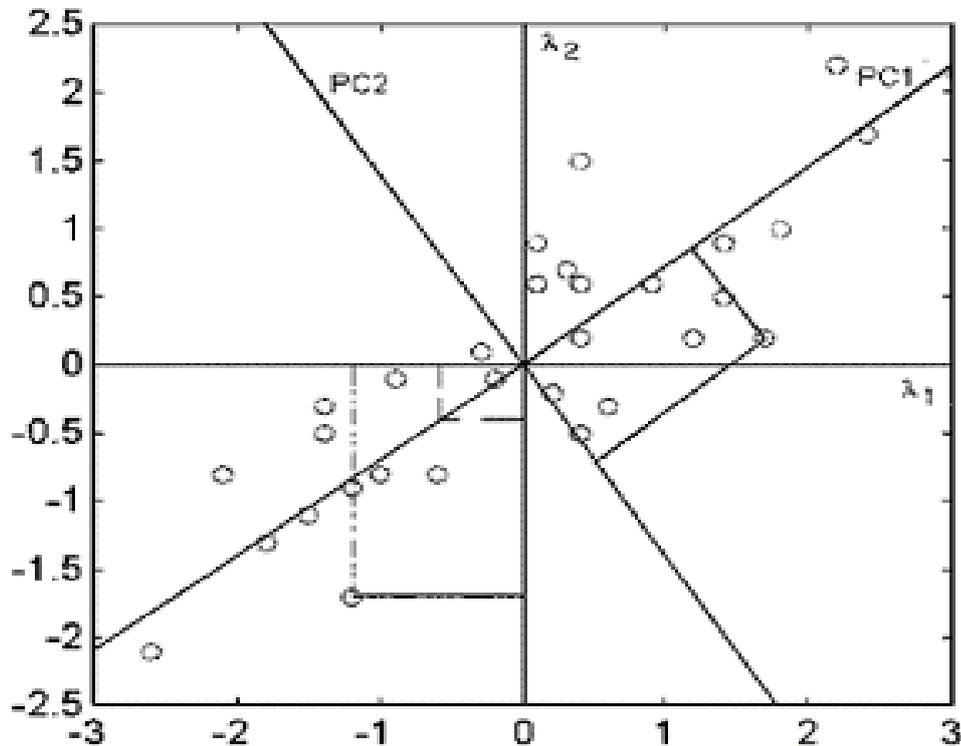
Dentre essas técnicas estão as **não supervisionadas**, que são a Análise de Componentes Principais (PCA, do inglês, "Principal Component Analysis") e Análise de Agrupamentos Hierárquicos (HCA, do inglês, "Hierarchical Cluster Analysis"), que servem para a análise exploratória dos dados. Ou seja, são usados para examinar similaridades ou diferenças entre amostras, bem como para identificar a formação de padrões no espaço multidimensional. Já as técnicas **supervisionadas**, a exemplo da Modelagem Independente e Flexível por Analogia de Classe (SIMCA, do inglês, "Soft Independent Modelling of Class Analogy") e Método do K-ésimo Vizinheiro mais Próximo (KNN, do inglês, "K-Nearest Neighbor"), servem para classificação das amostras ou agrupamento das mesmas em classes de acordo com suas características<sup>52</sup>.

### 2.6.2.2 - Análise das Componentes Principais

A PCA é a base fundamental da maioria dos métodos modernos para o tratamento de dados multivariados, sendo aplicada para compressão desses dados sem que haja perda de informações importantes. A redução dos dados permite que seu conjunto possa ser representado com menor número de novas variáveis, denominadas de fatores ou componentes principais (PC, do inglês, "Principal Component"), que são combinações lineares das variáveis originais<sup>53,54</sup>.

Esta técnica atua decompondo os dados originais, transformando-os em uma soma de matrizes, produtos de vetores chamados de escores e pesos, que podem ser calculados por mínimos quadrados. A Figura 5 mostra um gráfico bidimensional de um conjunto de trinta amostras ( $n = 30$ ). A matriz de dados consiste, neste caso, de duas colunas representando as medidas de intensidades registradas para dois comprimentos de onda  $\lambda_1$  e  $\lambda_2$  nas 30 amostras<sup>53,54</sup>.

**Figura 5.** Gráfico de um conjunto de dados bidimensionais ( $\lambda_1$  e  $\lambda_2$ ), mostrando os eixos das componentes principais (PC1, PC2)<sup>54</sup>



A primeira componente principal (PC1) é definida na direção de máxima variância do conjunto de dados. A segunda componente principal (PC2) é definida na direção que descreve a máxima variância no espaço da PC1, de forma que cada componente principal (PC1, PC2,

PC3, etc.) seja responsável pela fração sucessiva de variância de dados, consistindo em um sistema de coordenadas ortogonais entre si e, portanto, não correlacionadas. Normalmente as primeiras PCs, explicam a maior parte da variância total contida nos dados e podem ser usadas para representá-las.

Na análise exploratória são examinadas as relações entre as amostras e as variáveis, através de gráficos de escores ("scores") e pesos ("loadings") respectivamente, os quais permitem também avaliar a influência de cada variável em cada amostra, encontrando similaridade ou diferenças nos dados<sup>53,54</sup>.

Resumidamente, com o método *PCA* é possível<sup>55</sup>:

- ✓ demonstrar as diferenças entre as formas das amostras;
- ✓ definir quais variáveis contribuem mais para estas diferenças;
- ✓ identificar quais variáveis contribuem de forma correlacionadas ou independentemente uma da outra;
- ✓ detectar padrão de amostras (agrupamentos);
- ✓ diferenciar informações úteis de ruídos;
- ✓ identificar amostras anômalas ("outliers"), etc.

#### 2.6.2.3 - Amostras anômalas

Caracterizadas, geralmente, por serem bastante diferentes ou apresentarem erros grosseiros quando comparadas a maioria do dados, as amostras anômalas, a depender da amplitude do seu afastamento, às vezes, podem levar a conclusões errôneas. Como causas mais decorrentes deste fenômeno, estão: erros do analista, experimentais, instrumentais ou presença de compostos químicos de diferentes composições<sup>56</sup>.

Dentre diversas possibilidades já encontradas, a forma mais comum de anomalia corresponde àquela proveniente de amostras, especificamente, oriunda de dados de alguma amostra que difere do conjunto total de calibração por causa de ruídos ou porque esta reflete alguma propriedade única (um fenômeno físico ou algum tipo particular de não linearidade)<sup>56</sup>.

Amostras inicialmente separadas e designadas como díspares das demais, devem, com cautela, ser diferenciadas e definidas como uma anomalia ou uma amostra informativa. E, a partir disso, definir qual procedimento adotar: eliminá-las ou mantê-las.

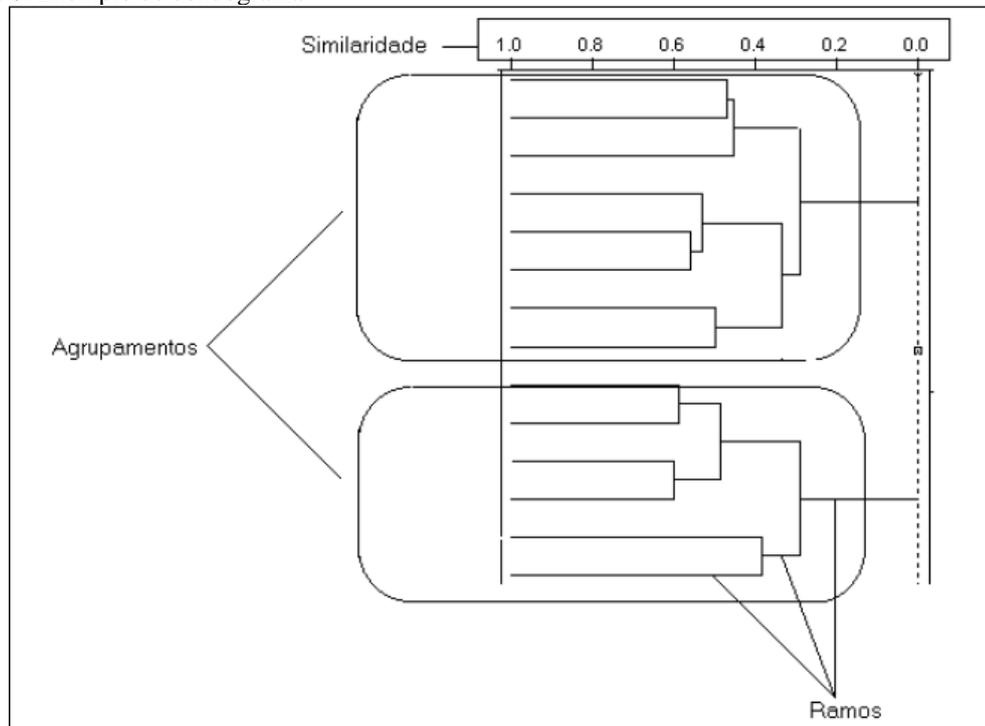
#### 2.6.2.4 - Análise de Agrupamentos Hierárquicos

Sem nenhuma caracterização prévia, as técnicas de agrupamento averiguam as relações existentes dentro de um conjunto de dados. Estas técnicas podem ser de dois tipos:

divisiva ou aglomerativa. No primeiro caso, todas as amostras constituem um único grupo, que será sequencialmente subdividido em grupos menores, subgrupos de acordo com a similaridade entre as amostras, até que cada uma delas dê origem a um grupo. Na técnica aglomerativa ocorre o contrário, onde, inicialmente, cada amostra é considerada um grupo e, conforme semelhanças apresentadas, vão sendo agrupadas de maneira sequencial e contínua até que formem um único grupo<sup>55</sup>.

Definida como aglomerativa e não supervisionada, a HCA observa as distâncias entre os pontos de todas as amostras do conjunto de dados, demonstrando os resultados em um dendograma (Figura 6), gráfico bidimensional com informações características, onde é possível visualizar os agrupamentos e similaridades entre amostras e/ou variáveis<sup>55,57</sup>.

**Figura 6.** Exemplo de dendograma<sup>57</sup>



A construção dos dendogramas é fundamentada na proximidade existente entre as amostras no espaço. Realizada em todas as amostras do conjunto, em pares, as distâncias calculadas definem a matriz de similaridade, indicando a distância entre os agrupamentos, da maior para menor, respectivamente (Equação 1). A cada passo, os dois grupos mais similares vão se juntando e o processo vai se repetindo até que forme um único agrupamento<sup>54</sup>.

$$S_{ij} = 1 - \frac{d_{ij}}{d_{\max}} \quad (1)$$

onde :

$S_{ij}$  é a similaridade entre duas amostras (ou agrupamentos);

$d_{ij}$  é a distância entre as amostras;

$d_{\max}$  é a maior distância encontrada entre todas as amostras do conjunto.

A aproximação de dois subgrupos pode ocorrer por três métodos: do vizinho mais distante, do vizinho mais próximo e pela média. Esta última forma pode ser calculada de várias maneiras, sendo a mais simples e utilizada para a medida de similaridade entre pontos num conjunto de dados, a distância euclidiana<sup>57</sup>.

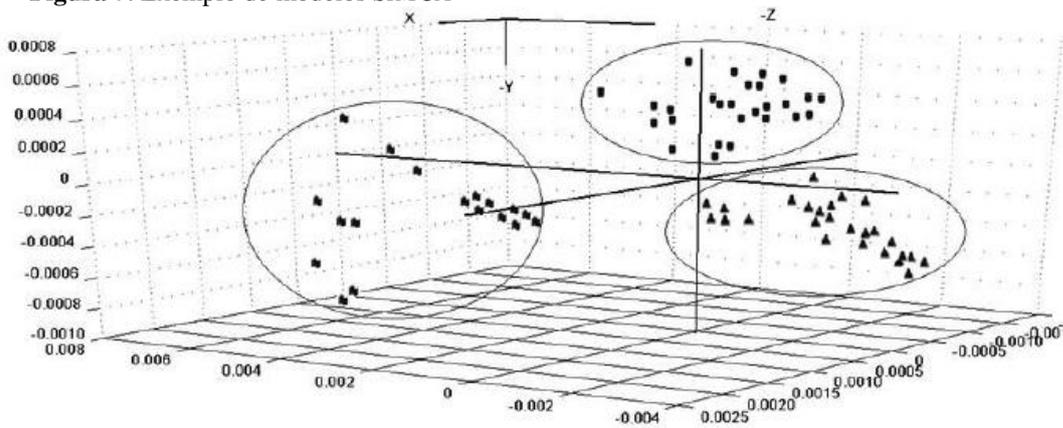
Esse método pode ser aplicado na identificação de grupos dentro de um conjunto de dados, para testar hipóteses de agrupamentos, na identificação de membros de um dado grupo ou na formação mais conveniente de agrupamentos com características um tanto quanto diversas<sup>55,57</sup>.

#### 2.6.2.5 - Modelagem Independente e Flexível por Analogia de Classe

A SIMCA é uma técnica de reconhecimento padrão supervisionada, utilizada para classificação de amostras futuras, que faz uso da PCA para modelar a forma e posição do objeto no espaço, definindo-lhe uma classe. Para esta técnica, um modelo multidimensional é construído para cada classe e a classificação das amostras é efetuada determinando dentro de qual classe a amostra está posicionada. Para isso, um número ótimo de PC's é previamente definido para cada classe e o modelo final é obtido através da demarcação dos limites entre cada PCA modelado. De maneira sequencial, por consequência, uma hipercaixa com tantas dimensões quantas forem o número de PC mais adequados para cada classe é obtida (Figura 7).

Para modelar ou delimitar as fronteiras onde estão localizadas as amostras, é utilizado um conjunto de treinamento contendo amostras de todas as classes e cuja identidade de cada amostra deste conjunto é conhecido a priori. As fronteiras de cada classe são modeladas ou delimitadas usando apenas as amostras daquela classe conhecida no conjunto de treinamento. Uma vez definidas as hipercaixas é importante verificar se existem classes com sobreposição

**Figura 7.** Exemplo de modelos *SIMCA*<sup>58</sup>



A capacidade de discernimento dos modelos SIMCA se fundamenta na inexistência de sobreposição entre duas classes diferentes, bem como na maior distância entre as classes possível. A classificação de uma amostra externa àquelas empregadas na modelagem SIMCA se baseia nas diferentes projeções possíveis no espaço dos escores, avaliando se a mesma encontra-se dentro dos limites definidos para a respectiva classe em todas as projeções previamente obtidas<sup>39</sup>.

Apresentando resultados expressos em uma tabela, onde asteriscos demarcam as amostras classificadas e as classes pertencentes, o modelo construído é passível de erros, podendo os mesmos serem do tipo I e do tipo II, caracterizados pela não inclusão da amostra na classe a qual pertence e a inserção da amostra em uma classe a qual não pertence, respectivamente<sup>58</sup>.

O SIMCA é um método bem estabelecido na literatura e é largamente utilizado para classificação de amostras em conjuntos de dados com alta dimensionalidade. Tem sido aplicado com sucesso em diferentes matrizes, incluindo: medicamentos<sup>5,53,59-61</sup>, alimentos<sup>63,64</sup>, biodiesel<sup>58,65,66</sup>, entre outras<sup>67</sup>.

#### 2.6.2.6 - K-ésimo Vizinho mais Próximo

Técnica de reconhecimento de padrão supervisionada, KNN é caracterizada pela comparação entre as distâncias das amostras, fazendo uso da distância euclidiana, calculada para cada par de amostras (Equação 2)<sup>68</sup>.

$$d_{ab} = \left[ \sum_j^m (x_{aj} - x_{bj})^2 \right]^{1/2} \quad (2)$$

A classificação de amostras externas às amostras modeladas (utilizadas na construção de classes), é determinada através do cálculo da distância multivariada destas novas amostras com relação as k-ésimas amostras vizinhas empregadas na modelagem. O método mais comum para a validação deste tipo de classificação baseia-se na validação cruzada para a qual cada uma das amostras do conjunto de modelagem é retirada uma de cada vez. Com isto, são determinados quantos vizinhos k-ésimos devem ser considerados para posterior classificação. Finalmente, para estimar a classe de uma amostra externa são verificadas quantas vezes amostras de uma mesma classe são atribuídas como vizinhas à amostra em questão, sendo considerada a classe final aquela a qual maior número de vizinhos seja relacionado a esta amostra<sup>51</sup>.

# **CAPÍTULO III**

### 3.0 - MATERIAL E MÉTODOS

#### 3.1 - Obtenção das amostras

Para realização das análises foram selecionados comprimidos de analgésico, contendo três fármacos em associação, são eles: dipirona (300mg), citrato de orfenadrina (35mg) e cafeína (50 mg).

As amostras foram constituídas por medicamentos de quatro marcas distintas, sendo uma de referência e três similares. Os excipientes que compõem cada medicamento estão indicados na Tabela 3.

Para realização das análises foram utilizados 10 lotes diferentes de cada similar para realização das análises e 20 lotes diferentes do medicamento de referência. Visto que o medicamento de referência é o material de comparação aos demais e uma maior quantidade de lotes permite garantir um método estatisticamente mais confiável.

**Tabela 3.** Lista de excipientes utilizados na produção dos medicamentos, segundo seus fabricantes.

Medicamento	Excipientes
Referência 1	amido de milho, amidoglicolato de sódio, talco e estearato de magnésio
Similar 1	amido de milho, álcool etílico, celulose microcristalizada MC 102, polivinilpirrolidona e estearato de magnésio
Similar 2	álcool etílico, amido de milho, amidoglicolato de sódio, dióxido de silício coloidal, edetato dissódico, estearato de magnésio, lactose, metabissulfito de sódio, celulose microcristalina e polivinilpirrolidona
Similar 3	álcool etílico, amido de milho, estearato de magnésio, polivinilpirrolidona e talco

As amostras foram adquiridas em farmácias e drogarias, em várias cidades dos estados do Ceará e da Paraíba.

#### 3.2 - Espectrofotômetro de Infravermelho Próximo

Os espectros de absorção molecular, na região do infravermelho próximo, foram obtidos em um espectrofotômetro *XSD near-infrared " Rapid Content<sup>TM</sup> Analyzer"* (FOSS). Na região de 800 a 2499,5 nm, as medidas espectrais de refletância difusa foram realizadas em comprimidos intactos, com resolução espectral de 0,5 nm.

Toda as medidas espectrais foram feitas na Empresa Brasileira de Pesquisa Agropecuária - EMBRAPA ALGODÃO, com sede em Campina Grande - Paraíba.

Para controle, obtenção, armazenamento e tratamento dos dados foi utilizados um microcomputador com processador Intel® Core I5 2.5 GHz.

### ***3.3 - Aquisição dos espectros***

Inicialmente, cada lote teve 3 (três) comprimidos analisados frente e verso, minimizando possíveis distorções causadas por vincos ou má uniformidade do fármacos, totalizando 300 espectros. Posteriormente, os pares de espectros (frente e verso), de cada comprimido, foram reduzidos a 1(um), mediante a obtenção da média, diminuindo para 150 espectros o conjunto de dados a ser utilizado na classificação.

### ***3.4 - Tratamento dos dados e softwares empregados***

O pré-processamentos dos dados foi realizado no software Unscrambler® 9.8, com seleção da faixa espectral de 1100 a 2499,5 nm e emprego do filtro Savitzky-Golay com primeira derivada, janela de 17 pontos e polinômio de segunda ordem. O mesmo programa computacional foi utilizado na construção dos modelos PCA e SIMCA.

Para verificação dos agrupamentos das marcas dos comprimidos, a partir do software Statistica® 9.0, o modelo de HCA foi construído e expresso em dendograma. Por fim, fazendo uso do software MATLAB® 7.11.0.584 (R2010b), foram realizados procedimentos como: seleção do conjunto de classificação, construção do modelo KNN e teste para construção do modelo SIMCA, utilizando o algoritmo de Kennard-Stone.

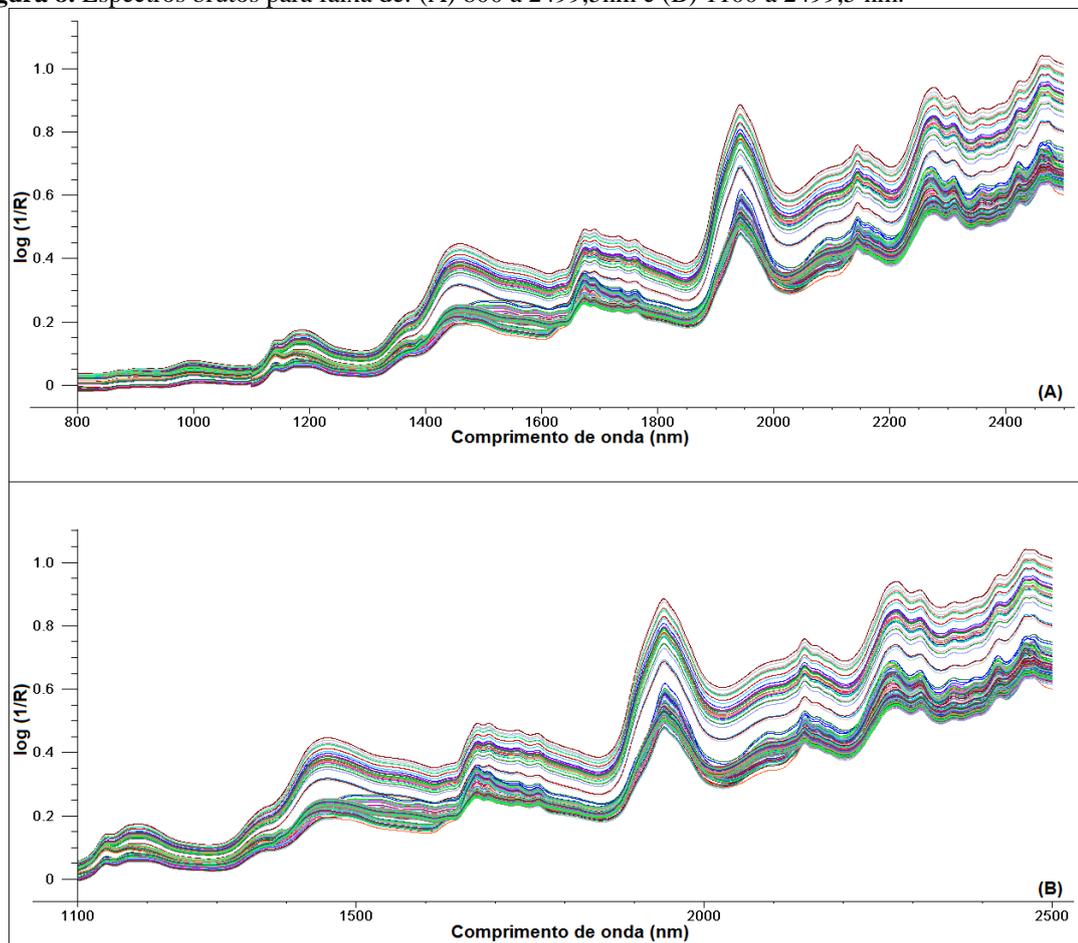
# **CAPÍTULO IV**

## 4.0 - RESULTADOS E DISCUSSÃO

### 4.1 - Pré-tratamento dos dados espectrais

Na Figura 8A, o conjunto de espectros obtidos diretamente do equipamento "XSD near-infrared", com região de 800 a 2499,5 nm, é apresentado. A partir da figura, é possível perceber que a faixa espectral compreendida entre 800 a 1099,5 não apresenta sinal, expressando apenas a linha de base. Em decorrência da ausência de informações relevantes para classificação das amostras, as variáveis foram selecionadas a priori, na faixa de 1100 a 2499,5 nm, Figura 8B.

**Figura 8.** Espectros brutos para faixa de: (A) 800 a 2499,5nm e (B) 1100 a 2499,5 nm.

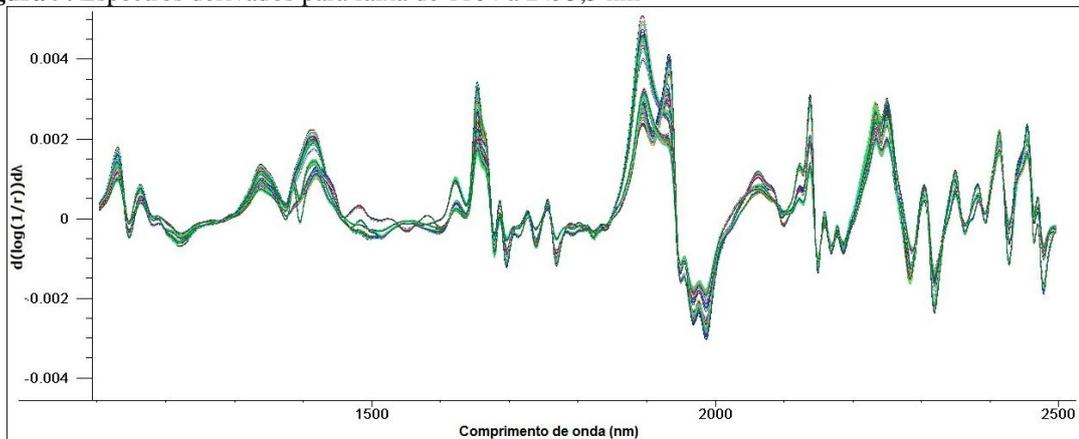


Posteriormente, os dados na faixa de variáveis selecionada a priori sofreram pré-processamento: primeira derivada pelo método de Savitzky-Golay, com janela de 17 pontos e polinômio de segunda ordem. Os dados pré-processados estão apresentados na Figura 9. Este tipo pré-processamento permite correção da linha de base, problema comum a espectroscopia NIR, que faz decair o sinal com a diminuição dos comprimentos de onda. Além disso, sinais

espúrios podem ser eliminados no processo e derivação, utilizando para isso os dados delimitados pela janela selecionada.

Com o pré-processamento, a faixa de trabalho foi de 1104,0 a 2495,5 nm, devido a resolução de 0,5 nm e visto que na derivação dos dados, para uma janela de 17 pontos, sendo um ponto central mais oito para cada lado. Os oito primeiros e os oito últimos pontos não possuem a janela preenchida para obterem valores e assim, os dados são excluídos no resultado final do pré-processamento.

**Figura 9.** Espectros derivados para faixa de 1104 a 2495,5 nm



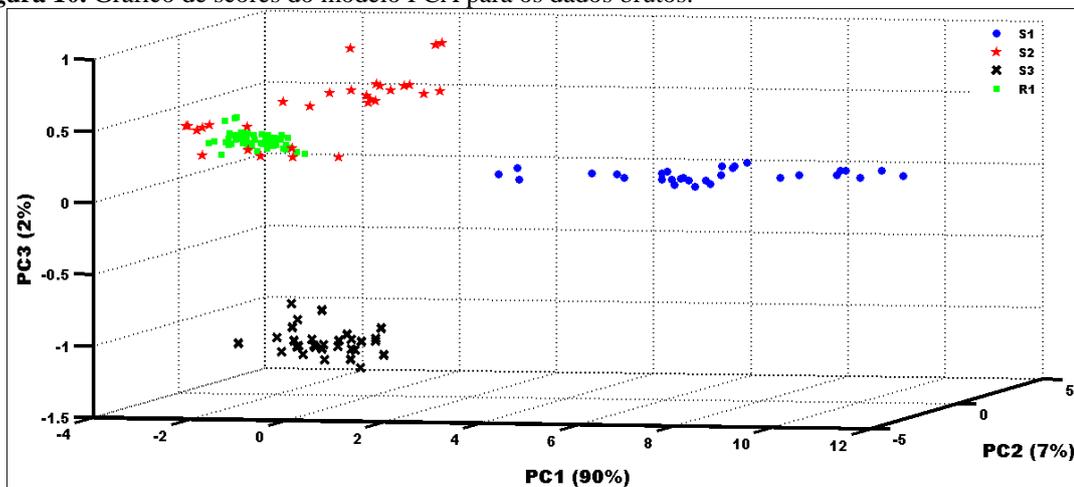
Em suma, o procedimento aplicado permitiu realçar variações de sinais que poderiam ficar ocultas durante o tratamento quimiométrico, influenciando diretamente na classificação. Em prova disso, é possível notar que, mesmo as faixas de comprimentos que apresentam absorbâncias e sinais reduzidos (ex. 1108,5 a 1200 nm) foram ressaltadas e passaram a apresentar picos mais expressivos.

#### 4.2 - Análise por PCA

Visando a análise exploratória dos dados, um tratamento quimiométrico por análise dos componentes principais foi aplicado. Para construção do modelo foram utilizados os 150 espectros disponíveis, com utilização de 3 PC's e aplicou-se procedimentos de validação cruzada completa, (do inglês, "full cross validation"). Neste tipo de procedimento, a partir de um grupo específico, uma primeira amostra é retirada do conjunto, permitindo a construção do modelo com as demais amostras pertencentes ao grupo, e, em seguida, o modelo é testado com a amostra suprimida inicialmente. O mesmo processo se repete com todas as amostras, uma a uma, até que a última seja eliminada e testada. Por fim, todos os modelos dão origem, por média, a um único e robusto modelo.

Na Figura 10 está apresentado a disposição espacial do gráfico de scores do modelo PCA, para os dados brutos, o qual permite avaliar o comportamento das amostras e grupos entre si. Apesar de expressivos 99 % de variância explicada conjunta, o modelo foi incapaz de separar todos os agrupamentos ali presentes. Responsáveis pela separação de dois dos quatro grupos, as PC1 (90%) e PC3 (2%) foram responsáveis, cada uma, por separar o conjunto de medicamentos S1 e S3, respectivamente. Enquanto que, a PC2 (7%) não conseguiu separar nenhum dos agrupamentos.

**Figura 10.** Gráfico de scores do modelo PCA para os dados brutos.



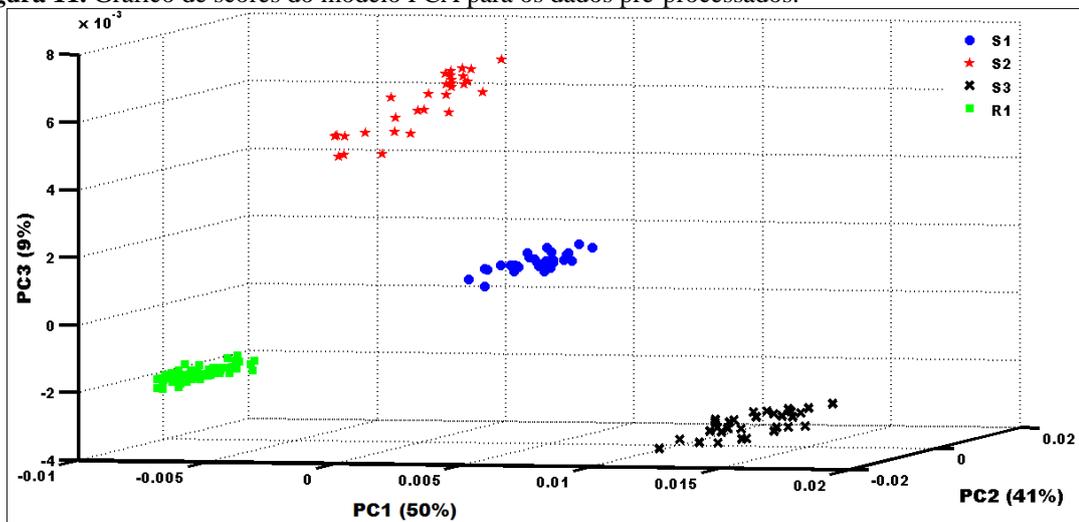
No geral, é perceptível problemas como a sobreposição de grupos (S2 e R1) e o distanciamento de amostras pertencentes ao mesmo agrupamento (S1 e S2), fatores responsáveis por possíveis erros na classificação das amostras. Esta sobreposição nos dados brutos corrobora a necessidade de um pré-processamento de dados, conforme descrito no item 4.1.

Por outro lado, diante de um pré-processamento adequado, é possível perceber que os sinais de absorbância das amostras tornou a separação dos grupos mais nítida, baseando-se em PC1 versus PC2 versus PC3, o modelo PCA para espectros derivados exhibe quatro agrupamentos bem definidos, com ausência de sobreposição e melhor compactação entre as amostras de cada grupo (Figura 11).

O gráfico de scores do modelo PCA dos dados pré-processados apresentou uma variância explicada conjunta de 100%, que, quando comparada ao gráfico dos dados brutos, demonstra uma distribuição mais uniforme da variância entre suas 3 PC's. No mesmo, a PC1 (50%) foi responsável pela separação dos grupos R1 e S3 de S1/S2, já a PC2 (41%) separou o grupo S1 dos demais e, por fim, a PC3 foi capaz de separar os grupos S2 e S1 de R1/S3.

Além da separação adequada de todos os agrupamentos, neste modelo, foi possível perceber que as amostras dos grupos S1 e S2, anteriormente dispersas, se apresentaram mais agrupadas. Ainda foi possível destacar que, dentre os agrupamentos visíveis nas Figura 10 e Figura 11, as amostras do medicamentos de referência se mantiveram bem agrupadas, onde, este resultado aponta para uma maior uniformidade na fabricação dos comprimidos, mesmo considerando lotes distintos e em maior quantidade do medicamento de referência em relação aos demais.

**Figura 11.** Gráfico de scores do modelo PCA para os dados pré-processados.



Desta maneira, ao se confrontar os dados expressos nos resultados e as diferenças entre as quatro formulações foi possível notar que: a similaridade entre os agrupamentos R1 e S2 pode ser decorrente da presença do excipiente amidoglicolato de sódio, uma vez que o mesmo está presente apenas nestas formulações. Com menor influência do que o amidoglicolato de sódio, o talco se apresenta como possível expoente da similaridade entre S3 com R1, estando presente nas duas formulações.

A possível justificativa para a dissimilaridade do agrupamento dos medicamentos similar 1 para com os demais grupos, ficou a caráter da ausência dos dois excipientes responsáveis pela aproximação dos conjuntos referência 1, similar 2 e similar 3, amidoglicolato de sódio e talco.

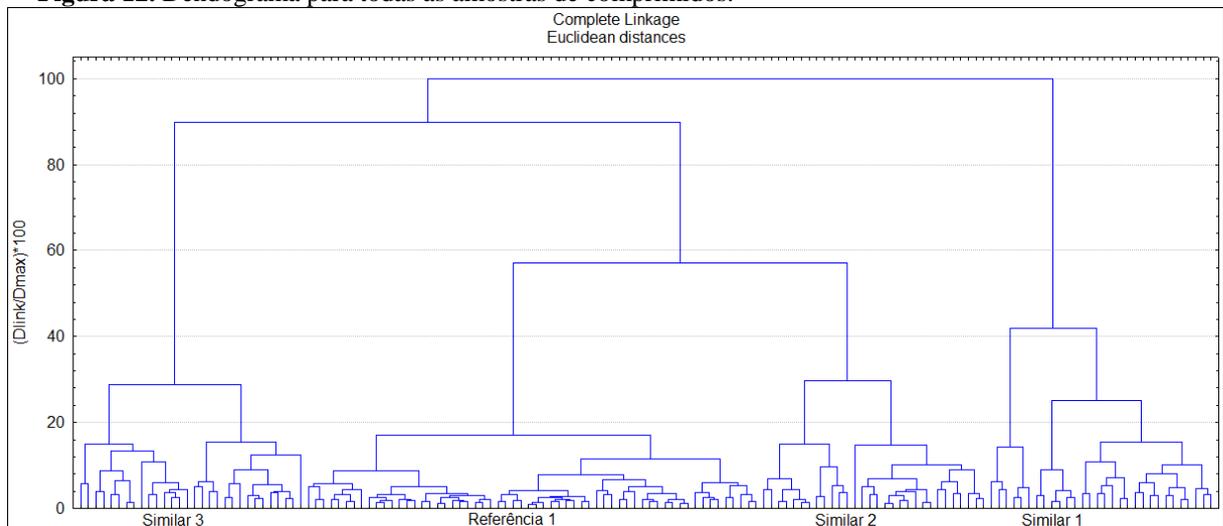
### 4.3 - Análise por HCA

Resultante de uma análise hierárquica de agrupamentos, o dendograma, Figura 12, foi constituído para todas as 150 amostras de comprimidos e utilizou a distância euclidiana como métrica de distância entre as amostras, com a técnica de ligação completa (do inglês,

"complete linkage"), responsável por apontar a proximidade entre os agrupamentos das amostras por meio da regra dos vizinhos mais distantes, e correlação da distância euclidiana, determinando a extensão em que valores de duas variáveis são proporcionais entre si, para medida de dissimilaridades entre amostras.

Conforme o máximo distanciamento das ligações obtidas, os agrupamentos formados apresentaram distâncias de ligações aproximadas, sendo estas: agrupamento do medicamento de referência (distância de ligação de 17,01 %), agrupamento do medicamento similar 1 (distância de ligação de 41,88 %), agrupamento do medicamento similar 2 (distância de ligação de 29,67 %) e agrupamento do medicamento similar 3 (distância de ligação de 28,89 %). Entre os agrupamentos, verificou-se a diferenciação nas distâncias, com base na distância de ligação máxima de: agrupamento do referência 1 para o similar 2 (distância de ligação de 57,09 %), agrupamento do similar 3 para o agrupamento R1/S2 (distância de ligação de 89,70 %) e agrupamento do similar S1 para o agrupamento de R1/S2/S3 (distância de ligação de 100 %).

**Figura 12.** Dendrograma para todas as amostras de comprimidos.



A partir da análise do dendrograma e das distâncias de ligação intragrupois pode-se afirmar que o conjunto de amostras do medicamento de referência foi o que apresentou maior similaridade entre suas amostras, corroborando com os dados do expressos no PCA. Por outro lado, com maior média para as distâncias de ligação, o conjunto das amostras do medicamento similar 1 despontou como o agrupamento de menor uniformidade. Já na análise das ligações intergrupais, o agrupamento dos medicamentos de referência para o similar 2 apresentou a maior similaridade (57,09 %), enquanto que, o agrupamento de R1/S2/S3 para S1 se apresentou completamente distinta, atingindo a máxima distância de ligação (100 %).

As similaridades e dissimilaridades apresentadas entre os grupos de fármacos são, possivelmente, oriundas dos tipos e quantidades distintas de excipientes utilizados na fabricação de cada medicamento. Distinções provenientes de princípios ativos foram descartadas, tendo em vista que, os ingredientes ativos em todas as composições são do mesmo tipo e devem estar presentes em quantidades iguais, se não, semelhantes, com variações inferiores a 5%.

Desta maneira, o modelo HCA construído para as amostras de medicamentos corrobora com os resultados apresentados no modelo PCA, onde se pode verificar a presença de quatro agrupamentos associados a marca de cada comprimido.

Considerando as análises exploratórias, PCA e HCA, estudadas para verificar a formação de agrupamentos das amostras de acordo com as classes definidas, é factível que os medicamentos de marcas diferentes possam ser separados utilizando técnicas supervisionadas de classificação.

#### **4.4 - Modelagem SIMCA**

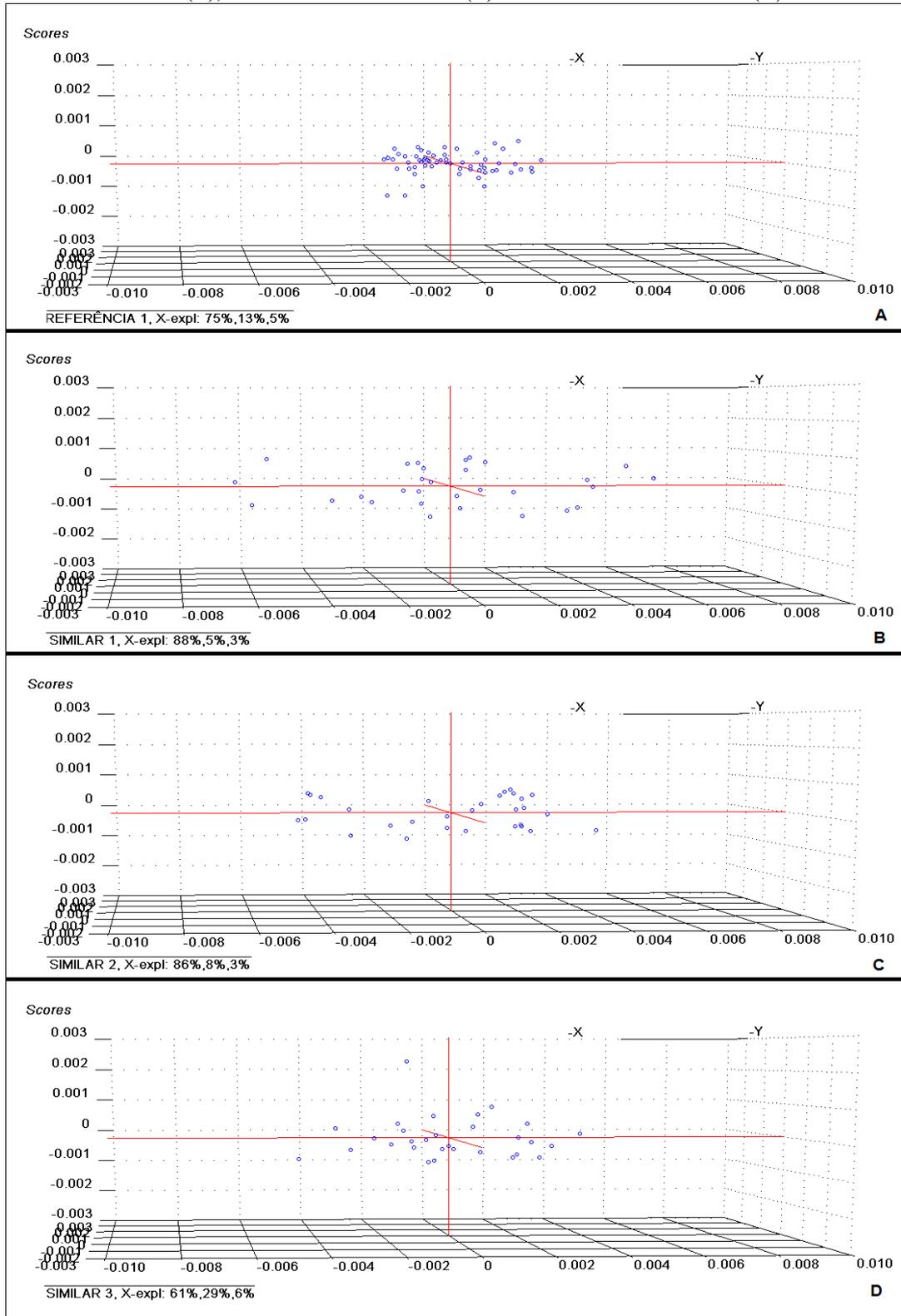
Como fase inicial para construção de um modelo SIMCA, os modelos de PCA para cada classe estudada são necessárias. Desta maneira, foram construídos 4 (quatro) modelos distintos de PCA, cada um para a sua respectiva classe de medicamentos, estando estes representados em gráficos de scores (Figura 13) e gráficos de loadings (Figura 14).

Em uma análise de componentes principais, o agrupamento das amostras define a estrutura dos dados através de gráficos de scores e loadings, cujos eixos são componentes principais (PCs), nos quais os dados são projetados. Os scores fornecem a composição das PCs em relação às amostras, enquanto os loadings fornecem essa mesma composição em relação às variáveis. Como, as PCs são ortogonais é possível examinar as relações entre amostras e variáveis através dos gráficos dos scores e dos loadings.

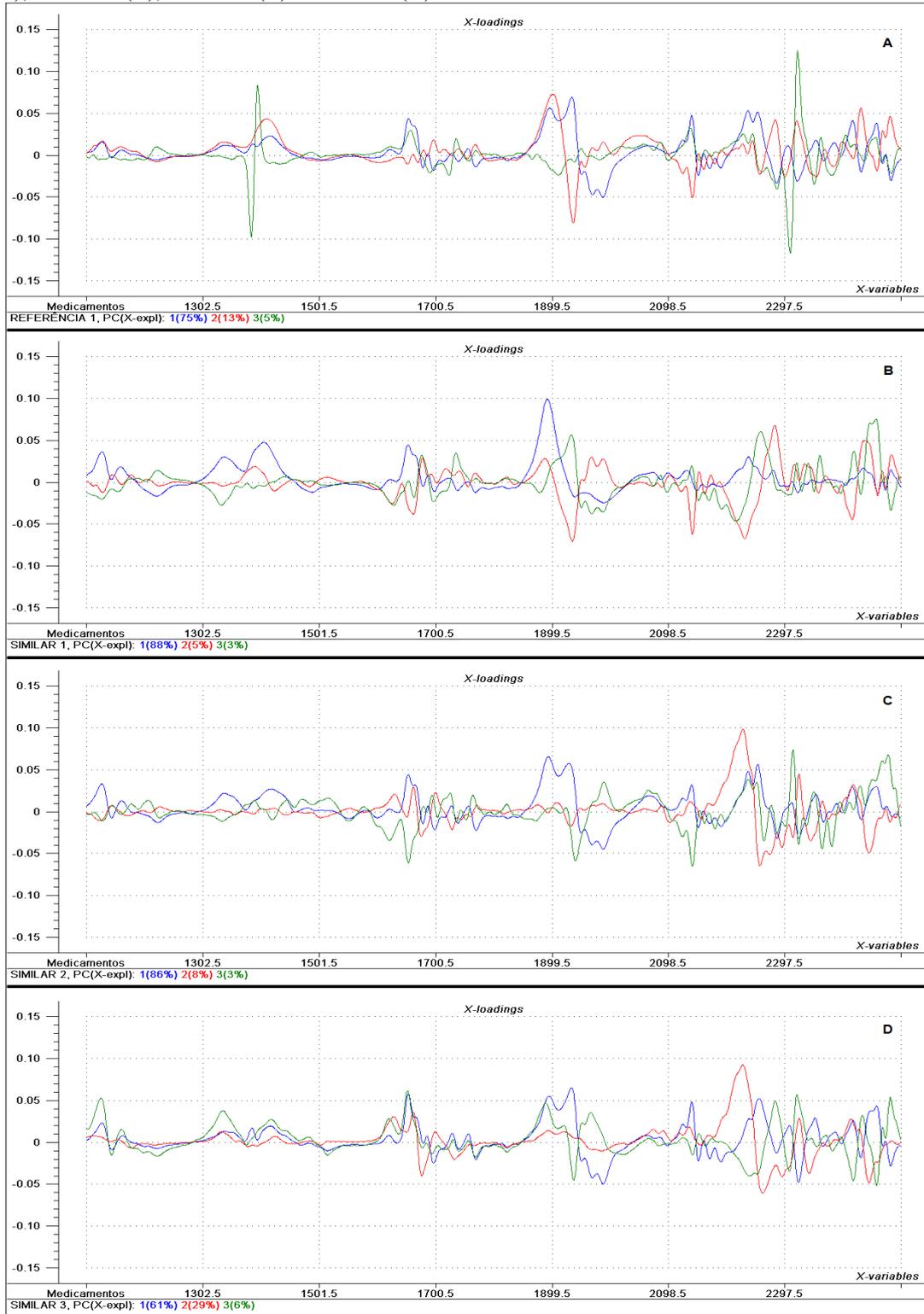
Fazendo uso das mesmas proporções dimensionais, facilitando a visualização e interpretação do dados, todos os gráficos de scores foram construídos empregando as 3 PC's utilizadas nos respectivos modelos de PCA. A partir da análise do conjunto de gráficos de scores e das informações ali contidas, pode-se notar que: o grupo R1 (Figura 13A) apresentou uma uniformidade amostral superior aos demais componentes do estudo, o que, apesar da maior quantidade de amostras, se era esperado para um agrupamento de medicamentos de referência. De maneira contrária, o conjunto de amostras do medicamento similar 1 (Figura 13B) se apresentou bastante disperso, evidenciando a disparidade entre suas amostras e, por fim, os agrupamentos S2 (Figura 13C) e S3 (Figura 13D) exibiram uniformidades

compatíveis entre si. Assim, todos os resultados aqui impressos confirmam as informações descritas no item 4.3.

**Figura 13.** Gráfico de scores de PC1 versus PC2 versus PC3 para medicamentos de referência (A); medicamentos similares 1 (B); medicamentos similares 2 (C) e medicamentos similares 3 (D).



**Figura 14.** Gráfico linear de loadings para PC1, PC2 e PC3 dos modelos PCA dos medicamentos: Referência 1 (A); Similar 1 (B); Similar 2 (C) e Similar 3 (D)



Responsáveis por expressar a composição das PC's em relação as variáveis, os gráficos loadings foram construídos para que pudessem corroborar com os demais resultados já

obtidos, proporcionando uma maior solidez aos mesmos e aumentando a confiabilidade no processo de classificação.

A partir dos mesmos é possível apontar quais PC's e respectivos comprimentos de ondas são influentes na atribuição de semelhanças e diferenças entre os compostos. A exemplo, é possível verificar a presença de picos distintos com máxima nos comprimentos de ondas 1386, 1399, 2310 e 2322 nm na PC3 do medicamento de referência, o que os responsabilizam por possíveis dissimilaridades.

Também pôde-se notar a presença de picos semelhantes expressos em PC's diferentes, como foi o caso dos comprimentos de onda com máxima em aproximadamente 1940nm, encontrados na PC2 e na PC3 para os conjunto de medicamentos R1 e S1 e os agrupamentos S2 e S3, respectivamente. Em ambos os casos, as bandas de transição expressam a presença de agrupamentos -OH, provavelmente, em decorrência da presença de álcool etílico e/ou água utilizados no processo de fabricação.

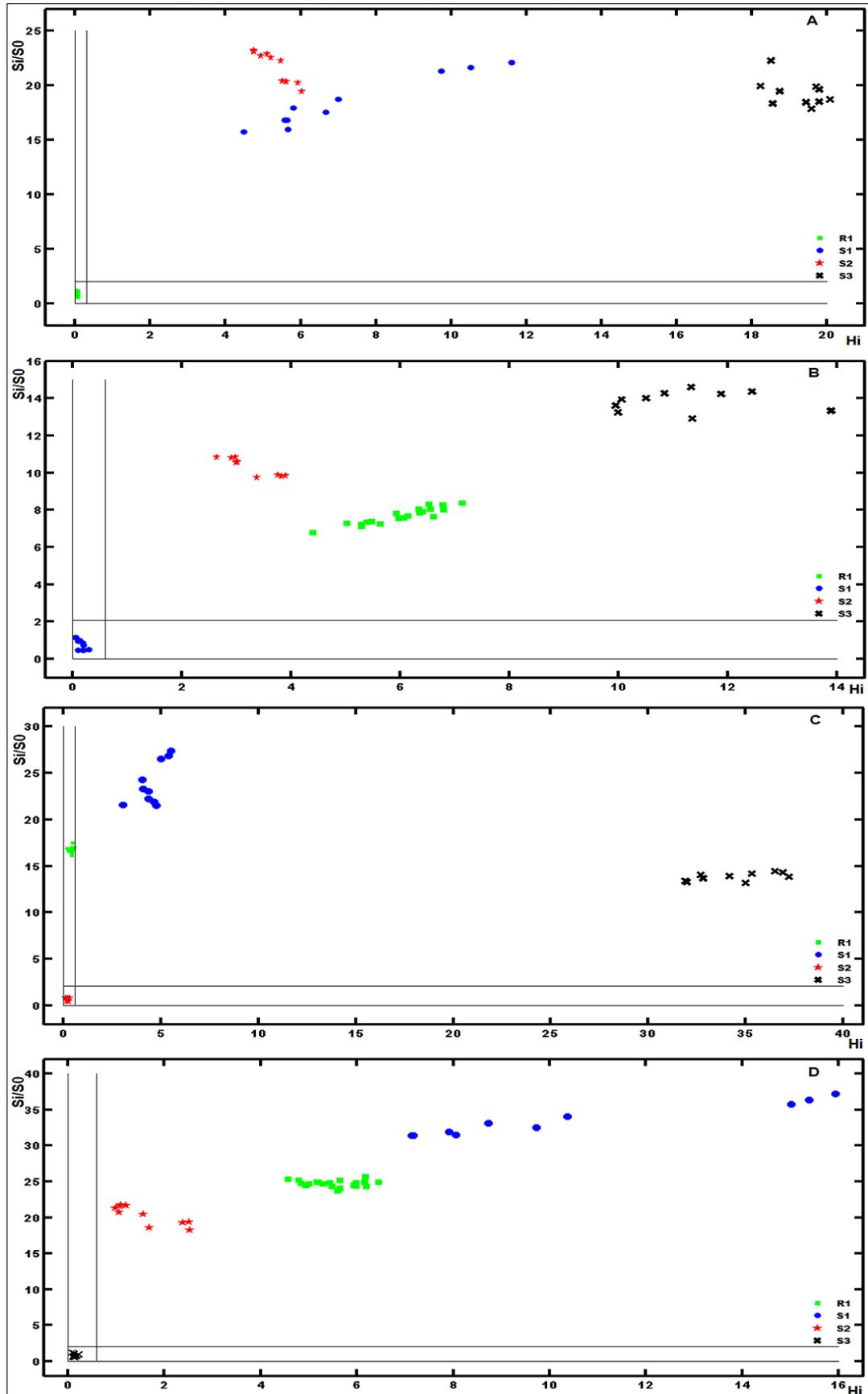
Para construção dos modelos SIMCA, o algoritmo de Kennard-Stone foi utilizado para particionamento das amostras, dando origem a um conjunto de treinamento, utilizado para construção propriamente dita do modelo e sua validação interna, e um conjunto de teste, utilizado para aplicar o modelo a amostras não utilizadas em sua construção.

Com um total de 150 espectros, o conjunto de treinamento foi compostos por 20 amostras de cada medicamento similar (S1, S2 e S3) e por 40 amostras do fármaco de referência (R1). Constituído pelos 100 espectros restantes no banco de dados, o conjunto teste foi formado por 10 espectros para os medicamentos similares e 20 para os de referência. Vale salientar que um modelo de classificação deve ser construído com, no mínimo, 10 amostras para que se permita precisão adequada para a construção do modelo.

Os gráficos mostram os dois limites utilizados para classificar uma amostra a um dado modelo, em que  $S_i/S_0$  é a distância relativa de amostras ao modelo por meio do desvio padrão residual e o leverage ( $H_i$ ) é a distância da nova amostra ao centro do modelo, que indica o nível de influência da amostra no modelo.<sup>59</sup> Nos gráficos  $S_i/S_0$  versus  $H_i$  dos modelos SIMCA para as amostras dos medicamentos, Figura 15, é possível verificar as distâncias ortogonais das amostras dos conjuntos de teste, indicando os limites para classificação dentro de cada classe, a 95 % de confiança estatística.

Visualizando os gráficos do modelo SIMCA podemos perceber que todos as amostras foram enquadrados dentro dos dois limites,  $S_i/S_0$  e  $H_i$ , caracterizando que as mesmas pertencem a aquela classe em questão. Assim sendo, a classificação foi realizada sem a presença de erros.

**Figura 15.** Gráficos dos resultados Si/S0 vs Hi do modelo SIMCA, para o conjunto de amostras do medicamento: referência (A); similar 1 (B); similar 2 (C) e similar 3 (D).



Na Figura 15C, o gráfico  $S_i/S_0$  versus  $H_i$  do modelo SIMCA para o conjunto de medicamentos similar 2 apresentou o agrupamento dos medicamentos de referência bem próximo ao centro do modelo, demonstrando a influência destas amostras sobre a sua construção. Estas informações, corroboram com a análise de cluster, onde os agrupamentos em questão apresentam uma maior similaridade dentre os quatro conjuntos estudados.

Como se era esperado, na Figura 15D, o gráfico expressou informações condizentes com o que já se pode notar na análise do dendograma e dos gráficos de loadings, onde o conjunto de amostras do medicamento similar 3 apresentou maior proximidade e similaridade dos agrupamentos R1 e S2, e se distanciou bastante do grupo S1, demonstrando a baixa influência deste sobre a construção do seu modelo.

A modelagem SIMCA apresentou bom desempenho em todos os níveis de significância estatística testados (1%, 5%, 10% e 25%), classificando adequadamente todas as amostras, não apresentando erros do tipo I e tipo II

#### **4.5 - K-ésimo Vizinho mais Próximo - KNN**

De maneira complementar as técnicas anteriormente citadas, ainda neste estudo, o conjunto de amostra foi submetido ao algoritmo KNN. O modelo KNN foi construído comparando a distância das amostras do conjunto de teste com as amostras de cada uma das classes do conjunto de classificação em relação ao centroide de cada uma das classes. Foi considerado dentro de uma dada classe, a amostra que tivesse o maior número de vizinhos da referida classe. Assim, foram testadas as amostras até o seu limite e as quatro classes de medicamentos (S1, S2, S3 e R1) foram devidamente separadas, possibilitando a diferenciação entre produto de referência e seus similares, e viabilizando uma futura distinção de possíveis produtos falsificados.

# CAPÍTULO V

## 5.0 - CONCLUSÃO

Comprovando a viabilidade deste trabalho, os resultados obtidos expressam uma classificação 100% correta, alocando todas as amostras em suas devidas classes. Sendo tal precisão oriunda do ótimo desempenho das técnicas investigativas (SIMCA e KNN) que classificaram perfeitamente todas as amostras do conjunto de teste, sem apresentar qualquer tipo de erro.

Alcançados com rapidez, baixo custo e sem geração de resíduos, os resultados demonstram que é possível classificar e diferenciar os mais diversos tipos de amostras de medicamentos, o que possibilitou uma distinção entre medicamentos de referência e similares. Desta forma, pode-se concluir que a utilização de espectroscopia NIR com métodos quimiométricos são capazes de fornecer informações valiosas sobre a distinção e classificação dos medicamentos a partir das suas características.

### 5.1 - Propostas Futuras

A partir das características do método apresentado nesta dissertação, como: ser não destrutivos, possibilitar a realização de análises *in situ* e *on line* em processos de produção, não gerar resíduos e, principalmente, possuir um baixo tempo de análise, pode-se recomendar uma infinidade de propostas que tomariam proveito destas características importantes.

De maneira sugestiva, como expansão do método aqui utilizado, novos estudos devem ser realizados em outros medicamentos que necessitam de um controle mais rígido na linha de produção, a exemplo dos fármacos de alta potência, onde quantidades ínfimas apresentam diferença significativa nos resultados terapêuticos, podendo trazer riscos à saúde humana.

Outra sugestão, esta de caráter complementar, diz respeito ao aumento da robustez, onde o método, agora, deverá não só classificar, mas também quantificar os constituintes presentes na amostra. Assim como proceder com a validação do método.

## REFERÊNCIAS

1. RUIZ, A.M.; OSORIO-DE-CASTRO, C.G.S. *Medicamentos: falando de qualidade*. Rio de Janeiro, Digital Ponto Gráfica, 2008.
2. MALUF, D.F.; PONTAROLO, R.; CORDEIRO, G.A.; NAGATA, N.; PERALTA-ZAMORA, P. *Determinação de sibutramina em formas farmacêuticas através de espectroscopia no infravermelho com refletância difusa e métodos de calibração multivariada*. *Química Nova*,33(3): 649-652, 2010.
3. MEDEIROS, V.M.; et al. *Screening analysis of river seston downstream of an effluent discharge point using near-infrared reflectance spectrometry and wavelet-based spectral region selection*. *Water Research*, 39: 3089–97, 2005.
4. RUISÁNCHEZ, A.P.; BOQUÉ, R.; RIUS, F.X. *Uncertainty of results in routine qualitative analysis*. *Trends in Analytical Chemistry*,22(10): 647-54, 2003.
5. ROSSIGNOLI, P.; PONTAROLO, R.; CORRER, C.J. *Determinação espectroscópica multivariada de glucosamina e condroitina em formulações farmacêuticas*. *Química Nova*, 31(6): 1285-89, 2008.
6. MANTANUS, J.; et al. *Moisture content determination of pharmaceutical pellets by near infrared spectroscopy: Method development and validation*. *Analytica Chimica Acta*, 642: 186–92, 2009.
7. GIL, E.S. *Controle físico-químico de qualidade de medicamentos*. 3 ed. São Paulo, Pharmabooks, 2010.
8. ANVISA. *RDC nº 17, de 16 de abril de 2010*.
9. GOMES, M.J.V.M.; REIS, A.M.M. *Ciências Farmacêuticas: Uma abordagem em farmácia hospitalar*. 1 ed. São Paulo, Atheneu, 2006.
10. ORGANIZAÇÃO MUNDIAL DE SAÚDE. *The importance of Pharmacovigilance: Safety monitoring of medicinal products*. 2002.

11. AMES, J.; SOUZA, D.Z. *Falsificação de medicamentos no Brasil*. Revista de Saúde Pública. 46(1):154-9, 2012.
12. ORGANIZAÇÃO MUNDIAL DE SAÚDE. *Medicamentos falsificados: Diretrizes para desenvolvimento de medidas de combate a medicamentos falsificados*. Brasília, 2005.
13. RODIONOVA, O.Y.; et al. *NIR spectrometry for counterfeit drug detection: a feasibility study*. Analytica Chimica Acta, 549(1):151-158, 2005.
14. NEWTON, P.N.; GREEN, M.D.; FERNÁNDEZ, F.M.; DAY, N.P.J.; WHITE, N.J. *Counterfeit anti-infective drugs*. The Lancet Infectious Diseases, 6(9):602-613, 2006.
15. NOGUEIRA, E.; NETO, G.V. *Falsificação de medicamentos e a lei n. 11.903/09: aspectos legais e principais implicações*. Revista de Direito Sanitário, 12(2):112-139, 2011.
16. AGÊNCIA NACIONAL DE VIGILÂNCIA SANITÁRIA. *Mais de 170 toneladas de medicamentos são apreendidas*. Disponível em: <<http://www.anvisa.gov.br/divulga/noticias/2009/160409.html>>, Acessado em 28 de janeiro de 2013.
17. FEDERAÇÃO NACIONAL DOS FARMACÊUTICOS. *Anvisa lança campanha para orientar sobre medicamento falsificado*. Disponível em: <<http://www.fenafar.org.br/portal/medicamentos/62-medicamentos/492-anvisa-lanca-campanha-para-orientar-sobre-medicamento-falsificado.html>>, Acessado em 14 de fevereiro de 2013.
18. WEINERT, P.L.; PEZZA, L.; PEZZA, H.R. *A simplified reflectometric method for the rapid determination of dipyrone in pharmaceutical formulations*. Journal of the Brazilian Chemical Society, 18(4):846, 2007.
19. SILVA, P. *Farmacologia*. 7 ed. Rio de Janeiro, Guanabara Koogan, 2006.
20. BRUNTON, L.L.; LAZO, J.S.; PARKER, K.L. *Goodman & Gilman: As Bases Farmacológicas da Terapêutica*. 11 ed. Rio de Janeiro, McGraw-Hill, 2006.
21. *Farmacopeia Brasileira*, 5 ed., volume 2, Brasília, 2010.
22. SANCHES, F.A.C. *Desenvolvimento de uma Metodologia para o Controle de Qualidade de Medicamento Injetável, sem Violação da Ampola, usando Espectroscopia NIR e Técnicas*

*Quimiométricas*. João Pessoa, Programa de Pós-Graduação em Química - UFPB, 2009. Dissertação de mestrado.

23. HAMERSCHLAK, N.; CAVALCANTI, A.B. *Neutropenia, agranulocytosis and dipyrrone*. São Paulo Medical Journal, 123(5): 247-249, 2005.

24. IBANEZ, L.; *et al.* *Agranulocytosis associated with dipyrrone (metamizol)*. European Journal of Clinical Pharmacology, 60(11): 821-829, 2005.

25. EDWARDS, J.E.; *et al.* *Single dose dipyrrone for acute postoperative pain (Cochrane Review)*. The Cochrane Library, 2, Oxford Update Software, 2005.

26. BOEHRINGER INGELHEIM. *Anador tenta ganhar mercado do Tylenol*. Disponível em: <[http://www.boehringer-ingelheim.com.br/conteudo\\_imprensa\\_texto.asp?conteudo=12&texto=731](http://www.boehringer-ingelheim.com.br/conteudo_imprensa_texto.asp?conteudo=12&texto=731)>, Acessado em 28 de janeiro de 2013.

27. FOLHA DE SÃO PAULO. *Pesquisa revela remédios mais consumidos pelos brasileiros*. Disponível em: <<http://www1.folha.uol.com.br/equilibrioesaude/1183441-pesquisa-revela-remedios-mais-consumidos-pelos-brasileiros.shtml>>, Acessado em 22 de abril de 2013.

28. *British Pharmacopoeia*, volume I e II, 2009.

29. CHOU, R, PETERSON, K., HELFAND, M. *Comparative Efficacy and Safety of Skeletal Muscle Relaxants for Spasticity and Musculoskeletal Conditions: A Systematic Review*. Journal of Pain and Symptom Management, 28(2), 2004.

30. MARIA, C.A.B.; MOREIRA, R.F.A. *Caféina: Revisão sobre métodos de Análise*, Química Nova, 30(1): 99-105, 2007.

31. DI NEZIO, M.S. *A Home-made hybrid system for the simultaneous determination of ergotamine, dipyrrone and caffeine in pharmaceutical preparations*. Journal of the Brazilian Chemical Society. 18(7): 1439-42, 2007.

32. *European Pharmacopoeia*, volume II, 5 ed, 2005. Disponível em: <<http://lib.njutc.edu.cn/yaodian/ep/EP5.0/index.htm>>, Acessado em 28 de janeiro de 2013.

33. *United States Pharmacopoeia 30*, NF 25, 2007. CD-ROM.

34. DOS SANTOS, W.T.P.; GIMENES, D.T.; DE ALMEIDA, E.G.N.; EIRAS, S.P.; ALBUQUERQUE, Y.D.T.; RICHTER, E.M. *Simple flow injection amperometric system for simultaneous determination of dipyrone and paracetamol in pharmaceutical formulations*. Journal of the Brazilian Chemical Society, 20(7):1249-1255, 2009.
35. AMÉRICO, M.A.; MOSSIN, S.A.G.; NISHIYAMA, P. *Perfil de fármacos por espectrofotometria no ultravioleta*. Revista Brasileira de Análises Clínicas, 40(4): 257-59, 2008.
36. DE SOUZA, J.S; FERRÃO, M.F. *Aplicações da espectroscopia no infravermelho no controle de qualidade de medicamentos contendo diclofenaco de potássio. Parte I : Dosagem por regressão multivariada*. Revista Brasileira de Ciências Farmacêuticas, 42(3): 437-45, 2006.
37. PEREIRA, A.V., PENCKOWSKI, L.; VOSGERAU, M.; SASSÁ, M.F. *Determinação espectrofotométrica de dipirona em produtos farmacêuticos por Injeção em fluxo pela geração de íons tri-iodeto*. Química Nova, 28(4): 553-557, 2002.
38. BONATO, P.S.; JABOR, V.A.P. *Análise enantiosseletiva de fármacos: Contribuições da cromatografia líquida de alta eficiência e eletroforese capilar*. Química Nova, 28(4): 683-691, 2005.
39. SKOOG, D.A.; HOLLER, F.J.; NIEMAN, T.A. *Princípios de análise instrumental*. 5. ed. Porto Alegre: Bookman Editora, 2002.
40. BLANCO, M.; ALCALÁ, M. *Simultaneous quantitation of five active principles in a pharmaceutical preparation: Development and validation of a near infrared spectroscopic method*. European Journal of Pharmaceutical Sciences, 27: 280–286, 2006.
41. DINC, E.; ONUR, F. *Application of a new spectrophotometric method for the analysis of a ternary mixture containing metamizol, paracetamol and caffeine in tablets*. Analytica Chimica Acta, 359: 93-106, 1998.

42. Scotter, C.N.G. *No-destructive spectroscopic techniques for the measurement of food quality*. Trends in Food. Science & Technology. 8(9): 285-292, 1997.
43. CORDEIRO, G.A. *Desenvolvimento de Metodologias Espectroscópicas Multivariadas para a Quantificação de Fármacos em Formas Farmacêuticas*. Curitiba, Programa de Pós-Graduação em Química - UFPR, 2006. Dissertação de mestrado.
44. RUFINO, J.L. *Desenvolvimento de metodologias analíticas para controle de qualidade de fármacos, utilizando-se técnicas espectroscópicas (NIR e MID) e processos de calibração multivariada (PLS)*. Curitiba, Programa de Pós-Graduação em Química - UFPR, 2004. Dissertação de mestrado.
45. ESTEBAN-DIEZ, I.; GONZALEZ-SAIZ, J.M.; SAENZ-GONZALEZ, C.; PIZARRO, C. *Coffee varietal differentiation based on near infrared spectroscopy*. Talanta, 71: 221-229, 2007.
46. BLANCO, M.; VILLAESCUSA, V.; *Use of NIR spectroscopy in the production of modified industrial resins*. Talanta, 71: 1333-38, 2007.
47. POPPI, R.J.; COSTA FILHO, P.A. *Aplicação de algoritmos genéticos na seleção de variáveis em espectroscopia no Infravermelho médio - Determinação simultânea de glicose, maltose e frutose*. Química Nova, 25(1): 46-52, 2002.
48. PARISOTTO, G.; SOUZA, J.S.; FERRAO, M.F.; FURTADO, J.C.; MOLZ, R.F. *Análise Exploratória Aplicada no Estudo de Medicamentos Contendo Piroxicam*. Revista Brasileira de Ciências Farmacêuticas, 41 (4): 499-505, 2005.
49. SIMÕES, S.S. *Desenvolvimento de métodos validados para a determinação de captopril usando espectrometria NIRR e calibração multivariada*. João Pessoa, Programa de Pós-Graduação em Química - UFPB, 2008. Tese de doutorado.
50. PASQUINI, C.; *Near infrared spectroscopy: Fundamentals, practical aspects and analytical applications*. Journal of the Brazilian Chemical Society, 14(2): 198-219, 2003.

51. DA SILVA, B.D. *Desenvolvimento de Metodologias Espectroscópicas Multivariadas para Quantificação de Isoniazida e Rifampicina em Medicamentos*. Curitiba, Programa de Pós-Graduação em Química - UFPR, 2008. Dissertação de mestrado.
52. MASSART, D.L.; VANDEGINSTE, B.G.M.; BUYDENS, S.J.; LEWI, P.J.; SMEYERS-VERBEKE, J. *Handbook of Chemometrics and Qualimetrics: Parte B*, Amsterdam, Elsevier, 1997.
53. SABIN, J.G.; FERRÃO, M.F.; FURTADO, J.C. *Análise multivariada aplicada na identificação de fármacos antidepressivos. Parte II: Análise por componentes principais (PCA) e o método de classificação SIMCA*. Revista Brasileira de Ciências Farmacêuticas. 40(3): 387-96, 2004.
54. FERREIRA, M.M.C; ANTUNES, A.M.; MELGO, M.S.; VOLPE, P.L.O. *Quimiometria I: Calibração multivariada, um tutorial*. Química Nova. 22(5): 724-31, 1999.
55. BEEBE, K. R.; PELL, R.J.; SEASHOLTZ, B. *Chemometrics – a practical guide*. New York: Wiley, 1998.
56. MOREIRA, A.B. *Desenvolvimento de métodos analíticos para a determinação de ácido acetilsalicílico, paracetamol e cafeína em matriz sólida por espectroscopia de fluorescência*. São Paulo, UNICAMP, 2005. Tese de doutorado.
57. PARREIRA, T.F. *Utilização de métodos quimiométricos em dados de natureza multivariada*. Campinas, Programa de pós-graduação em Química - UNICAMP, 2003. Dissertação de mestrado.
58. VÉRAS, G.; et al. *Classificação de biodiesel na região do visível*. Química Nova. 35(2), 315-318, 2012.
59. EL-HAGRASY, A.S.; DELGADO-LOPEZ, M.; DRENNEN III, J.K. *A Process Analytical Technology approach to near-infrared process control of pharmaceutical powder blending: Part II: Qualitative Near-Infrared Models for Prediction of Blend Homogeneity*. Journal of Pharmaceutical Science. 95(2):422-434, 2006.

60. SENA, M.M.; FREITAS, C.B.; SILVA, L.C.; PÉREZ, C.N.; DE PAULA, Y.O. *Determinação espectrofotométrica simultânea de paracetamol e ibuprofeno em formulações farmacêuticas usando calibração multivariada*. Química Nova. 30(1): 75-79, 2007.
61. PASCHOAL, L.R.; FERREIRA, W.A.; PRADO, M.R.D; VILELA, A.P.O. *Aplicação do método da espectrofotometria de derivadas na identificação e doseamento simultâneo de sistemas multicomponentes*. Revista Brasileira de Ciências Farmacêuticas.39(1): 105-13, 2003.
62. WEINERT, P.L.; PEZZA, L.; PEZZA, H.R. *A simplified reflectometric method for the rapid determination of dipyrone in pharmaceutical formulations*. Journal of the Brazilian Chemical Society. 18(4): 846-854, 2007.
63. VIEIRA, R.G.; et al. *Chemical pattern of brazilian apples. A chemometric approach based on the Fuji and Gala varieties*. Ciência e Tecnologia de Alimentos. 31(2): 418-426, 2011.
64. HELFER, G.A.; FERRÃO, M.F.; FERREIRA, C.V.; HERMES, N. *Aplicação de métodos de análise multivariada no controle qualitativo de essências alimentícias empregando espectroscopia no infravermelho médio*. Ciência e Tecnologia de Alimentos. 26(4): 779-786, 2006.
65. VERAS, G.; GOMES, A.A.; SILVA, A.C.; BRITO, A.L.B; ALMEIDA, P.B.A; MEDEIROS, E.P. *Classification of biodiesel using NIR spectrometry and multivariate techniques*. Talanta. 83: 565-568, 2010.
66. SILVA, G.W.B.; GOMES, A.A.; SILVA, P.; COSTA, G.B.; FERNANDES, D.D.S.; FONTES, M.M.; VERAS, G. *Biodiesel/Diesel Blends Classification with Respect to Base Oil Using NIR Spectrometry and Chemometrics Tools*. Journal of the American Oil Chemists' Society. 89(7): 1165-71, 2012.
67. MOREIRA, E.D.T.; PONTES, M.J.C.; GALVAO, R.K.H.; ARAÚJO, M.C.U. *Near infrared reflectance spectrometry classification of cigarettes using the successive projections algorithm for variable selection*. Talanta. 79: 1260–1264, 2009.

68. SANTOS, F.C. *Variações do método KNN e suas aplicações na classificação automática de textos*. Goiás, Programa de Pós-Graduação do Instituto de Informática - UFG, 2009. Dissertação de mestrado.

# APÊNDICE A

## Classification of Tablets containing Dipyrone, Caffeine and Orphenadrine by Near Infrared Spectroscopy and Chemometric Tools

Carlos Alan D. Melo,<sup>a</sup> Priscila Silva,<sup>b</sup> Adriano de Araújo Gomes,<sup>c</sup>  
David Douglas S. Fernandes,<sup>b</sup> Germano Vêras<sup>b</sup> and Ana Claudia D. Medeiros<sup>\*,a</sup>

<sup>a</sup>Laboratório de Desenvolvimento e Ensaio em Medicamentos (LABDEM),  
Centro de Ciências Biológicas e da Saúde and <sup>b</sup>Laboratório de Química Analítica e Quimiometria,  
Centro de Ciências e Tecnologia, Universidade Estadual da Paraíba,  
58429-500 Campina Grande-PB, Brazil

<sup>c</sup>Departamento de Química, Centro de Ciências Exatas e da Natureza (CCEN),  
Universidade Federal da Paraíba, 58051-970 João Pessoa-PB, Brazil

O objetivo deste estudo foi classificar amostras de comprimidos contendo dipirone, cafeína e orfenadrina usando espectroscopia no infravermelho próximo e técnicas quimiométricas. O conjunto de dados foi de 300 espectros de amostras de três comprimidos por lote e quatro diferentes produtores. O pré-processamento foi realizado pelo algoritmo Savitzky-Golay com primeira derivada, janela de 17 pontos e polinômio de segunda ordem. A classificação dos comprimidos foi conduzida usando modelos quimiométricos baseados na análise de componentes principais (PCA), modelagem independente flexível por analogias de classes (SIMCA), algoritmo genético- (GA-LDA) e algoritmo das projeções sucessivas-análise discriminante linear (SPA-LDA). Pela análise PCA, observou-se agrupamentos para cada conjunto de comprimidos. Para o modelo SIMCA, utilizou-se 15 e 30 medidas espectrais para o conjunto de treinamento dos medicamentos similares e de referência, respectivamente. Para o modelo GA-LDA, utilizou-se 12 variáveis, enquanto que o modelo SPA-LDA selecionou somente dois comprimentos de onda, 1572 e 1933 nm. Os modelos classificaram corretamente todas as amostras. A metodologia permitiu uma classificação rápida e não destrutiva das amostras e sem necessidade de determinações analíticas convencionais.

The objective of this study was to classify samples of tablets containing dipyrone, caffeine and orphenadrine using near infrared (NIR) spectroscopy and chemometric techniques. The data set had 300 spectra of samples from three tablets *per* batch and four different manufacturers. The pre-processing was accomplished by Savitzky-Golay algorithm with the first derivative, window with 17 points and second-order polynomial. The tablet classification was performed using chemometric models based on principal component analysis (PCA), soft independent modeling of class analogies (SIMCA), genetic algorithm- (GA-LDA) and successive projection algorithm-linear discriminant analysis (SPA-LDA). For PCA analysis, clusters were observed for each group of tablets. The SIMCA model was built using 15 and 30 spectral measures for the training set of similar drugs and reference drugs, respectively. The GA-LDA model used 12 variables, whereas SPA-LDA selected only two wavelengths, 1572 and 1933 nm. The methodology allowed a quick and non-destructive classification of the tablets and without the need for conventional analytical determinations.

**Keywords:** chemometrics, drug screening, near infrared spectroscopy, pharmaceuticals, quality control

### Introduction

Counterfeiting of medicines is part of a broader process involving the distribution of drugs that do not meet the

standards of quality, safety and efficacy. According to the World Health Organization (WHO), spurious/falsely-labeled/falsified/counterfeit (SFFC) medicines are those wrongly labeled, deliberately or misleading with respect to their identity or source. Tampering includes reference, similar generic products and may include

\*e-mail: anaclaudia@uepb.edu.br

products with correct, incorrect, insufficient or missing and/or with fake packaging active ingredients.<sup>1-3</sup>

In most of the developed countries with effective systems of regulation and market control (i.e., United States, Australia, Canada, Japan, New Zealand and most European Union countries), the incidence of SFFC drugs is low, less than 1% of the market, according to estimates from the countries concerned.<sup>3,4</sup> In contrast, the highest incidence occurs in regions where the regulatory and supervisory systems are weak. In some developing countries, the SFFC drugs reach alarming 25% of the local market, which represents about 10% of the global pharmaceutical market.<sup>4,5</sup>

Due to high amount of fraud and risk posed by these drugs, pharmaceutical products (medicines, cosmetics and related) have been subjected to safety requirements and quality assurance through technical regulations set by government authorities. These regulations are supported by voluntary technical activities that contribute to the quality of products, such as ISO 9001 standard.<sup>5</sup>

In Brazil, the inspection agencies seized SFFC drugs, among them contraceptives, antibiotics and painkillers, more often containing dipyrone in the composition.<sup>6</sup> These drugs continue to be subject to forgery by their popularity and acceptance due to the high levels of marketing and consumption.

According to information publicly provided by the Boehringer Ingelheim Company, three main substances of the class of painkillers, that although being different molecules have the same purpose of pain sedation, holds 95% of the Brazilian market. Drugs with dipyrone lead with 39%, followed by paracetamol with 30% and aspirin with 26% of the market.

Dipyrone (sodium salt of 1-phenyl-2,3-dimethyl-4-methylaminomethane sulphonate-5-pyrazolone) is a therapeutic agent commonly used as analgesic, antipyretic and antispasmodic in various pharmaceutical formulations. Restricted in some countries like the United States, dipyrone is commercially available in Brazil mainly due to its strong analgesic effect and relatively low cost.<sup>6-10</sup> In some drugs, dipyrone is present in association with orphenadrine [(*RS*)-*N,N*-dimethyl-2-[(2-methylphenyl)phenylmethoxy]ethanamine dihydrogen 2-hydroxypropane-1,2,3-tricarboxylate], analgesic with weak antihistaminic property, useful for relieving pain associated with traumatic or inflammatory muscle contractions,<sup>11</sup> and with caffeine (1,3,7-trimethylxanthine), an alkaloid responsible for promoting increased alert capacity and reduced fatigue,<sup>10,12-14</sup> and further has central stimulating action and may exacerbate the excitatory symptoms of both previous drugs.

According to Santos *et al.*,<sup>7</sup> in samples with two or more active ingredients present in a single formulation, the quantification must be performed by high performance liquid chromatography (HPLC) with UV detection. Although presenting accuracy and precision, the official method in question is also characterized by being laborious and expensive, often requiring pre-treatment of samples, ultra-pure reagents and specialized operators, and sample degradation and production of organic waste harmful to the environment take place.<sup>15</sup>

However, numerous other analytical techniques have been proposed for the analysis of drugs, among which the near infrared (NIR) spectroscopy,<sup>16-20</sup> a rapid and non-destructive technique based on the absorption of electromagnetic radiation between 14000 and 4000  $\text{cm}^{-1}$  (780 to 2500 nm).

The use of multiple analytical channels in the acquisition of chemical information for samples, as NIR spectroscopy, may be adequately exploited using multivariate analysis, extracting as much information as possible of data sets. In this context, pattern recognition techniques and NIR spectroscopy have been reported for the development of screening methodology for quality control of various matrices as fuel,<sup>20-23</sup> drink<sup>24</sup> and food.<sup>25</sup>

Soft independent modeling of class analogies (SIMCA)<sup>26</sup> is a well known supervised pattern recognition method that uses principal component analysis (PCA) to model the hyperspace of each class. The PCA method promotes compression of a large data set and the variance is concentrated in few variables called principal components, i.e., for a set of  $k$  objects measured in sensors  $j$  generating the matrix  $\mathbf{X}_{kj}$ , PCA reduced the matrix  $\mathbf{X}$  into a product of two other arrays of low dimensionality  $\mathbf{T}_{kxA}$  (scores) and  $\mathbf{L}_{jxA}$  (loadings). The new variables in  $\mathbf{T}$  present the advantage of being mutually orthogonal and  $A$  represents the number of new variables considered to be significant for the model of each class. The classification of new samples is carried out by means of an  $F$ -test at a given significance level.

The new variables present advantage of being mutually orthogonal, allowing the use of all spectral information in the construction of the SIMCA model, known as full spectrum method. This characteristic permits the detection of anomalous samples or outliers, present in the data set.<sup>26,27</sup>

However, when employing full spectrum in the construction of mathematic models, many variables are redundant and/or non-informative, and their inclusion may affect the quality of the final model. Currently, a well-succeeded alternative to overcome this drawback is the use of variable selection techniques.<sup>28-31</sup>