



Ronaldo Pedro da Silva

**Ciência de Dados & Química do Estado Sólido:
Uma Plataforma para Competitividade da Indústria
Farmoquímica e Farmacêutica em Mercado Emergentes**

Tese de Doutorado

Tese apresentada ao Programa de Pós-graduação em Engenharia de Produção da PUC-Rio como requisito parcial para obtenção do grau de Doutor em Engenharia de Produção.

Orientador: Prof. Eugenio Kahn Epprecht
Co-orientador: Prof. Roberto Ribeiro de Avillez

Rio de Janeiro
Março de 2018



Ronaldo Pedro da Silva

**Ciência de Dados & Química do Estado Sólido: Uma
Plataforma para Competitividade da Indústria
Farmaquímica e Farmacêutica em Mercado Emergentes**

Tese apresentada como requisito parcial para obtenção do grau de
Doutor pelo Programa de Pós-Graduação em Engenharia de Produção da
PUC-Rio. Aprovada pela Comissão Examinadora abaixo assinada.

Prof. Eugenio Kahn Epprecht

Orientador

Departamento de Engenharia Industrial - PUC-Rio

Prof. Roberto Ribeiro de Avillez

Co-orientador

Departamento de Engenharia dos Materiais - PUC-Rio

Prof. Antônio Márcio Tavares Thomé

Departamento de Engenharia Industrial - PUC-Rio

Prof. Fernando Antonio Ribeiro Serra

Universidade Nove de Julho – UNINOVE

Prof. Maria Simone de Menezes Alencar

Universidade Federal do Estado do Rio de Janeiro – UNIRIO

Prof. Lorenzo do Canto Visentin

Nanobusiness – Informação e Inovação

Prof. Márcio da Silveira Carvalho

Coordenador Setorial do Centro

Técnico Científico - PUC-Rio

Rio de Janeiro, 19 de março de 2018

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

Ronaldo Pedro da Silva

Graduou-se em Administração de Empresas pelo IAG/PUC-Rio em 2004. Possui especialização técnica em difração de raios-X por policristais e mestrado em Engenharia de Produção pelo Departamento de Engenharia Industrial da PUC-Rio em 2007. Possui experiência de cerca de 30 anos em ambiente empresarial, de pesquisa industrial e pesquisa científica associada à concepção e gestão de desenvolvimento de inovações em processos, produtos e serviços. Foi consultor na área de nanotecnologia da coordenação-geral de micro e nanotecnologias do Ministério de Ciência, Tecnologia e Inovação (MCTI); Secretaria de Inovação do Ministério do Desenvolvimento, Indústria e Comércio Exterior (MDIC) e da Organização das Nações Unidas para a Educação, a Ciência e a Cultura (UNESCO). É sócio-fundador da NanoBusiness Informação e Inovação, empresa *spin-off* do Departamento de Engenharia Química e de Materiais da PUC-Rio. Atualmente é professor da Universidade Federal do Rio de Janeiro (UFRJ) Campus Professor Geraldo Cidade.

Ficha Catalográfica

Silva, Ronaldo Pedro da

Ciência de dados & química do estado sólido: uma plataforma para competitividade da indústria farmoquímica e farmacêutica em mercado emergentes / Ronaldo Pedro da Silva; orientador: Eugenio Kahn Epprecht; co-orientador: Roberto Ribeiro de Avillez. – 2018.

229 f.; 30 cm

Tese (doutorado) – Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Industrial, 2018.

Inclui bibliografia

1. Engenharia Industrial – Teses. 2. Química do estado sólido. 3. Ciência dos dados. 4. Text mining. 5. Relações Fuzzy. 6. Gestão de P&D. I. Epprecht, Eugenio Kahn. II. Avillez, Roberto Ribeiro de. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Industrial. IV. Título.

CDD: 658.5

Dedico este trabalho a minha querida mãe, Tereza Vicença Viana, PhD com pós-doutorado na ciência em ser feliz. Tenho uma gratidão imensa e eterna por me conceber e deixar como herança a capacidade e força em sair do zero e realizar sonhos. Sua bibliografia, metodologia e resultados, ou seja, oferecer mais do que tinha para deixar sua família e amigos felizes, foi sua melhor tese de vida!

Agradecimentos

A Deus, pela existência, oportunidades e saúde;

Ao meu orientador, Professor Eugenio Kahn Epprecht, pela objetividade, franqueza, parceria e, sobretudo, pela confiança depositada na concepção desta tese de doutorado. Tenho muito a agradecer por confiar no meu desempenho no programa de pós-graduação do Departamento de Engenharia da PUC-Rio, desde os meus esforços para ingressar no mestrado.

Ao meu co-orientador, Professor Roberto Ribeiro de Avillez, pela flexibilidade, oportunidades de estudos e aprendizado ao longo destes 22 anos de atividade produtiva, o que contribuiu, inclusive, com as minhas iniciativas como pesquisador-empresendedor;

Essa tese não seria possível sem o apoio da UFRJ, PUC-Rio e Inmetro, essa última instituição representada pelo Dr. Oleksii Kuznetsov. Agradeço também ao apoio financeiro do CNPq e Faperj e, em especial, ao apoio financeiro dos sócios fundadores da NanoBusiness, particularmente, meu sócio e sogro Jesus Lago Ramirez;

Estendo os meus sinceros agradecimentos à toda a equipe NanoBusiness e aqueles que direta ou indiretamente colaboraram com a construção do conhecimento acumulado, o qual viabilizou essa tese, desde a criação da empresa em dezembro de 2008. Aqueles que me conhecem sabem que a minha maior motivação para a realização do doutorado está associada a expectativa da aplicação dos seus resultados. Assim, espero que os artigos e patentes potenciais gerados possam contribuir para a sustentabilidade e o crescimento do segmento de química de estado sólido, bem como empresa e sua equipe, diante do atual cenário de crise financeira e de valores em que se encontra o nosso país.

Sabemos que o desenvolvimento científico e tecnológico depende da postura colaborativa de competências específicas e multidisciplinares. Deste modo, reforço meus agradecimentos pelo apoio técnico e aprendizado obtido por meio dos mestres e doutores Adriano Koshiyama, Daniel Giacometti, Lorenzo Visentin e Luciana Piovesan.

Desde criança construí na minha mente um pai ou amigo imaginário e aprendi ao longo desses 49 anos de idade a extrair o melhor das pessoas. Então, esse pai ou amigo imaginário resiste até hoje com pedaços de muitas pessoas que tenho grande respeito e admiração. Uma pequena amostra do meu amigo imaginário é formada pelo meu querido irmão Marzin, Sr. Amastor (hoje com 102 anos de idade), Célio Esteves Barbosa, José Carlos Pinto, Ricardo Henrique Leal, Roberto de Avillez, Fernando Rizzo, Fernando Galembeck, José Carlos Lucena, Jesus Lago Ramirez, entre outros chefes, patrões, professores, pesquisadores, colegas de trabalho e membros da família que tive a honra e oportunidade de trabalhar e conviver. Meu coração se enche de felicidade ao lembrar e reconhecer suas orientações, conselhos e apoio na construção da minha formação profissional e

peçoal. Hoje, depois de inúmeras dificuldades - comum a todos ao longo da vida - posso garantir que Deus existe e é um pai ou amigo real e não imaginário. Quero, portanto, agradecer a Deus e a meus amigos reais.

Escrever aqui uma frase de agradecimento à minha família está longe de ser uma tarefa trivial, pois lembranças surgem, emocionam e enchem o meu coração de alegrias e saudades. São tantas pessoas a agradecer, em especial a minha esposa Elvira, que foi incansável na sua compreensão, incentivo e apoio, e a meus queridos filhos Júlia, Pedro e Luísa, pela paciência e compreensão nos momentos que deixamos de estar juntos devido não apenas ao doutorado. Muito mais complexo do que fazer um doutorado e ser empreendedor é viver sem vocês família! Muito obrigado!

Resumo

da Silva, Ronaldo Pedro; Epprecht, Eugenio Kahn (Orientador); Avillez, Roberto Ribeiro de (Co-orientador). **Ciência de Dados & Química do Estado Sólido: Uma Plataforma para Competitividade da Indústria Farmoquímica e Farmacêutica em Mercado Emergentes**. Rio de Janeiro, 2018. 229p. Tese de Doutorado – Departamento de Engenharia Industrial, Pontifícia Universidade Católica do Rio de Janeiro.

A área de química do estado sólido ocupa uma posição cada vez mais importante nas atividades de pesquisa e desenvolvimento farmacêuticas. A compreensão das propriedades do estado sólido de um insumo farmacêutico ativo (IFA) mostra-se crítica no desenvolvimento de formulações em função de seus impactos na biodisponibilidade e solubilidade dos fármacos, sendo essencial para garantir o benefício terapêutico, otimizar o desenvolvimento e garantir a proteção da propriedade intelectual. Esta tese investiga indicadores científicos e tecnológicos na área de química do estado sólido utilizando ferramentas de ciência dos dados a partir de publicações científicas e depósitos de patentes, visando contribuir para o aumento da competitividade da indústria farmoquímica e farmacêutica brasileira e de outros mercados emergentes. A partir da utilização de ferramentas de ciência dos dados é proposta uma metodologia baseada em técnicas de *text mining* associadas a relações *fuzzy*. Essa metodologia de identificação de competências específicas aplicada na área de química do estado sólido tem como estudo de caso a descoberta de uma nova forma polimórfica para o IFA acetato de dexametasona. Os resultados revelam que existem competências científicas em química do estado sólido no Brasil. Contudo, quando comparada com a interação universidade-empresa mundial, a indústria farmoquímica e farmacêutica local perde em estágio de competitividade e desenvolvimento. Por outro lado, os resultados demonstram a robustez da metodologia e sua capacidade de identificar pesquisadores em área específicas, oferecendo soluções para apoio a tomada de decisão e identificação de pesquisadores relevantes para o desenvolvimento do setor farmoquímico e farmacêutico.

Palavras-chave

Química do estado sólido; Ciência dos dados; Text Mining; Relações Fuzzy; gestão de P&D; Interação universidade-empresa; Parceria Público-Privado.

Abstract

da Silva, Ronaldo Pedro; Epprecht, Eugenio Kahn (Advisor); Avillez, Roberto Ribeiro de (Co-advisor). **Data Science & Solid State Chemistry: a platform for the competitiveness of the pharmaceutical industry in emerging markets**. Rio de Janeiro, 2018. 229p. Tese de Doutorado – Departamento de Engenharia Industrial, Pontifícia Universidade Católica do Rio de Janeiro.

The solid-state chemistry area has received increased attention in the pharmaceutical research and development activities. The comprehension of the solid-state properties of an active pharmaceutical ingredient (API) is critical in the development of formulations due to their impact on the bioavailability and solubility of the final drug, being essential to ensure therapeutic benefit, optimize development and allow a proper intellectual property protection. This research investigates science and technology indicators in the solid-state chemistry area using data science tools applied to scientific publications and patent documents, aiming to contribute to the increase of the competitiveness of the pharmaceutical industry in Brazil and in other emerging markets. Through data science tools, a methodology based on text mining techniques associated to fuzzy relations is proposed. This methodology for identifying specific competencies is applied in the solid-state chemistry area exploring a case study of the discovery of a new polymorphic form of the API dexamethasone acetate. The results reveal the existence of scientific competencies in solid-state chemistry in Brazil. However, when compared to the global university-company interaction, the local pharmaceutical industry shows a lower stage of competitiveness and development. On the other hand, the results indicates the robustness of the methodology and its ability to identify researchers in specific areas, offering solutions to support the decision making and identification of researchers relevant to the development of the pharmaceutical sector.

Keywords

Solid State Chemistry; Data Science; Text Mining; Fuzzy Relation; R & D management; University-company interaction; Public-Private Partnership

Sumário

1 Introdução.....	15
1.1. Objetivos.....	20
2 Pesquisa e Desenvolvimento na Indústria Farmacêutica e a Química do Estado Sólido.....	22
2.1. Dinâmicas da P&D na Indústria Farmacêutica.....	22
2.2. Química do Estado Sólido	28
3 Ciência dos dados na identificação de competências científicas.....	46
3.1. Extrair informações do conhecimento científico e tecnológico.....	48
3.2. Uso de indicadores bibliométricos de C&T	50
3.3. Identificação de lideranças e competências científicas	52
4 Metodologia.....	56
4.1. Etapa 1: Procedimento geral para Elaboração de indicadores de C&T.....	57
4.2. Etapa 2: Modelo de Correlação entre palavras-chave e tópicos.....	65
4.3. Etapa 3: Modelo de identificação de competências específicas – NB-RQ.....	72
5 Resultados e discussão	77
5.1. Indicadores científicos e tecnológicos em química do estado sólido no mundo	77
5.2. Indicadores científicos e tecnológicos em química do estado sólido no Brasil.....	93
5.3. Identificação de competências: estudo de caso acetato de dexametasona....	103
5.3.1 Identificação de lideranças e competências científicas no Mundo e no Brasil.....	104
5.3.1.1 Identificação de lideranças e competências científicas em volume de publicações no mundo	104
5.3.1.2 Identificação de lideranças e competências científicas em validade das publicações no mundo.....	106
5.3.1.3 Identificação de lideranças e competências científicas em valor acadêmico no mundo	107
5.3.1.4 Identificação de lideranças e competências científicas em valor industrial no mundo.....	108
5.3.1.5 Identificação de lideranças e competências científicas no mundo com valor para o caso acetato de dexametasona	110
5.3.1.6 Identificação de lideranças e competências científicas no mundo – NB-RQ: caso acetato de dexametasona.....	111
5.3.1.7 Identificação de lideranças e competências científicas no Brasil – NB-RQ: caso acetato de dexametasona.....	114
6 Conclusões.....	116
6.1. Química do Estado Sólido	116
6.2. Da ciência ao mercado.....	117

6.3. Ciência de Dados (Método NB-RQ)	119
6.4. Trabalhos e projetos realizados em andamento	120
7 Referências bibliográficas	121
Anexos	126

Lista de quadros

Quadro 1 - Técnicas analíticas comumente utilizadas para caracterização do estado sólido	37
Quadro 2 – Definições de termos comumente utilizados	49
Quadro 3 - Tópicos selecionados em química do estado sólido na indústria farmacêutica.....	66
Quadro 4 - Indicadores de perfil dos principais países em pesquisa em química do estado sólido	78
Quadro 5 - Principais depositantes de patente em química do estado sólido e país de origem.....	92
Quadro 6 – Principais depositantes de patente em química do estado sólido no Brasil.....	102
Quadro 7 – Instituições de pesquisa e empresa com vínculo com os pesquisadores Top10 NB-RQ – Mundo	113

Lista de figuras

Figura 1 - Mercado farmacêutico brasileiro (U\$).....	18
Figura 2 - Representação esquemática dos vários tipos de formas sólidas	30
Figura 3 - Linha do tempo de eventos envolvendo impactos na indústria farmacêutica devido a formas cristalinas de fármacos	32
Figura 4 - Propriedades que podem diferir entre formas cristalinas de um IFA ...	35
Figura 5 - Crescimento anual de registros indexados na base de dados Web of Science.....	47
Figura 6 - Etapas da metodologia proposta	56
Figura 7 - Tabela Base que contém as principais informações extraídas da publicação.....	65
Figura 8 - Evolução das publicações científicas em química do estado sólido (1997- 2016)	77
Figura 9 - Rede de colaboração dos principais países com publicações em química do estado sólido	79
Figura 10 - Evolução temporal das publicações em química do estado sólido por país	80
Figura 11 - Evolução temporal das categorias <i>Web of Science</i>	81
Figura 12 - Palavras-chave mais frequentes nas publicações em química do estado sólido	82
Figura 13 - Evolução temporal das palavras-chave mais frequentes nas publicações em química do estado sólido (1997-2016).....	83
Figura 14 - Evolução temporal das principais técnicas analíticas nas publicações em química do estado sólido (1997-2016).....	84
Figura 15 - Mapa de densidade da rede de palavras-chave mais frequentes.....	85
Figura 16 - Mapa de rede de co-ocorrência de palavras-chave com o termo cocrystals	86
Figura 17 - Principais organizações em publicações científicas na área de química do estado sólido	87
Figura 18 - Principais empresas em publicações científicas na área de química do estado sólido	88

Figura 19 - Patenteamento em química do estado sólido por ano de prioridade (1997-2015).....	90
Figura 20 - Principais países de prioridade dos documentos de patente	91
Figura 21 - Evolução das publicações científicas de autores brasileiros em química do estado sólido (1997- 2016)	93
Figura 22 - Distribuição das principais categorias de assunto das publicações em química do estado sólido	94
Figura 23 - Palavras-chave mais frequentes nas publicações em química do estado sólido de autores brasileiros	95
Figura 24 - Principais países com os quais o Brasil possui colaboração em publicações científicas em química do estado sólido	97
Figura 25 - Evolução do total de publicações de autores brasileiros e publicações com colaboração internacional	98
Figura 26 - Principais organizações em publicações científicas em química do estado sólido no Brasil.....	99
Figura 27 - evolução das principais organizações brasileiras em número de publicações em química do estado sólido (1997-2016).....	100
Figura 28 - Rede de colaboração entre as principais organizações brasileiras em publicações em química do estado sólido.....	101
Figura 29 – Pesquisadores Top10 pela perspectiva (<i>Vpub</i> *) Volume de Publicações - Mundo	105
Figura 30 – Pesquisadores Top10 pela perspectiva (<i>Vval</i> *) Validade das Publicações - Mundo	106
Figura 31 - Pesquisadores Top10 pela perspectiva (<i>Vacad</i> *) Valor Acadêmico - Mundo	108
Figura 32 - Pesquisadores Top10 pela perspectiva (<i>Vind</i> *) Valor Industrial – Mundo.....	109
Figura 33 - Pesquisadores Top10 pela perspectiva (<i>Vproj</i> *) Valor do Projeto – Mundo	110
Figura 34 - Pesquisadores Top10 pela perspectiva NB-RQ – Mundo	112
Figura 35 - Pesquisadores Top10 pela perspectiva NB-RQ – Brasil	114

I think I'm not the kind of person that's just curious. I valued science ultimately for the good it can do. And I've gotten a lot of satisfaction out of trying to create science and engineering that can do good rather than just be very curious.

Robert Langer

Introdução

Ao longo das últimas décadas, o Brasil foi capaz de criar infraestrutura e base científica fortes no que se refere à geração de conhecimento, reduzindo gradativamente a distância que o separava dos países líderes em pesquisa científica no mundo.

Entre 2006 e 2016, a produção científica brasileira aumentou 125% levando o país a ocupar a 13^a posição no ranking de países com maior número de publicações científicas com participação de aproximadamente 3% no total de publicações. Embora ocupe uma posição modesta, a publicação de artigos científicos no Brasil cresce a um ritmo superior à média mundial e, somente em 2016, apresentou um aumento de cerca de 10% em contraste com 7% ao redor do mundo (AJE, 2016).

Por outro lado, apesar da grande relevância deste crescimento, o Brasil tem sistematicamente reduzido a qualidade do conteúdo da sua produção científica, característica indicada pela perda na sua posição no ranking de países ordenados por citações recebidas por artigo publicado. Deste modo, a ciência produzida em escala no Brasil, na média, tem tido pouca repercussão internacional o que poderia ser atribuído a ausência de diferenciação ou qualidade (MORAES, 2017).

Uma das principais causas deste cenário está associada à ausência de estratégias de longo prazo que viabilizem a cooperação de cientistas brasileiros com pesquisadores estrangeiros, bem como com o desafio em criar condições que favoreçam o aumento na formação e titulação do número de doutores brasileiros, com capacidade de realizar interações com redes internacionais de pesquisa. Apesar dos ótimos resultados de oportunidades de doutoramento no país criados pela comunidade científica brasileira a partir da década de 1980, o foco prolongado na nacionalização reduziu a intensidade da criação de redes e parcerias internacionais de pesquisa, agravado pela barreira linguística.

Neste contexto, oito universidades públicas brasileiras respondem por aproximadamente 2/3 da produção científica do país. Por outro lado, essas e todas as instituições de pesquisa e universidades públicas nacionais enfrentam

importantes obstáculos de planejamento e gestão que, atualmente, vem se agravando com a crise econômica e política.

A partir dos esforços do governo para superar a rigidez na gestão pública, soluções e casos de sucesso estão sendo estabelecidas através de modelos onde entidades privadas sem fins lucrativos administram centros de pesquisa públicos sob contrato com agências federais. Entre os exemplos de sucesso temos o Instituto de Matemática Pura e Aplicada (IMPA) e o Centro Nacional de Pesquisa em Energia e Materiais (CNPEM).

Ainda que estes casos de sucesso possuam grande relevância, a transferência da produção científica brasileira e seu impacto na sociedade são baixos quando comparada com outras economias emergentes. Em especial, destaca-se o baixo número de depósitos de patentes de origem brasileira, sendo este um indicador importante da capacidade de um país de transbordar seu esforço científico a favor da sociedade e competitividade da nação. Além disso, os países que construíram o desenvolvimento a partir do conhecimento científico possuem um maior volume de cientistas em atividade de pesquisa em empresas. Ao contrário dos países desenvolvidos, o Brasil possui um maior número de cientistas dedicados a realização de pesquisas nas universidades.

Desse modo, evidencia-se um relativo descompasso entre a intensiva e crescente produção de conhecimento científico e a geração e exploração de tecnologias inovadoras por empresas brasileiras, limitando os benefícios desse conhecimento ao âmbito acadêmico.

Tal fato mostra-se crítico ao considerarmos a crescente incorporação e ganho de importância do conhecimento nas atividades produtivas observado no mundo nas últimas décadas, tornando a inovação cada vez mais relevante enquanto variável estratégica para a competitividade. Nesse contexto, ideias e conhecimento passaram a ser as fontes mais importantes do crescimento econômico ao passo que a inovação se tornou um dos elementos fundamentais da competitividade das nações. A produção de conhecimento e seu fluxo dinâmico geram novos desafios para as economias, transformando as características da competição global e incitando os países a buscarem novos patamares por meio do investimento em ciência, tecnologia e inovação (MARZANO, 2011).

A produtividade é considerada um fator fundamental para a sustentabilidade da prosperidade econômica das nações e, em linhas gerais,

economias mais competitivas tendem a gerar altos níveis de renda para a população.

A inovação, sobretudo a tecnológica, mostra-se especialmente importante na medida em que as economias alcançam as fronteiras do conhecimento científico. Deste modo, a possibilidade de gerarem mais valor apenas pela integração e adaptação de tecnologias exógenas tende a desaparecer.

Para que as economias em estágios de desenvolvimento mais intensivos em inovação possam manter vantagens competitivas e avançar em atividades de maior valor agregado, suas empresas precisam estar na vanguarda do desenvolvimento de produtos e processos à nível mundial. Portanto, as atividades científicas podem contribuir com a qualidade de vida das pessoas, através da formação de um ambiente propício para geração de produtos, empresas e oportunidades de trabalho com valor agregado.

Nesse contexto, tornam-se fundamentais a existência de um ambiente de inovação propício e do apoio e articulação dos setores público e privado. Particularmente, a progressão para este estágio requer investimentos suficientes em pesquisa e desenvolvimento (P&D), principalmente pelo setor privado; a existência de instituições de pesquisa científica de qualidade capazes de gerar o conhecimento necessário para criação de novas tecnologias; extensa colaboração entre a indústria e as universidades para pesquisa e desenvolvimento tecnológico; e a proteção da propriedade intelectual.

Mostra-se evidente que o Brasil ainda não possui uma economia madura e intensiva em inovação já que a maior parcela de sua economia ainda permeia por estágios de desenvolvimento focados em fatores e eficiência. Entretanto, algumas exceções retratam casos de sucesso como a indústria aeronáutica brasileira que desempenha um papel importante na produção de bens de alto valor agregado e geração de empregos altamente qualificados. Deste modo, o país possui modelos de referências nacionais e internacionais que deveriam ser aplicados e adaptados em outras economias solidárias a outras competências científicas disponíveis no Brasil e no mundo.

Hoje o Brasil não pode depender única e exclusivamente dos esforços das atividades de pesquisa científicas internas para garantir sua capacidade competitiva. Assim, mostra-se necessária a prática de modelos de ciência aberta em rede e a nível global, além da adequação de modelos internacionais de

sucesso que colabore com a redução da distância entre as atividades científicas e empresariais. Além disso, os fluxos de conhecimento devem ganhar maior dinamismo, integrando as atividades de pesquisa aos desenvolvimentos de setores industriais.

Entre os setores que poderiam se beneficiar a partir da incorporação de modelos de inovação aberta em uma perspectiva global destacam-se as indústrias farmoquímica e farmacêutica. Nos últimos dez anos, o mercado farmacêutico brasileiro iniciou uma trajetória de forte ampliação, apresentando taxas de crescimento acima de dois dígitos ao ano, impulsionado pelos medicamentos genéricos e pela ascensão social das classes mais baixas de renda. Mediante uma adequada coordenação entre o setor público e o privado, as empresas farmacêuticas de capital nacional conseguiram aproveitar as oportunidades e ampliaram significativamente sua participação no mercado brasileiro, tornando-se majoritárias nas receitas provenientes do varejo farmacêutico no país.

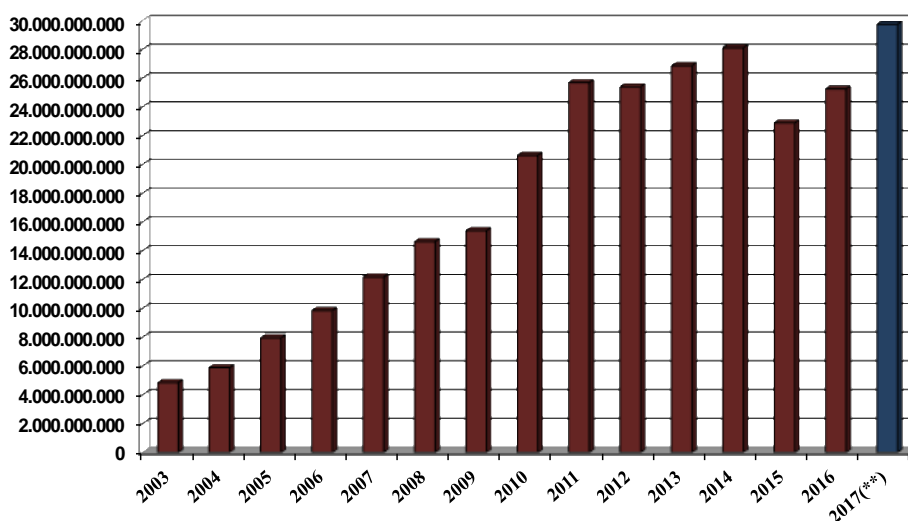


Figura 1 – Mercado farmacêutico brasileiro (US\$).

(**)12 meses móveis até Novembro/2017.

Fonte: Sindusfarma, 2017.

No entanto, o fortalecimento da indústria farmacêutica brasileira não gerou as mesmas oportunidades para a indústria farmoquímica, produtora da principal matéria-prima dos medicamentos. Na média dos últimos dez anos, a participação das importações no consumo de insumos farmacêuticos ativos (IFA) por empresas farmacêuticas brasileiras foi de aproximadamente 90%. Parte relevante desses insumos consumidos pela indústria farmacêutica brasileira é proveniente dos

países asiáticos. O deslocamento expressivo da produção de princípios ativos para a região nas últimas décadas, em particular para Índia e China, transformou-a na maior produtora mundial de farmoquímicos.

Estima-se que cerca de dois terços da produção de IFAs destinada à venda no mundo sejam provenientes da Ásia. Não obstante a reduzida participação do mercado doméstico, a indústria farmoquímica brasileira exportou, em média, 65% de sua produção no período 2009-2013, revelando um descompasso entre a demanda brasileira e a capacidade de oferta das empresas locais.

De modo geral, o crescimento das empresas farmacêuticas brasileiras nas últimas décadas foi impulsionado pelo lançamento de medicamentos genéricos que não necessitam de esforços significativos de P&D, caracterizando uma baixa geração de inovações para o mercado em função dos poucos investimentos em P&D realizados por essas empresas. Entretanto, ainda que os níveis de investimento sejam baixos, o valor dos dispêndios com P&D das empresas vem crescendo ao longo dos anos e atingiu R\$ 2,28 bilhões em 2014 (IBGE, 2016).

Diante das dificuldades relacionadas ao avanço de sua capacidade de inovação, a grande indústria farmacêutica mundial tem buscado permanecer na fronteira tecnológica do setor por meio de parcerias, sobretudo, com instituições de ciência e tecnologia (ICTs) ou empresas *startups* intensivas em P&D. No caso brasileiro, observa-se uma tendência semelhante na medida em que as empresas nacionais identificam as ICTs como principais parceiras para suas atividades. Soma-se ainda o fato do governo brasileiro buscar promover essa interação, favorecendo a articulação de esforços desses atores (PARANHOS, 2010).

Entretanto, para que a indústria farmacêutica brasileira possa avançar em direção às atividades mais intensivas em conhecimento para geração de inovações, muitos desafios ainda precisam ser superados. De modo geral, inovações na indústria farmacêutica podem partir de modificações de uma ou mais propriedades dos compostos como estrutura química, método de síntese do IFA, forma farmacêutica, farmacodinâmica, farmacocinética, bem como outras propriedades terapêuticas (ARONSON, 2008).

Em especial, aspectos relacionados à área de química do estado sólido e às propriedades do estado sólido de materiais farmacêuticos tem recebido cada vez mais atenção, ocupando uma posição de grande importância nas atividades de pesquisa e desenvolvimento farmacêuticas (BYRN et. al, 2017). A compreensão

das propriedades do estado sólido de um IFA mostra-se crítica no desenvolvimento de formulações em função de seus impactos na biodisponibilidade e solubilidade dos fármacos. Por meio da caracterização e estudo do estado sólido pode-se garantir o benefício terapêutico, otimizar o desenvolvimento de produtos e garantir a proteção da propriedade intelectual.

Tendo em vista a crescente relevância da química do estado sólido nas atividades de P&D farmacêuticas, além das grandes oportunidades de avanço das indústrias farmoquímica e farmacêutica brasileira por meio da geração de inovações, esta tese tem como tema central a investigação de indicadores científicos e tecnológicos com vistas à contribuição para o monitoramento da pesquisa e do desenvolvimento tecnológico na área de química do estado sólido. Em especial, busca-se propor alternativas que permitam a identificação de parceiros para colaboração em atividades de pesquisa, desenvolvimento e inovação, favorecendo o avanço do setor.

1.1

Objetivos

A presente pesquisa tem como objetivo investigar indicadores científicos e tecnológicos na área de química do estado sólido utilizando ferramentas de ciência dos dados a partir de publicações científicas e depósitos de patentes que possam contribuir para o aumento da competitividade da indústria farmoquímica e farmacêutica brasileira e em outros mercados emergentes, tendo como foco o estudo de caso de uma nova forma polimórfica descoberta para o Insumo Farmacêutico Ativo (IFA) acetato de dexametasona.

Para tanto, foram estabelecidos os seguintes objetivos específicos:

- Rever a conjuntura científica e tecnológica atual da química do estado sólido e suas implicações para a indústria farmacêutica mundial;
- Elaborar indicadores científicos e tecnológicos a partir de dados bibliográficos estruturados de publicações científicas e de documentos de patentes em química do estado sólido a nível mundial com vistas à sua contribuição para o monitoramento científico e tecnológico da área;

- Elaborar indicadores científicos e tecnológicos a partir de dados bibliográficos estruturados de publicações científicas e de documentos de patentes em química do estado sólido no Brasil com vistas à sua contribuição para o monitoramento científico e tecnológico da área;
- Identificar as principais técnicas de caracterização de polimorfismo em fármacos sólidos;
- Identificar os principais parâmetros químicos do estado sólido com vistas à contribuição para a tomada de decisão em atividades de pesquisa e desenvolvimento nos setores farmoquímico e farmacêutico;
- Aplicar um método de identificação de competências para identificação de potenciais parceiros nacionais e internacionais que possam contribuir com as pesquisas aplicadas na área;
- Aplicar o método de identificação de competências para o caso do insumo farmacêutico ativo acetato dexametasona.

1.2

Estrutura do trabalho

Esta tese está organizada em seis capítulos, sendo o primeiro introdutório. O Capítulo 2 é dedicado a revisão de literatura sobre a P&D e a química do estado sólido na indústria farmacêutica. O Capítulo 3 aborda os conceitos gerais envolvidos na ciência dos dados aplicada na identificação de competências científicas, bem como a extração de informações do conhecimento científico e tecnológico, o uso de indicadores de C&T e a identificação de lideranças e competências científicas. A metodologia da pesquisa é apresentada no Capítulo 4, o qual inclui as etapas propostas e o detalhamento de sua aplicação para os fins da tese. O Capítulo 5 apresenta a discussão dos resultados obtidos e o Capítulo 6 as conclusões da pesquisa, seguido das referências bibliográficas.

2

Pesquisa e Desenvolvimento na Indústria Farmacêutica e a Química do Estado Sólido

2.1

Dinâmicas da P&D na Indústria Farmacêutica

Segundo estimativas do QuintilesIMS, o mercado farmacêutico mundial deve alcançar a marca de US\$ 1,48 trilhão em 2020, representando um crescimento de aproximadamente 35% em relação ao registrado em 2016. Esse crescimento deve ser impulsionado sobretudo pela expansão do mercado de países emergentes e por tendências demográficas decorrentes do avanço no envelhecimento da população e do aumento de sua expectativa de vida (QuintilesIMS, 2016).

De modo geral, o setor farmacêutico é altamente dependente de suas atividades de pesquisa e desenvolvimento (P&D) e a expansão deste mercado deve ser acompanhada por uma grande e crescente demanda por inovação nos próximos anos. A descoberta de novos compostos com atividade terapêutica e a inserção bem-sucedida de medicamentos no mercado encontra-se intrinsecamente ligada ao avanço das inovações científicas e tecnológicas, sendo, portanto, um pilar fundamental para o setor.

Nesse aspecto, deve-se ressaltar que o desenvolvimento de novos medicamentos se mostra longo, extremamente custoso, de alto risco e com reduzidas chances de sucesso (SCHUHMACHER *et al.*, 2016). O processo de P&D na indústria farmacêutica consiste basicamente em cinco etapas que envolvem desde a descoberta de um novo composto até a comercialização do medicamento, passando por quatro fases de testes clínicos. Esse processo leva em média entre 10 e 15 anos.

A primeira etapa, de descoberta, dura em média de quatro a cinco anos e envolve a busca por novos compostos químicos que possam apresentar função terapêutica. Por meio de técnicas de “random screening” e “rational drug design”, entre 5.000 e 10.000 compostos, em média, podem ser avaliados para que se tenha

um candidato promissor (MALERBA & ORSENIGO, 2002). Tendo sido identificado um composto com função terapêutica, inicia-se então a segunda etapa de desenvolvimento de testes pré-clínicos com duração de cerca de um ano que envolvem avaliação de toxicidade, determinação farmacocinética, modificações de estrutura molecular, avaliação de rotas de síntese alternativas, avaliação e preparação de modelos de formulação, estudos de formas cristalinas e propriedades do estado sólido, testes de estabilidade, preparação de formulações para a etapa de testes clínicos, entre outros. Juntas, as etapas de descoberta e testes pré-clínicos representam aproximadamente 33% de todos os custos do processo de desenvolvimento e possuem menos de 1% de chances de sucesso (PARANHOS, 2010; IFPMA, 2016; PAUL *et al.*, 2010).

A terceira etapa, por sua vez, envolve três fases de testes clínicos e dura em média de seis a sete anos, sendo responsável por até 65% dos custos totais do desenvolvimento (PAUL *et al.*, 2010). Durante a Fase I são realizados testes com voluntários saudáveis para exame de segurança, tolerância, relação entre dose e efeito, duração do efeito e efeitos colaterais. Na Fase II, iniciam-se testes do composto em pequenos grupos de pacientes para avaliação da bioequivalência, eficácia e segurança. Por fim, a Fase III envolve a avaliação da eficácia terapêutica, segurança e valor em tratamento de longo prazo em pacientes, além da definição de dosagem e indicações, contraindicações, efeitos colaterais e interações medicamentosas (PARANHOS, 2010). Embora os custos dessa etapa sejam consideravelmente altos, as chances de sucesso das três etapas mostram-se relativamente altas: 65% para a Fase I, 40% para a Fase II e 50% para a Fase III. (IFPMA, 2016).

Ao término dos testes clínicos, inicia-se então a quarta etapa de submissão do registro do medicamento para que seja aprovado pelas autoridades para comercialização. A etapa dura entre um e dois anos e representa apenas 5% dos custos do projeto. Tendo sido autorizado o início da comercialização, mais uma fase de testes clínicos após inserção no mercado é iniciada com um grande número de pacientes para que se possa identificar possíveis novos efeitos colaterais ou reações adversas que não tinham sido observadas nas fases II e III. (PARANHOS, 2010).

Como já destacado, o processo de P&D no setor farmacêutico envolve custos elevados, alto risco e longo período de conclusão. Atualmente, estima-se

que o custo do desenvolvimento bem-sucedido de um medicamento exceda 2,6 bilhões de dólares, enquanto na década de 1970 este valor não superava a cifra de 180 milhões. Este aumento expressivo reflete os diversos desafios técnicos, regulatórios e econômicos envolvidos nas atividades de P&D da indústria farmacêutica e tende a se acelerar com o passar do tempo (DIMASI *et al.*, 2016).

De modo geral, parte considerável dos investimentos feitos por empresas farmacêuticas em P&D é perdida em projetos que não se concretizam em medicamentos aprovados para comercialização em função das altas taxas de fracasso nas atividades de P&D farmacêuticas. Ainda que um composto apresente perspectivas promissoras no início do desenvolvimento, somente os testes pré-clínicos e clínicos poderão demonstrar sua eficácia, qualidade e segurança. Assim, os gastos perdidos em P&D aumentam consideravelmente na medida em que o projeto fracassa em etapas mais avançadas do desenvolvimento (IFPMA, 2016).

2.1.1

Novos Modelos de Inovação e Negócio na P&D farmacêutica

Mesmo em face dos diversos desafios apresentados e do crescimento dos custos de desenvolvimento, a indústria farmacêutica não diminuiu seu ímpeto inovativo tendo passado a buscar a adoção de novos modelos de inovação para alcançar seus objetivos. Desse modo, a inovação aberta e a exploração novos modelos de negócio como *joint ventures* entre empresas farmacêuticas e outros atores externos tem se apresentado como alternativas interessantes para aumentar a produtividade da P&D no setor farmacêutico (IFPMA, 2016).

Nos últimos anos, observou-se um progressivo deslocamento do modelo de produção totalmente integrado para um modelo aberto na indústria farmacêutica. Esse processo é marcado pelo *outsourcing* de diversas etapas da cadeia produtiva e inovativa farmacêutica, sobretudo as com maiores exigências regulatórias e de maior especialização. Ao descrever esse processo, Paranhos (2010) destaca que atividades como a realização de P&D passaram a ser feitas em parceria com empresas *start-ups* e universidades (podendo ser totalmente terceirizadas), o desenvolvimento do produto e a produção clínica pelas *Contract Research Organizations* (CROs), a produção, validação e formulação das tecnologias pelas

Contract Manufacturing Organizations (CMOs) e a distribuição pelas Contract Sales Organizations (CSOs), tornando a empresa totalmente desintegrada.

Assim, embora a indústria farmacêutica tradicionalmente colaborasse com parceiros externos para acessar know-how específico, tais colaborações são cada vez mais fundamentais nos dias de hoje ao passo em que a complexidade e os custos da P&D farmacêutica aumentam consideravelmente. Essas colaborações permitem o acesso a competências e tecnologias específicas como, por exemplo, novos alvos de medicamentos, validação de alvos, modelos animais e biomarcadores, entre outras (SCHUHMACHER *et al.*, 2016).

A descentralização das atividades de P&D farmacêuticas possibilita o aumento de sua produtividade, reduzindo o tempo de desenvolvimento e otimizando gastos. Como exemplo, pode-se citar a iniciativa da empresa farmacêutica GlaxoSmithKline (GSK) que atualmente gasta quase metade de seu orçamento para P&D em colaboração com parceiros externos como universidades ou empresas de biotecnologia. Recentemente, a empresa criou um programa chamado *Discovery Partnerships with Academia* (DPAc), uma iniciativa que busca promover parcerias com universidades desde a etapa de descoberta até a otimização de produção em qualquer área ou modalidade de tratamento. Um dos aspectos mais interessantes do modelo adotado pela GSK reside no fato do parceiro acadêmico poder se beneficiar do *know-how* da empresa no descobrimento de novos compostos e de seus recursos em química medicinal, segurança pré-clínica e farmacocinética, enquanto a GSK pode acessar ideias da academia para promover seu potencial inovador (GSK, 2017).

Como já destacado, as atividades de P&D são fundamentais para o sucesso e avanço da indústria farmacêutica e sua importância é evidenciada pelo gasto cumulativo, tanto a nível de empresas individuais quanto da indústria como um todo. O gasto total em P&D de empresas farmacêuticas e de biotecnologia cresceu de 108 bilhões de dólares em 2006 para 141 bilhões de dólares em 2015 (IFPMA, 2016). De todos os setores industriais, a indústria farmacêutica é aquela que mais tem investido em P&D. Além disso, entre as dez empresas com maiores investimentos em P&D no mundo no ano fiscal de 2014/2015, quatro foram empresas farmacêuticas: Novartis (5), Roche (7), Johnson e Johnson (8); e Pfizer (10) (IFPMA, 2016).

Associado ao fortalecimento do *outsourcing* e do aumento dos gastos, observou-se ainda um outro processo na dinâmica da P&D no setor farmacêutico mundial nas últimas décadas: sua internacionalização. Ainda que mais lenta no setor farmacêutico do que em outros, sobretudo em função das diferentes regulações nos países, a internacionalização tem ganhado cada vez mais força (PARANHOS, 2010).

Neste contexto, novos desafios são apresentados para a dinâmica do processo sendo necessário reavaliar os modelos de inovação e negócio no setor farmacêutico para garantir sua dinâmica. Para empresas que buscam como estratégia ocupar uma posição de liderança na inovação na indústria farmacêutica, a geração do conhecimento ou sua integração mostra-se fundamental. Nesse modelo, a inovação é criada internamente ou com a incorporação de recursos externos, mantendo-se o domínio sobre o conhecimento gerado. Para essas empresas, são condições importantes a criação interna de competências-chave de P&D, criação de redes com parceiros acadêmicos e prestadores de serviço e a definição de estratégias de crescimento corretas para lidar com os altos custos do processo de inovação garantindo o investimento sustentável em P&D para o contínuo fluxo de lançamento de novos produtos. (SCHUHMACHER *et al.*, 2016).

Por outro lado, empresas farmacêuticas que não buscam uma estratégia de liderança em inovação podem direcionar seus esforços para as oportunidades de crescimento do mercado de genéricos e de países emergentes. Enquanto o mercado de medicamentos inovadores de prescrição apresenta uma expectativa de crescimento anual de aproximadamente 6% entre 2015 e 2020, o mercado de genéricos deve crescer 12% por ano no mesmo período (EVALUATE, 2015).

Independentemente da estratégia de cada empresa, os desafios impostos à indústria farmacêutica favorecem cada vez mais uma estrutura aberta de colaboração. Enquanto empresas inovadoras podem buscar no modelo aberto de inovação a possibilidade de reduzir gastos de P&D e otimizar o longo processo de lançamento de novos medicamentos, empresas focadas em genéricos com menor intensidade inovativa podem buscar no modelo de colaboração a oportunidade de reduzir seus custos para serem mais competitivas em um mercado ditado sobretudo pelo preço.

De modo geral, empresas farmacêuticas líderes em inovação e que se apropriam das maiores parcelas deste mercado estão localizadas no eixo dos países mais desenvolvidos – Estados Unidos, Europa e Japão -, os quais concentram a maior parte das atividades de P&D farmacêuticas (PARANHOS, 2010). Entretanto, o grande crescimento do mercado em países emergentes, sobretudo àquele relacionado aos medicamentos genéricos, tem favorecido a criação de competências específicas nesses países que podem, no longo prazo, permitir o avanço em direção a atividades mais intensivas em P&D no setor farmacêutico para geração de produtos inovadores.

Entretanto, deve-se ressaltar que os desafios enfrentados pela indústria farmacêutica não se limitam a dimensão gerencial do estabelecimento de modelos mais adequados de negócio e de inovação nos dias atuais. A dinâmica de inovação da indústria farmacêutica também se depara com desafios de origem técnica que surgem na medida em que o avanço tecnológico permite uma melhor compreensão dos diversos aspectos relacionados a ação dos medicamentos e à saúde humana. Como já destacado anteriormente, além do descobrimento de novos compostos com eficácia terapêutica, inovações na indústria farmacêutica podem partir de modificações de uma ou mais propriedades dos compostos como estrutura química ou método de síntese, forma farmacêutica, farmacodinâmica, farmacocinética, bem como outras propriedades terapêuticas (ARONSON, 2008).

Dentre as diversas dimensões nas atividades de P&D de medicamentos, aspectos relacionados à área de química do estado sólido e às propriedades do estado sólido de materiais farmacêuticos tem recebido cada vez mais atenção (BYRN *et al.*, 2017). Em especial, observa-se a crescente importância da compreensão das propriedades do estado sólido de um insumo farmacêutico ativo¹ (IFA) para o desenvolvimento de formulações, tendo em vista seus impactos na biodisponibilidade e solubilidade dos fármacos. A avaliação e controle de parâmetros do estado sólido permitem garantir o benefício terapêutico, otimizar o desenvolvimento de produtos e garantir a proteção da propriedade intelectual.

Assim, tendo em vista a relevância do avanço do processo de *outsourcing* e dos modelos de inovação aberta e colaboração no setor farmacêutico, bem como a

¹ Insumos farmacêuticos ativos (IFAs) ou fármacos podem ser conceituados como substâncias que apresentam propriedades farmacológicas com finalidade medicamentosa, utilizada para diagnóstico, alívio ou tratamento. É o princípio ativo do medicamento (ANVISA, 2017).

crecente importância da área da química do estado sólido no âmbito das atividades de P&D, a presente pesquisa busca aplicar ferramentas de ciência dos dados para elaboração de indicadores e identificação de competências específicas, tendo como objetivo investigar tendências e desafios científicos e tecnológicos que possam contribuir para o aumento da competitividade da indústria farmacêutica em países emergentes como o Brasil. Em especial, busca-se propor soluções que permitam a identificação de parceiros para colaboração em atividades de P&D, favorecendo o avanço do setor.

Tendo sido discutido brevemente as dinâmicas das atividades P&D no setor farmacêutico, a próxima seção apresenta uma discussão da importância da química do estado sólido na atualidade.

2.2

Química do Estado Sólido

Nas atividades de P&D de novos fármacos grande ênfase geralmente é dada à etapa de descoberta de um novo composto com atividade terapêutica para a cura, tratamento ou prevenção de doenças. Entretanto, ainda que a descoberta dos compostos seja de fato relevante, a capacidade de gerar um medicamento de modo eficiente e dentro de um prazo razoável depende diretamente das atividades de formulação e produção de fármacos estáveis e biodisponíveis para criação de produtos bem-sucedidos. Sendo assim, deve-se levar em conta que diversas características físicas e químicas desempenham um papel central no processo, já que um medicamento se configura não como um composto isolado, mas como parte de uma mistura de materiais com características físicas e químicas que podem determinar seu sucesso ou fracasso terapêutico.

Nesse contexto, a área denominada química do estado sólido e o estudo das propriedades do estado sólido de materiais farmacêuticos tem ocupado uma posição cada vez mais relevante não apenas em atividades de P&D farmacêuticas, mas em toda a cadeia do setor. Em especial, essa área desempenha um papel importante na compreensão e controle das propriedades físicas e químicas de IFAs para garantir sua eficácia terapêutica, otimizar o desenvolvimento de produtos e garantir a proteção da propriedade intelectual.

Tal fato mostra-se crítico ao considerarmos que aproximadamente 70% dos medicamentos disponíveis atualmente são produzidos com formulações sólidas, para uso oral, como comprimidos, cápsulas, drágeas, entre outras. (ABRIMOV, 2016). Além disso, de acordo com os dados da empresa norte-americana SSCI em estudo realizado em 2007, dentre 245 compostos analisados, aproximadamente 90% apresentaram variações de sua forma no estado sólido (STAHLY, 2007). Os dados refletem a complexidade da temática e a necessidade do controle e monitoramento contínuo do estado sólido na indústria farmacêutica.

A eficácia terapêutica de cada IFA está diretamente relacionada a suas características no estado sólido. Enquanto alguns IFAs podem apresentar somente uma ou duas formas polimórficas conhecidas, outros podem existir em mais de 20 formas polimórficas diferentes, cada uma com propriedades físicas e químicas distintas. Em alguns casos, as propriedades de formas sólidas diferentes de um mesmo IFA podem ser bastante similares. Em outros casos, podem apresentar diferenças significativas gerando impactos em suas propriedades físico-químicas, incluindo sua solubilidade, com consequências diretas no processo de produção do medicamento, a estabilidade de sua formulação e sua biodisponibilidade (HUANG & TONG, 2004; LLINÀS & GOODMAN, 2008).

De modo geral, os sólidos farmacêuticos podem ser classificados em duas categorias: cristalinos ou amorfos. Simplificadamente, pode-se dizer que os sólidos cristalinos (como polimorfos, sais, cocristsais e solvatos) apresentam um arranjo ordenado e periódico de suas moléculas, enquanto os sólidos amorfos são caracterizados por não apresentarem ordenação espacial definida (BYRN et al, 2017).

Dentre os tipos de sólidos que um IFA pode apresentar, destacam-se:

- Polimorfos: IFAs que apresentam a mesma fórmula química, mas diferente estrutura cristalina
- Solvatos: IFAs com moléculas de solventes integradas à sua rede cristalina (quando a água é o solvente presente na estrutura cristalina denomina-se hidrato)
- Sais: IFAs cristalizados com ácidos ou bases

- Cocristais: IFAs cristalizados com outras substâncias sólidas e neutras em condições ambientes, inclusive cristalizados com outros IFAs
- Amorfos: IFAs com estruturas cristalinas espacialmente desordenadas

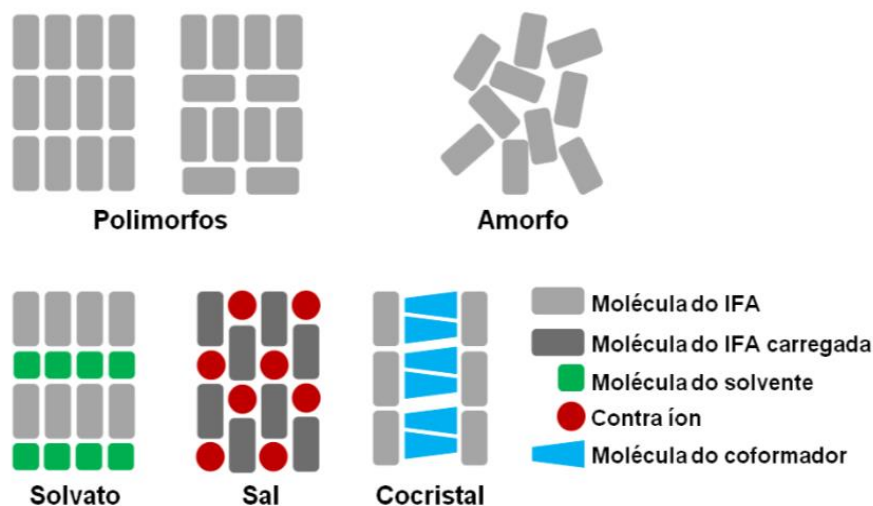


Figura 2 – Representação esquemática dos vários tipos de formas sólidas.
Fonte: PRADO & ROCHA, 2015.

No que se refere aos polimorfos, deve-se ressaltar ainda que existem duas definições comumente encontradas no âmbito farmacêutico. A primeira definição, difundida pela *International Committee on Harmonization* (ICH) e adotada por agências reguladoras como o *Food and Drug Administration* (FDA) nos EUA, engloba no conceito de polimorfo os solvatos e hidratos, bem como os materiais amorfos. Já a segunda definição, mais estrita, restringe o conceito de polimorfos a formas sólidas de um mesmo composto químico, não incluindo, portanto, solvatos e amorfos (PRADO & ROCHA, 2015).

As propriedades que podem apresentar diferenças entre formas sólidas de um IFA incluem cor, solubilidade, forma do cristal, absorção e sorção de água, tamanho de partícula, dureza, densidade, entre outras. Formas sólidas diferentes podem apresentar variações em seu ponto de fusão, propriedades espectrais e estabilidade termodinâmica.

Em um medicamento, essas variações das propriedades podem gerar diferenças na taxa de dissolução, absorção oral, biodisponibilidade, níveis de irritação gástricos e toxicidade. Assim, tanto a segurança quanto a eficácia do

medicamento são impactadas pelas propriedades que variam entre as diferentes formas sólidas. Em especial, grande atenção é dada a estabilidade tendo em vista a possibilidade do surgimento de formas sólidas indevidas ao longo das etapas de todo o processo de desenvolvimento. Transições de formas sólidas de um mesmo IFA podem ocorrer ao longo de toda a cadeia, desde a produção do IFA até sua formulação, armazenamento, distribuição e uso (BYRN et al., 2017)

Embora o fenômeno do polimorfismo seja conhecido desde o início do século 19, foi somente nas últimas duas décadas que a temática do controle e monitoramento do polimorfismo e da química do estado sólido tornou-se procedimento padrão na indústria farmacêutica (BHARDWAJ, 2016). Entre os fatores que contribuíram para que as empresas farmacêuticas passassem a tratar a temática de forma séria, destacam-se a identificação de problemas de produção como no caso do IFA ritonavir envolvendo o surgimento de um polimorfo mais termodinamicamente estável com menor solubilidade e biodisponibilidade levando a retirada do produto do mercado; e o início de disputas patentárias envolvendo formas polimórficas, sobretudo com a difusão dos medicamentos genéricos (BERNSTEIN, 2002; BHARDWAJ, 2016)

O caso do IFA ritonavir tornou-se emblemático nas discussões sobre química do estado sólido na indústria farmacêutica ao demonstrar o impacto do polimorfismo na dissolução, com implicações para biodisponibilidade. A empresa farmacêutica Abbott Laboratories havia introduzido no mercado o medicamento Norvir, uma formulação sólida do IFA ritonavir com atividade antirretroviral, sendo utilizado para tratar infecção pelo vírus HIV e AIDS. A formulação do medicamento envolvia a única forma polimórfica conhecida do ritonavir à época e foi comercializada durante 18 meses. Entretanto, o surgimento de uma nova forma mais termodinamicamente estável e menos solúvel durante a produção resultou na diminuição da biodisponibilidade do medicamento. Tendo sido identificado o problema, a empresa precisou retirar do mercado lotes do medicamento com grandes prejuízos financeiros. Caso a Abbott tivesse explorado o estado sólido do ritonavir para conhecer suas possíveis formas e suas respectivas propriedades previamente, os riscos aos pacientes poderiam ter sido evitados, bem como os gastos com a retirada do produto e desenvolvimento de uma nova formulação (BAUER et al. 2001).

A figura 3 a seguir apresenta uma linha do tempo dos eventos envolvendo impactos do polimorfismo e outras formas sólidas no desempenho e na qualidade de medicamentos. Como pode ser observado, os diversos eventos ressaltam a complexidade e importância do controle e monitoramento do estado sólido de produtos farmacêuticos, sobretudo para que se possam evitar consequências mais sérias para os usuários dos medicamentos.

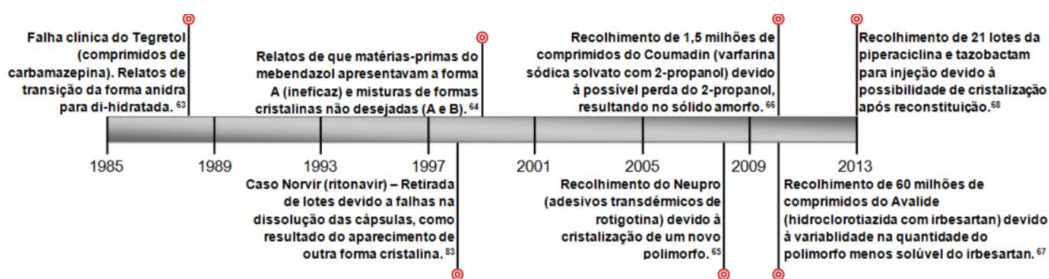


Figura 3 – Linha do tempo de eventos envolvendo impactos na indústria farmacêutica devido a formas cristalinas de fármacos. Fonte: PRADO & ROCHA, 2015.

O reconhecimento da importância do controle e monitoramento do estado sólido dos IFAs foi acompanhado ainda pelo avanço regulatório sobre o tema ao redor do mundo. Em especial, destaca-se o lançamento do guia do *International Commitee on Harmonization (ICH) Q6A* em 1999. Esse guia trouxe diretrizes para que se determinasse qual a forma sólida presente em um IFA, bem como em seu produto farmacêutico (medicamento), fornecendo árvores de decisão para serem utilizadas como roteiro para o estabelecimento de critérios específicos. Desde então, muitos avanços ocorreram tanto na perspectiva regulatória quanto da compreensão e controle do estado sólido em produtos farmacêuticos.

2.2.1

Química do estado sólido no desenvolvimento de novos fármacos

O processo de desenvolvimento de novos fármacos inicia-se com a descoberta de um composto químico (também conhecido como nova entidade química) com potencial atividade terapêutica. Tendo sido identificado um candidato promissor, iniciam-se as etapas de teste pré-clínicos incluindo atividade biológica, testes de metabolismo, determinação farmacocinética, estudos toxicológicos, entre outros. Em seguida, na medida em que se avança nos estudos

pré-clínicos e se inicia o estágio de desenvolvimento para realização dos estudos clínicos, esforços são direcionados para tornar o composto descoberto em um insumo farmacêutico ativo (IFA) para utilização em uma formulação, geralmente na forma sólida. Inicia-se então o desenvolvimento do novo fármaco como prova de conceito, o que envolve estudos clínicos iniciais em humanos (Fases I e II). Nas etapas seguintes, a formulação para o produto acabado (medicamento) e o desenvolvimento de processos passam a ser os principais focos, de modo a garantir a eficácia, segurança e qualidade do medicamento desde sua produção até o armazenamento (ABRIMOV, 2016).

Inicialmente, quando o IFA é produzido, deve-se assegurar que a forma sólida desejada pode ser obtida de maneira pura, consistente, reprodutível. Em seguida, ao inserir o IFA em uma formulação para produção do medicamento, mostra-se imprescindível garantir que nenhuma transição da forma sólida do IFA ocorra ao longo do processo.

Do mesmo modo, após sua produção, deve-se assegurar que nenhuma transição ocorra ao longo do tempo de armazenamento até o consumo, já que propriedades importantes podem ser afetadas com a mudança da forma sólida. Nesse sentido, a estabilidade deve ser avaliada ao longo de toda a cadeia no que se refere às condições ambientais, de armazenamento e de embalagem (BYRN et al., 2017)

Além de desempenhar um papel fundamental no desenvolvimento de novos fármacos, a profunda compreensão das propriedades do estado sólido de determinado IFA também contribuem para a proteção da propriedade intelectual relacionada, já que formas sólidas e/ou o processo para sua obtenção são passíveis de proteção por meio do depósito de patente.

Como já destacado, variações da forma sólida podem levar a alterações das propriedades físico-químicas do IFA, com potenciais consequências para o seu desempenho e segurança (HUANG & TONG, 2004). Nesse sentido, as agências reguladoras ao redor do mundo – como a *Food and Drug Administration* (FDA) nos EUA e a Agência Nacional de Vigilância Sanitária (ANVISA) no Brasil – atuam continuamente para garantir a adequação das empresas farmacêuticas às práticas recomendadas no âmbito da avaliação do estado sólido.

Segundo definição adotada pela ANVISA (2014), a biodisponibilidade “indica a velocidade e a extensão de absorção de um princípio ativo em uma

forma de dosagem, a partir de sua curva de concentração/tempo na circulação sistêmica ou sua excreção na urina”. A seleção da forma sólida de um determinado IFA possui impactos diretos sobre a solubilidade e a taxa de dissolução do produto farmacêutico, as quais se relacionam diretamente com a biodisponibilidade e o desempenho do medicamento sendo, portanto, duas das principais propriedades a serem controladas.

As diferenças mais significativas em solubilidade são observadas entre formas cristalinas e amorfas de IFAs (HANCOCK & PARKS, 2000; MURDANDE et al., 2010). Enquanto os sólidos cristalinos anidros são conhecidos por apresentarem maior solubilidade em água do que os hidratos, solvatos formados a partir de solventes miscíveis em água são comumente mais solúveis em água do que sua correspondente forma anidra não solvatada (PUDIPEDDI & SARAJUDDIN, 2005). De modo geral, a utilização de sais e cocristsais aumenta significativamente a taxa de dissolução em formulações sólidas sendo, portanto, os principais tipos de sólidos escolhidos para aprimorar a solubilidade (dissolução) de IFAs pouco solúveis (MAHESHWARI et al., 2012). Entretanto, embora as principais diferenças estejam relacionadas a variações do tipo de forma sólida, formas polimórficas de um mesmo IFA também podem apresentar diferenças relevantes devendo ser observadas. (ABRIMOV, 2016).

Ao discutirem o impacto do polimorfismo sobre a biodisponibilidade, Prado e Rocha (2015) apresentam o exemplo do antibiótico palmitato de cloranfenicol (CAP), um pró-fármaco do cloranfenicol que possui três formas cristalinas conhecidas: A, B e C. Como destacam os autores:

“Para que ocorra a absorção, o pró-fármaco deve ser hidrolisado por esterases intestinais. A forma metaestável do CAP, forma B, apresenta uma bioatividade oito vezes maior que a forma A (forma estável) e, se administrada em humanos, pode causar efeitos adversos como aplasia medular. Este fato pode ser atribuído à maior solubilidade da forma B, que resulta em uma alta taxa de hidrólise e, conseqüentemente, em uma maior biodisponibilidade.” (PRADO & ROCHA, 2015, p. 11).

Prado e Rocha (2015) ressaltam ainda que existem diversos medicamentos disponíveis comercialmente que deram origem à medicamentos genéricos que possuem formas cristalinas diferentes, particularmente entre o medicamento inovador (referência) e o genérico. Como exemplo, pode-se citar os casos do cloridrato de ranitidina, varfarina sódica, famotidina, entre outros.

Além da biodisponibilidade, diferenças de propriedades de formas sólidas podem ter impactos diretos sobre a processabilidade e a estabilidade dos produtos farmacêuticos. A processabilidade contempla o conjunto de características que durante a fabricação de um medicamento podem sofrer influência, incluindo fluxo do pó, compressibilidade, entre outras. Já a estabilidade, tanto química quanto física, mostra-se relevante tendo em vista as possíveis transições de formas sólidas e as diferenças na reatividade e taxas de degradação entre formas distintas (PRADO & ROCHA, 2015). A figura abaixo apresenta um resumo das propriedades que podem ser impactadas pela variação da forma sólida.

<p style="text-align: center;">Empacotamento</p> <ul style="list-style-type: none"> • Volume da cela unitária • Densidade • Índice de refração • Higroscopicidade 	<p style="text-align: center;">Termodinâmicas</p> <ul style="list-style-type: none"> • Temperatura de fusão • Entalpia e Entropia • Solubilidade • Pressão de vapor
<p style="text-align: center;">Mecânicas</p> <ul style="list-style-type: none"> • Dureza • Força de tensão • Compressibilidade • Fluxo 	<p style="text-align: center;">Cinéticas</p> <ul style="list-style-type: none"> • Taxa de dissolução • Taxas de reações no estado sólido • Estabilidade
<p style="text-align: center;">Superficiais</p> <ul style="list-style-type: none"> • Energia livre de superfície • Tensão interfacial • Morfologia • Área superficial 	<p style="text-align: center;">Espectroscópicas</p> <ul style="list-style-type: none"> • Transições eletrônicas, vibracionais e rotacionais • Transições de spin nuclear

Figura 4 – Propriedades que podem diferir entre formas cristalinas de um IFA.
Fonte: PRADO & ROCHA, 2015.

Assim, como pode ser observado, mudanças do estado sólido podem impactar propriedades importantes para o desenvolvimento de um medicamento eficaz e seguro. Para tanto, mostra-se fundamental compreender detalhadamente as características específicas de cada substância no que se refere a suas formas sólidas possíveis e respectivas propriedades. O controle e monitoramento do estado sólido deve ocorrer ao longo de toda a cadeia farmacêutica, desde a descoberta inicial até o desenvolvimento, produção, armazenamento e distribuição.

2.2.2

Controle e monitoramento do estado sólido

Para a correta compreensão, controle e monitoramento do estado sólido na indústria farmacêutica diversas técnicas analíticas desempenham um papel central, tanto na caracterização inicial do IFA quanto no produto acabado (medicamento). Entre os métodos disponíveis estão os cristalográficos, espectroscópicos, térmicos e microscópicos.

Geralmente, diversas técnicas são empregadas conjuntamente para uma correta compreensão do estado sólido. Quando um novo composto químico é descoberto, um amplo número de técnicas analíticas é necessário para que se possa caracterizar e conhecer em detalhes as propriedades do estado sólido do novo fármaco a ser desenvolvido. Em especial, a difração de raios-X por monocristal possui grande relevância nessa etapa inicial, fornecendo identificação definitiva de formas cristalinas por fornecer dados sobre empacotamento, conformação das moléculas e interações. Entretanto, com o avanço do desenvolvimento e a obtenção de determinada forma sólida adequada para a formulação, outras técnicas menos complexas passam a ser suficientes para o controle e monitoramento do estado sólido como a difração de raios-X por policristais (PRADO & ROCHA, 2015).

Embora a difração de raios-X por policristais seja a mais indicada para determinação de formas sólidas em materiais farmacêuticos sólidos, outras técnicas correlatas como análise térmica podem ser fundamentais para solução de problemas específicos. O quadro abaixo apresenta as principais técnicas analíticas comumente utilizadas para caracterização de formas sólidas.

Quadro 1 – Técnicas analíticas comumente utilizadas para caracterização do estado sólido.

Técnica Analítica	Informações	Vantagens	Desvantagem/Dificuldades
Difração de raios-X por policristais (PXR)	<ul style="list-style-type: none"> - Informação estrutural - Picos de difração únicos para formas cristalinas - Halo para amorfo 	<ul style="list-style-type: none"> - Não destrutível - Análise quali e quantitativa 	<ul style="list-style-type: none"> - Orientação preferencial
Difração de raios-X por monocristal	<ul style="list-style-type: none"> - Informação estrutural - Resolução de estruturas cristalinas 	<ul style="list-style-type: none"> - Não destrutível 	<ul style="list-style-type: none"> - Requer um monocristal de boa qualidade
Calorimetria Diferencial Exploratória (DSC)	<ul style="list-style-type: none"> - Temperatura de transição vítrea, de cristalização e de fusão, capacidade calorífica, calor de fusão, transição e cristalização - Interações IFA-IFA e IFA-excipientes 	<ul style="list-style-type: none"> - Pequena quantidade de amostra - Análise quali- e quantitativa 	<ul style="list-style-type: none"> - Destrutível - Sem informação da natureza do evento térmico
Termogravimetria (TG)	<ul style="list-style-type: none"> - Transições envolvendo ganho ou perda de massa - Temperatura de decomposição 	<ul style="list-style-type: none"> - Pequena quantidade de amostra - Análise quantitativa da estequiometria de solvatos 	<ul style="list-style-type: none"> - Destrutível - Inadequado para amostras que degradam em baixas temperaturas
Infravermelho médio (MID-IR)	<ul style="list-style-type: none"> - Vibrações intramoleculares, ligações de hidrogênio - Bandas únicas ou deslocamentos em polimorfos 	<ul style="list-style-type: none"> - Rapidez - Pequena quantidade de amostra 	<ul style="list-style-type: none"> - Preparação de amostra pode induzir transição de fase - Interferência da umidade
Infravermelho próximo (NIR)	<ul style="list-style-type: none"> - Sobretons e combinações de vibrações na região de infravermelho médio - Sensível aos diferentes estados da água 	<ul style="list-style-type: none"> - Rapidez - Não destrutível - Sem preparação da amostra - Habilidade de penetrar por frascos de vidro 	<ul style="list-style-type: none"> - Baixa sensibilidade e seletividade - Deslocamento da linha base
Raman	<ul style="list-style-type: none"> - Vibrações intramoleculares - Bandas únicas ou deslocamentos em polimorfos - Complementar ao infravermelho 	<ul style="list-style-type: none"> - Rapidez - Não destrutível - Sem preparação da amostra - Habilidade de penetrar por frascos de vidro - Pequena quantidade de amostra - Em baixas 	<ul style="list-style-type: none"> - Aquecimento local da amostra - Fluorescência da amostra - Fotodegradação

		frequências pode analisar vibrações de rede - Insensível à água	
Ressonância magnética nuclear no estado sólido (SSNMR)	- Ambiente do núcleo e químico dentro da molécula - Dinâmica molecular - Interações IFA-IFA e IFA-excipientes	- Não destrutível - Análise quali- e quantitativa sem calibração	- Espectro afetado pela água - Deslocamento da linha base - Preparação de amostra pode induzir transição de fase
Microscopia de luz polarizada (PLM)	- Cristalinidade (birrefringência) - Morfologia, cor e tamanho	- Pequena quantidade de amostra - Fácil de usar - Pequena ou nenhuma preparação de amostra	- Informação quantitativa não disponível
Microscopia eletrônica de varredura (MEV)	- Propriedades topográficas	- Pequena quantidade de amostra - Maior resolução do que a microscopia óptica	- Requer preparo da amostra e de condições do equipamento
Termomicroscopia (TOA)	- Informações complementares de transição de fase e perda de solvente	- Variação da temperatura	- Destrutível

Fonte: PRADO & ROCHA, 2015

2.2.3

Descoberta de uma nova forma polimórfica de Acetato de Dexametasona

Durante as atividades de pesquisa planejadas para realização desta tese de doutorado, foi descoberta uma nova forma polimórfica do IFA Acetato de Dexametasona pela equipe de pesquisadores da NanoBusiness Informação e Inovação Ltda. (uma *CRO spin-off* do Departamento de Engenharia Química e de Materiais PUC-Rio) em parceria com professores da UFF e UFRJ que colaboraram com a atividade de caracterização desse novo polimorfo. Essa iniciativa de pesquisa entre uma empresa de base tecnológica e um grupo de professores universitários teve apoio do CNPq, Faperj e Inmetro.

Assim, com apoio de órgãos de fomento à pesquisa e com base nas demandas recorrentes do setor farmacêutico, foi estabelecida na empresa uma linha de pesquisa de descoberta de novas formas polimorfas. Esse esforço de P&D deu origem a um artigo intitulado “*New Polymorph Form of Dexamethasone Acetate*”, o qual foi aceito recentemente para publicação em periódico internacional (**ANEXO A**). Os dados cristalográficos da nova forma polimórfica descoberta foram obtidos pela técnica de difração de raios-X por monocristal e depositados no *Cambridge Crystallographic Data Centre* (CCDC) pelos mesmos autores do artigo supracitado (**Anexo B**). Como será discutido na sessão seguinte, caso sejam atendidas as exigências da legislação patentária nacional e internacional, os esforços de P&D realizados na descoberta dessa nova forma polimórfica do Acetato de Dexametasona poderão ser passíveis a proteção por meio de depósito de patente. Dados do mercado de Acetato de Dexametasona são apresentados no **Anexo C**.

Considerando que as atividades de pesquisa e prestação de serviços supracitados estão associadas a metrologia científica, o Laboratório de Caracterização de Materiais e Química do Estado Sólido da NanoBusiness recebeu recentemente a acreditação pela Coordenação Geral de Acreditação do Inmetro (Cgcre), segundo os requisitos estabelecidos na ABNT NBR ISO/IEC 17025:2005, sendo, portanto, reconhecida a competência da empresa para a realização de ensaios de caracterização por difração de raios-X, sob o número CRL-1264. Nesse sentido, a NanoBusiness é atualmente pioneira e única empresa privada na América Latina com a Acreditação ISO/IEC 17025:2005 para caracterização de polimorfismo em fármacos.

Para atender às exigências e necessidades do mercado e das atividades da metrologia científica, três iniciativas de pesquisa foram concebidas. A primeira buscou desenvolver novas abordagens metodológicas para validação metrológica e avaliação de incerteza de medição para difração de raios-X por policristalinas (DRXP). Seus resultados estão em fase de publicação no artigo intitulado: “*New approaches for validation and evaluation of measurement uncertainty by x-ray powder diffraction (XRPD)*”, tech.memo, N°01/18, DEI/PUC-Rio, Rio de Janeiro, 2018 – (**APENDICE A**).

A segunda iniciativa envolve o planejamento e execução de um programa de ensaio de proficiência proposto pela NanoBusiness conjuntamente com o Inmetro. Esse programa de proficiência, pioneiro no mundo, tem como base inicial os resultados obtidos a partir da aplicação de um protocolo desenvolvido pela empresa aplicado a análise quantitativa de fases polimórficas por difração de raios-X por policristais, denominado “Comparação Interlaboratorial para caracterização e quantificação de fases polimórficas de insumos farmacêuticos ativos por difração de raios X por policristais”. As atividades da iniciativa geraram uma publicação internacional intitulada “*Validation of the method of quantitative phase analysis by X-ray diffraction in API: case of Tibolone*” (ANEXO D).

Por fim, a última iniciativa está associada à competências em química do estado sólido em fármacos e em ciência de dados. Visto que o entendimento do comportamento no estado sólido de todos os componentes de um medicamento (IFAs e excipientes) é essencial para o desenvolvimento de medicamento e para a farmacovigilância, a NanoBusiness deu início a um projeto inovador objetivando proporcionar aos especialistas oriundos das mais diversas áreas um acesso rápido, fácil e o mais completo possível às informações necessárias e pertinentes para a tomada de decisão relacionada ao polimorfismo em IFAs e às propriedades pertinentes à avaliação do risco sanitário associado. Essa atividade de pesquisa e inovação possui como plataforma inicial o “Banco de Dados e-Diffraction Pharma”, o primeiro banco de dados privado nacional dedicado a caracterização de polimorfismo em fármacos desenvolvido pela NanoBusiness, onde estão armazenadas informações cruciais para as análises qualitativa e quantitativa de polimorfismo, aplicadas ao controle de qualidade de IFAs, medicamentos e excipientes. Atualmente, esse banco de dados dispõe de mais de 350 princípios ativos catalogados e armazenados, em suas inúmeras formas polimórficas, sendo constante a sua atualização. O projeto conceitual dessa iniciativa receberá apoio preliminar do CNPq/ANVISA e Inmetro.

2.2.4

Patentes de invenção relativas a polimorfos

Como destacado anteriormente, o progressivo ganho de importância da temática da química do estado sólido na indústria farmacêutica, sobretudo no que tange ao controle e monitoramento do polimorfismo, foi diretamente influenciado não somente pela identificação de problemas de produção e na eficácia de medicamentos, mas também pelo avanço das disputas patentárias envolvendo formas polimórficas de IFAs (BERNSTEIN, 2002; BHARDWAJ, 2016). Impulsionadas sobretudo pela difusão dos medicamentos genéricos, tais disputas geraram importantes discussões no âmbito das legislações nacionais e internacionais em matéria de propriedade intelectual, sendo marcadas por grandes controvérsias.

De modo geral, a temática do patenteamento de formas polimórficas é marcada por pressões de duas forças principais: por um lado, grandes empresas farmacêuticas inovadoras utilizam patentes de polimorfos como um meio de prolongar seus direitos exclusivos associados a seus medicamentos cujas patentes aproximam-se do prazo de expiração; por outro, fabricantes de medicamentos genéricos buscam desenvolver novas formas polimórficas de modo a contornar patentes vigentes, além de buscarem contrapor os esforços de extensão das grandes empresas inovadoras para que possam comercializar versões genéricas de seus produtos (BERNSTEIN, 2002; LAIRD, 2004).

Para que se possa compreender melhor esta problemática do patenteamento de invenções relacionadas a formas polimórficas, mostra-se relevante apresentar algumas considerações acerca dos principais conceitos envolvidos. Inicialmente, deve-se ter em mente que uma patente pode ser conceituada como um direito que dá ao seu titular a exclusividade da exploração temporária (geralmente 20 anos) sobre uma determinada invenção em um determinado território. A patente assegura o direito temporário do titular de impedir que terceiros possam produzir, usar, colocar à venda, vender ou importar, sem o seu consentimento, produtos ou processos protegidos pelo instrumento em um determinado país. Em contrapartida pelo direito exclusivo, o conteúdo técnico da matéria protegida pela patente deve ser revelado tornando-se público (BARBOSA, 2014).

Cada país ao redor do mundo possui sua própria legislação nacional em matéria de patentes. Entretanto, a legislação da maior parte dos países segue os mesmos princípios por estes serem signatários de tratados internacionais² específicos sobre o tema como o Acordo sobre Aspectos dos Direitos de Propriedade Intelectual Relacionados ao Comércio (TRIPs - *Trade-Related Aspects of Intellectual Property Rights*). Em seu artigo 27, o Acordo TRIPs estabelece que qualquer invenção, seja de produto ou de processo, em todos os setores tecnológicos, será passível de patenteamento desde que atenda três requisitos: seja nova, envolva um passo inventivo e seja passível de aplicação industrial. Entretanto, os países-membros do Acordo possuem liberdade para interpretar tais requisitos de patenteabilidade segundo suas próprias políticas nacionais, bem como outros critérios como o de suficiência descritiva e de melhor forma de execução (BRASIL, 2013).

Assim, invenções relacionadas ao estado sólido - como no caso das formas polimórficas - enquadram-se no âmbito do mesmo princípio de todas as outras passíveis de proteção por meio de patente. Tais invenções podem envolver a descoberta de uma nova forma cristalina de um composto já conhecido que apresente propriedades desejadas para um determinado uso, ou ainda um novo processo para produção de um polimorfo. Entretanto, a temática do patenteamento de formas polimórficas envolve algumas questões específicas já que polimorfos não se diferem em sua composição química, mas somente em sua estrutura cristalina e, conseqüentemente, em seus parâmetros físico-químicos. Além disso, processos de produção de novas formas sólidas podem apresentar somente pequenas variações em relação a processos anteriores, não apresentando portanto grande distintividade (BYRN et al., 2017).

Desse modo, são observadas constantes controvérsias no que tange ao patenteamento de invenções relacionadas à formas polimórficas, tendo em vista as diferentes interpretações dos requisitos de patentabilidade (novidade, atividade inventiva e aplicação industrial) observadas em cada país. No que se refere a novos processos para obtenção de polimorfos, poucas controvérsias são observadas desde que se atendam os requisitos supracitados, bem como os critérios de suficiência descritiva e melhor forma de execução. Por outro lado, são

² Outros exemplos de tratados internacionais em matéria de propriedade intelectual são a Convenção da União de Pais (CUP) e o Tratado de Cooperação em Matéria de Patentes (PCT).

observadas grandes discussões quando se considera o patenteamento de novas formas polimórficas em si, sobretudo em relação a novas formas cristalinas de substâncias já conhecidas, já em domínio público ou não (BRASIL, 2013).

Ao discutirem o tema sob a perspectiva da legislação brasileira, LIMA & SILVEIRA (2011) abordam os principais aspectos levantados no debate em relação ao patenteamento ou não de polimorfos. Em primeiro lugar, as autoras indicam que há o questionamento se a maioria das patentes de formas polimórficas preenche o requisito de novidade, tendo em vista a grande carência de informações no estado da técnica que impediria que os interessados constatassem se o polimorfo já havia sido produzido antes e, portanto, estaria apenas sendo caracterizado no pedido de patente, ou se de fato evolveria um material novo.

Além disso, LIMA & SILVEIRA (2011) também ressaltam a controvérsia oriunda do fato de que o polimorfismo é uma propriedade intrínseca e natural de algumas substâncias, o que leva algumas correntes a acreditarem se tratar portanto de uma mera descoberta e não uma invenção ou criação humana. Segundo esta linha, a possibilidade de descobrir diferentes cristais seria óbvia quando o polimorfismo é encontrado, sendo esta uma propriedade intrínseca da matéria em seu estado sólido. Desse modo, embora um novo processo de obtenção de uma forma polimórfica poderia ser passível de patenteamento, o seu produto resultante, ou seja, a forma polimórfica em si, não seria patenteável por se tratar de uma descoberta. Segundo as autoras:

Assim, os opositores do patenteamento de polimorfos argumentam que sua produção se constitui em mera sequência de tentativas e erros, sem exigência de criatividade humana (INPI, 2007). Por outro lado, a obtenção de alguns desses materiais - tais como os mais estáveis, alguns solvatos, e o menos estável, ou seja, o amorfo - seria óbvia para um técnico no assunto, também carecendo de criatividade. tal fato é agravado pela impossibilidade de antecipação da maioria das propriedades de um polimorfo (LIMA & SILVEIRA, 2011, p. 1527).

Por outro lado, os defensores do patenteamento de polimorfos ressaltam a relevância dos esforços que resultam na obtenção de formas polimórficas que resolvem problemas técnicos específicos e possuiriam grande atividade inventiva em sua concepção. Desse modo, a depender da interpretação adotada por cada país quanto aos requisitos de patenteabilidade, existem países mais propensos a adotarem posicionamento menos restritivos e mais restritivos quanto ao

patenteamento de formas polimórficas. Tal fato ressalta a grande controvérsia envolvida na temática, a qual ganha maior relevância ao passo em que a indústria farmacêutica vem usando agressivamente o sistema de patentes para proteção não apenas de novas formas polimórficas, mas também de outras pesquisas triviais e incrementais, tendo como objetivo ampliar seu período de exclusividade de comercialização (BYRN et al., 2017; LIMA & SILVEIRA, 2011).

No Brasil, ainda que seja possível proteger as novas formas polimórficas, o exame de mérito realizado pelo INPI para estas invenções é rigoroso e o pedido deverá atender aos seguintes requisitos:

- i) Apresentação da caracterização da forma cristalina através de técnicas de caracterização físico-química de sólidos. Neste caso, a difração de raios X de monocristal é a técnica indicada e mais precisa para a caracterização da estrutura cristalina do novo polimorfo;
- ii) Caso não seja viável a obtenção dos dados por difração de raios X de monocristal, deve-se utilizar a técnica de difração de raios X pelo método do pó com metodologia para indexação de picos, bem como a outros métodos de identificação físico-química de sólidos como, por exemplo, DSC e RMN;
- iii) Com o objetivo de garantir a reprodutibilidade na produção do novo polimorfo é exigido a apresentação de todos os parâmetros do processo e etapas de cristalização para obtenção da nova forma cristalina;
- iv) Apresentação de dados empíricos que evidenciem que a nova forma cristalina é diferente das demais formas cristalinas já existentes. Nesse caso, deve-se comparar a diferença da taxa de dissolução do novo polimorfo com a taxa de dissolução da forma cristalina disponível no mercado; e
- v) Especificar a aplicação da nova forma cristalina e como esta nova estrutura cristalina deve ser utilizada no mercado farmoquímico e farmacêutico

Naturalmente, os requisitos supracitados são necessários no processo de pedido de proteção por patentes de uma nova forma polimórfica, mas não suficientes para que uma patente de invenção seja concedida. A referida invenção deve atender ainda aos requisitos de novidade, atividade inventiva e aplicação industrial.

3

Ciência dos dados na identificação de competências científicas

A quantidade de dados gerados no mundo tem aumentado de forma expressiva à medida em que cresce o número de máquinas, pessoas, coisas e processos conectados. Segundo estimativas da IBM, 90% dos dados digitais hoje disponíveis foram gerados somente nos últimos dois anos, sendo que 2.5 quintilhões de bytes são criados todos os dias. Esses dados são gerados a partir de diversas fontes como dispositivos móveis, sensores, redes sociais, transações comerciais, sinais de GPS, entre outras (IBM, 2017).

Neste contexto marcado pelo grande e crescente volume de dados disponíveis, empresas ao redor do mundo passaram a direcionar seus esforços para exploração de dados como vantagem competitiva. Ainda que em intensidades e formas distintas, esse processo pode ser observado em praticamente todos os setores industriais.

O volume e a variedade dos dados atualmente disponíveis extrapolam a capacidade de análise manual e desafiam as tecnologias de armazenamento convencionais. Desse modo, avanços importantes têm ocorrido na capacidade de processamento e armazenamento dos computadores e na infraestrutura de comunicação e de redes sem fio agora amplamente disponíveis. Além disso, modelos estatísticos e algoritmos específicos passaram a ser desenvolvidos, permitindo maior integração entre bases de dados distintas e uma análise mais profunda e ampla dos dados disponíveis.

Assim, na esteira desse processo e partir da convergência dos fenômenos supracitados, a Ciência dos Dados tem ganhado cada vez mais relevância na atualidade. O conceito de Ciência dos Dados pode ser definido de maneira ampla como uma área de estudo multidisciplinar que envolve os campos da estatística, ciência da computação e teoria da informação, de modo a permitir a exploração de quantidades massivas de dados estruturados e não estruturados para geração de insights, informações e conhecimentos relevantes.

A abordagem da Ciência dos Dados oferece soluções alternativas e inovadoras para solução de problemas e auxílio na tomada de decisão por meio de procedimentos analíticos como a mineração de dados, mineração de textos, aprendizado de máquina, entre outros. O avanço da Ciência dos Dados foi acompanhando ainda pelo surgimento de diversos conceitos e abordagens nos últimos anos como *Big Data*, *Data Analytics* e *Internt of Things* (IoT). Como já destacado anteriormente, o surgimento de tais abordagens encontra base comum no grande volume e variedade de dados hoje existentes e gerados a partir de diversas fontes.

O conhecimento científico não é exceção a esse intenso processo de geração de grandes volumes de dados. Somente a base de dados *Web of Science* (WoS) mantida pela Clarivate Analytics possui atualmente mais de 68 milhões de registros bibliográficos de publicações em diversas áreas do conhecimento. A cada ano, o número de publicações indexadas cresce e apenas em 2016 totalizou mais de 2.8 milhões de novos registros.

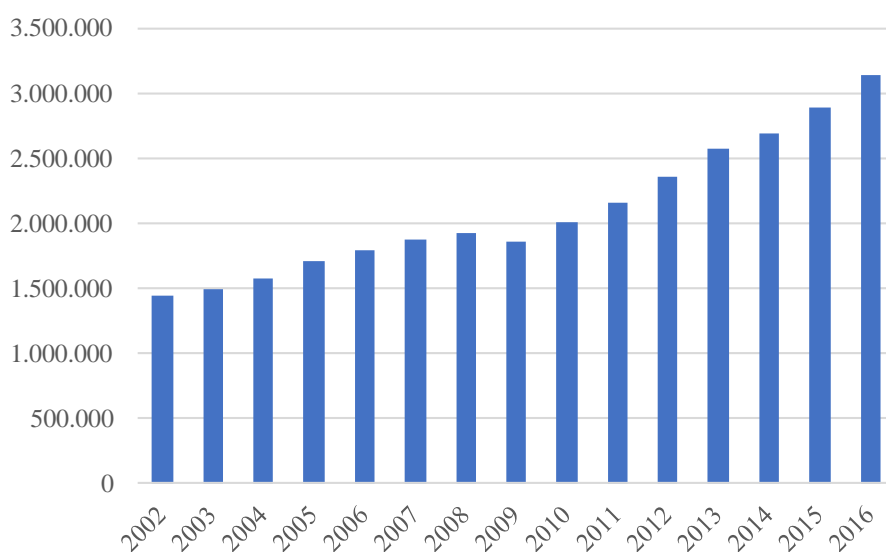


Figura 5 – Crescimento anual de registros indexados na base de dados Web of Science.

Do mesmo modo, o número de depósito de patentes também tem apresentado um crescimento expressivo. Somente em 2016, aproximadamente 3.2 milhões de patentes foram depositadas no mundo, um aumento de 8.3% em relação à 2015 (WIPO, 2017).

Neste contexto, com a crescente introdução de novos conceitos, invenções, conhecimentos, tecnologias e inovações por meio de publicações de todos os tipos (científica e não científica), incluindo artigos científicos e depósitos de patente, bem como sua indexação em bases de dados integradas e de ampla abrangência, acompanhar o desenvolvimento científico, tecnológico e comercial torna-se cada vez mais complexo, tanto para empresas quanto para instituições de pesquisa como universidades (MILES et al. 2016). Assim, as ferramentas e estratégias da Ciência dos Dados oferecem também alternativas interessantes para lidar com o vasto volume de conhecimento científico gerado atualmente.

3.1

Extraindo informações do conhecimento científico e tecnológico

Embora o conceito de Ciência dos Dados seja em si mais recente, abordagens específicas para busca, organização, processamento e extração de informações de literatura científica e tecnológica possuem longa tradição. Em especial, destacam-se a mineração de dados e de texto, procedimentos analíticos utilizados na elaboração de indicadores de ciência e tecnologia (C&T) e monitoramento tecnológico tendo como base publicações científicas e documentos de patentes. No entanto, são diversas as abordagens direcionadas para o propósito específico da análise do conhecimento científico.

Como destacam Miles et al. (2016), os termos frequentemente usados para descrever meios de se lidar com o grande volume de dados e informações científicas e tecnológicas incluem bibliometria, cientometria, mineração de dados, mineração de textos, mineração tecnológica e Big Data. Essas abordagens auxiliam na identificação de padrões e tendências no desenvolvimento do conhecimento científico e tecnológico, contribuindo para o processo de tomada de decisão. A seguir, são apresentados alguns conceitos referentes aos termos citados:

Quadro 2 – Definições de termos comumente utilizados.

Termo	Definição
Big Data Analytics	Desenvolvimento e investigação de novos métodos e algoritmos capazes de detectar padrões úteis ou correlações presentes em grandes volumes de dados.
Bibliometria	Utilização de análises quantitativas para medir publicações e resultados científicos. Comumente utilizado de maneira indistinta com o termo cientometria.
Cientometria	Análise matemática e estatística da ciência, tecnologia e inovação com o objetivo de avaliar e compreender citações científicas, mapear áreas científicas, identificar redes de pesquisa, medir o impacto de autores, artigos, revistas e instituições e gerar indicadores para definição de políticas públicas e apoio a tomada de decisão gerencial.
Mineração de dados	O processo computacional de descobrimento de padrões em grandes conjuntos de dados envolvendo métodos relacionados as áreas de inteligência artificial, aprendizado de máquina, estatística e sistemas de bases de dados.
Mineração de texto	Análise semântica de conteúdo de textos em linguagem natural para determinar o objetivo ou sentido de determinada comunicação e extrair informações úteis. A mineração de textos é considerada a aplicação da mineração de dados em textos.
Mineração Tecnológica	Mineração tecnológica refere-se a aplicação da mineração de textos à informações de ciência, tecnológica e inovação. Ela explora bases de dados de textos para extrair conhecimentos acerca de tecnologias emergentes.

Fonte: Adaptado de Miles et al. 2016.

Embora exista diferenças conceituais entre os termos, geralmente observa-se na literatura a utilização de alguns destes de maneira indistinta, em especial os conceitos de bibliometria e cientometria. Como destacam Miles et al. (2016), um fator que determina a escolha da abordagem é o fato do dado a ser analisado estar estruturado (ex.: publicações científicas indexadas e patentes), semiestruturado (ex.: dados de redes sociais) ou não-estruturado (sites, artigos de jornais e outros dados digitais). Enquanto os dados estruturados podem ser abordados por meio da análise quantitativa e estatística, dados não-estruturados são usualmente abordados por meio da análise semântica e ferramentas linguísticas.

Ao abordar esta temática, Cantú-Ortiz (2017) introduz ainda os conceitos de *Research Intelligence* e de *Research Analytics*. Segundo o autor:

Research intelligence pursuits compare metrics of achievement to produce rankings of various types, benchmarking dossiers, or scientific trend statements to offer feedback to those actors responsible for assessment of outcomes, appraisal of throughput, or strategy amendment. Research Analytics is understood as the

confluence of Data Analytics applied to scientific products found in Bibliometric databases and using Scientometrics and Research Intelligence methodologies in both prospective and prescriptive ways (CANTÚ-ORTIZ, 2017, p. 2)

Assim, o conceito de *Research Analytics* retrata uma área multidisciplinar e emergente que combina métodos e ferramentas como inteligência artificial, aprendizado de máquina, estatística e ciência da computação para medir o impacto da ciência e tecnologia. Trata-se, portanto, da sinergia entre a Ciência dos Dados, Bibliometria, Cientometria e *Research Intelligence*, tendo como objetivo a extração de insights, informações e conhecimentos a partir do grande e crescente conjunto de dados científicos disponíveis (CANTÚ-ORTIZ, 2017).

Não obstante a diversidade de conceitos existentes, deve-se ter em mente que o objetivo comum das diferentes abordagens reside em auxiliar na identificação de padrões e tendências no desenvolvimento do conhecimento científico e tecnológico, contribuindo para o processo de tomada de decisão. Para fins da presente pesquisa, optou-se pela utilização do termo bibliometria englobando o conceito da cientometria.

3.2

Uso de indicadores bibliométricos de C&T

Como já destacado, acompanhar o desenvolvimento científico, tecnológico e comercial mostra-se fundamental hoje para diversos fins como identificar competências científicas em diferentes áreas, gerar vantagem competitiva para empresas que queiram manter-se na vanguarda tecnológica, aprimorar a definição de políticas públicas e a alocação eficiente de recursos em países, avaliar os impactos sociais das atividades de pesquisa científicas, compreender o avanço científico e identificar desafios e tendências futuras, entre outros.

De modo geral, as publicações científicas configuram-se como o principal canal pelo qual pesquisadores divulgam suas descobertas sendo, portanto, um elemento central na geração e disseminação de conhecimento. As patentes, por sua vez, refletem os avanços tecnológicos e retratam a capacidade de se transferir resultados de pesquisas em aplicações, congregando grande parte dos conhecimentos tecnológicos já desenvolvidos.

Dados de publicações científicas e de documentos de patente podem ser analisados por meio de ferramentas de bibliometria e ciência dos dados, permitindo a elaboração de indicadores de ciência e tecnologia (C&T) e o monitoramento tecnológico, além da avaliação detalhada do avanço e das particularidades do conhecimento em diferentes áreas. Como destacam Moed et al. (2004), a elaboração e o uso desses indicadores quantitativos de C&T possui grande relevância e tradição, sendo observados avanços contínuos nas ferramentas para análise e extração de informação dos mesmos.

Ao abordar o tema, Gregolin et al. (2005) ressaltam que os indicadores quantitativos:

[...] podem ser compreendidos como dados estatísticos usados para medir algo intangível, que ilustram aspectos de uma realidade multifacetada. A construção e o uso de indicadores de produção científica são objeto de estudo de várias áreas do conhecimento, sendo usados tanto para o planejamento e a execução de políticas para o setor como também para que a comunidade científica conheça melhor o sistema no qual está inserida.

O uso de indicadores de C&T fornece uma base importante para as discussões que envolvem os aspectos políticos, sociais e econômicos do desenvolvimento de novas tecnologias. Dados de depósitos de patente podem ser utilizados para identificar o posicionamento e o foco de atuação de países, além de fornecer indícios do seu estágio de desenvolvimento e especificidades. Já os dados de publicações científicas podem contribuir para um entendimento mais robusto do avanço em determinada área, evidenciando as competências específicas de países e os principais centros geradores de conhecimento ao redor do mundo, os quais desempenham um papel importante na transferência de tecnologias para aplicações industriais.

Milanez (2015) destaca que países líderes mundiais em pesquisa científica e desenvolvimentos tecnológicos como EUA, Japão e Alemanha, possuem uma preocupação constante com a elaboração e análise de indicadores quantitativos para apoio na definição de políticas em ciência e tecnologia. O autor enfatiza ainda o forte investimento de recursos em órgãos que realizam estudos para o estabelecimento de políticas em C&T como o *Science and Technology Policies Group* da OECD, o *National Science Board* (divisão do *National Science Foundation* dos EUA) e o *Observatoire des Sciences et des Techniques* da França,

os quais tiveram orçamentos somados em 2013 na ordem de 22 milhões de dólares.

Além de auxiliar países na definição de políticas de ciência e tecnologia, a elaboração e análise destes indicadores também representa uma estratégia de vantagem competitiva para empresas. Ao integrar atividades de monitoramento de indicadores de C&T, torna-se possível avaliar a evolução recente de tecnologias de interesse, identificando oportunidades e ameaças tecnológicas, potenciais parceiros para projetos, principais concorrentes, principais aplicações das tecnologias, entre outras informações relevantes para a tomada de decisão estratégica e gerencial.

Frente ao crescente volume de dados científicos e tecnológicos e sua indexação em bases de dados online e de ampla abrangência, a bibliometria tem sido aplicada para mineração de dados e de textos (*data mining e text mining*) de publicações científicas, documentos de patentes e outras fontes, permitindo a elaboração de indicadores quantitativos a partir dos dados bibliográficos indexados nessas bases (MILANEZ, 2015).

3.3

Identificação de lideranças e competências científicas

Dentre as possíveis soluções apresentadas a partir da geração de indicadores de C&T e a aplicação de ferramentas de Ciência dos Dados está a identificação de competências específicas em determinada área do conhecimento. Identificar líderes na geração de conhecimento e tecnologia em determinada área pode desempenhar um papel importante na definição de parcerias estratégicas, seja na criação de redes de colaboração entre pesquisadores e instituições acadêmicas ou na identificação de potenciais parceiros para atividades de pesquisa e desenvolvimento de empresas.

Neste aspecto, deve-se ressaltar que as empresas adotam cada vez mais modelos abertos de gestão da inovação, incorporando conhecimentos externos para impulsionar suas atividades internas. Assim, a identificação de competências externas mostra-se fundamental, sendo importante o desenvolvimento de ferramentas que auxiliem esse processo. Do mesmo modo, a identificação de competências científicas e tecnológicas específicas pode auxiliar tomadores de

decisão a direcionarem maiores recursos para determinadas instituições, além de permitir a definição de políticas científicas e tecnológicas mais assertivas.

A avaliação do posicionamento comparativo de universidades, autores e revistas são cada vez mais comuns. No entanto, observa-se uma relativa falta de consenso quanto as abordagens a serem empregadas e até mesmo críticas quanto a proposta de se criar esses tipos de classificações (ADLER et al., 2008; OSTERLOH & FREY, 2009). De modo geral, as abordagens mais difundidas para identificação de competências e liderança baseiam-se em métodos bibliométricos de contagem de publicações e/ou citações recebidas.

Além disso, a maioria dos estudos geralmente agregam a ciência como um todo, não buscando uma distinção mais detalhada por áreas da ciência. Como consequência, instituições maiores tendem a ocupar posições elevadas nos rankings em função de sua maior produção científica como um todo. Quando alguma distinção é feita no que se refere áreas específicas do conhecimento, geralmente são utilizados como base sistemas de classificação de revistas oferecidos por bases de dados como Scopus e Web of Science (WoS).

Ao discutir o tema, Klevans & Boyack (2010) apresentam uma visão crítica quanto a tal abordagem, especialmente no que se refere a distinção de áreas com base nos sistemas de classificação de revistas científicas. Os autores destacam que a categorização por disciplinas utilizadas nesses sistemas não consegue distinguir as atividades multidisciplinares únicas dos pesquisadores retratadas em suas publicações. Desse modo, entende-se que haja uma dificuldade em identificar competências e lideranças específicas na medida em que os sistemas de classificação de áreas e disciplinas não capturam a essência da multidisciplinaridade. Para abordar esse problema, os autores propõem um novo método de classificação baseado na análise de co-citações para identificar clusters que, por sua vez, retratariam de forma mais acurada a área do conhecimento (retratada com paradigmas) quando comparada aos sistemas tradicionais. Embora forneça uma alternativa interessante, tal abordagem possui ainda limitações importantes. Em especial, destaca-se o fato de utilizar somente informações de publicações científicas, não contemplando dados de depósito de patente ou outras fontes.

Um aspecto importante e geralmente pouco abordado quando se busca identificar competências e liderança em determinada área envolve a adequação da

análise aos objetivos específicos e das respostas que se buscam. Desse modo, deve-se levar em conta que somente a liderança quantitativa, seja no número de publicações ou citações, não necessariamente retrata a competência a ser buscada. Uma empresa, por exemplo, pode considerar como variável relevante na identificação de competências a capacidade do pesquisador ou instituição de transformar seu conhecimento em aplicações tecnológicas. Desse modo, embora o número de publicações e citações seja importante, o número de patentes depositadas também poderia ser incorporado ao modelo, fornecendo um resultado muito mais adequado às suas intenções.

Ao avaliar a abordagem da bibliometria tradicional na avaliação da ciência, Bornmann & Leydesdorff (2014) destacam ainda que alguns dos indicadores comumente utilizados em análises apresentam limitações relevantes. Como exemplo, os autores citam o *Journal Impact Factor*, indicador mais conhecido para avaliar a relevância de revistas científicas, argumentando que:

“Since it gives the average number of citations for the papers in a journal during the preceding 2 years, it may be determined by a few highly cited papers and hardly at all by the mass of papers, which are cited very little or not at all (BORNMANN & LEYDESDORFF, 2014, p. 1229).

Outro exemplo citado pelos autores é do h-index, um indicador bibliométrico para quantificação da produtividade e impacto de publicações. Em sua essência, o h-index busca medir as publicações em um conjunto a partir de um número mínimo específico de citações (identificado pelo h), de modo que as poucas publicações com muitas citações possuem influência reduzida nos cálculos. Embora os autores ressaltem quem o h-index não seja afetado pelo mesmo problema de cunho estatístico do *Journal Impact Factor* por levar em conta o número mínimo de citações dentro do conjunto de publicações, outros fatores limitantes devem ser levados em conta. Como destacam Bornmann & Leydesdorff (2014, p. 1229-1230):

The h-index, however, has other weaknesses that make its use in research evaluation questionable; the arbitrary limit for the selection of the significant publications with at least h citations is criticized; it could just as well be h^2 citations.

De modo geral, tanto as abordagens supracitadas quanto variações e aprimoramentos posteriores como o g-index, tem como foco uma única classificação, de um único tipo de objeto de análise, tanto revistas ou pesquisadores (BOUYSSOU & MARCHANT, 2010; WOEGINGER, 2008). Como pode ser observado, muitos desafios ainda existem na identificação de competências específicas e lideranças. Tendo em vista essa visão, a presente pesquisa busca investigar as tendências e os desafios científicos e tecnológicos na área de química do estado sólido gerando inicialmente indicadores de C&T por meio da bibliometria associada a ferramentas de ciência dos dados para identificação de competências específicas na área.

4

Metodologia

Neste capítulo são apresentados os procedimentos metodológicos propostos para a investigação de indicadores científicos e tecnológicos na área de química do estado sólido utilizando ferramentas de ciência dos dados a partir de publicações científicas e depósitos de patentes. Além disso, foi aplicado um modelo para avaliação de competências específicas de modo a identificar atores (pesquisadores ou instituições) líderes em determinadas áreas da química do estado sólido no Brasil e no mundo, tendo como foco o estudo de caso de uma nova forma polimórfica descoberta para o IFA acetato de dexametasona.

Por meio desse modelo foi aprimorado o processo de busca de parceiros estratégicos, favorecendo o desenvolvimento de projetos conjuntos na área, contribuindo para o aumento da competitividade da indústria farmacêutica e farmacêutica brasileira e em outros mercados emergentes.

A metodologia proposta para o desenvolvimento da tese baseia-se em três etapas que envolvem a elaboração de indicadores de C&T por meio da bibliometria associada a ferramentas de ciência dos dados, conforme apresentado na Figura 6.

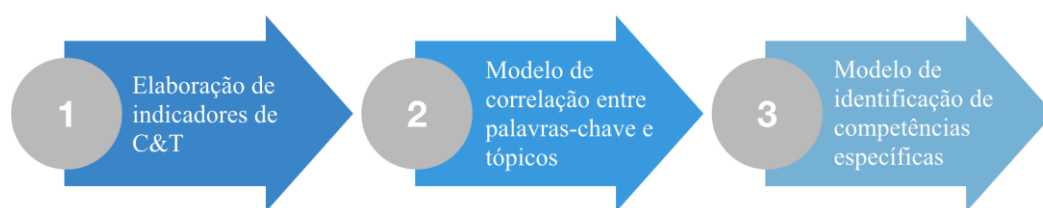


Figura 6 – Etapas da metodologia proposta.

A primeira etapa da metodologia envolve a aplicação de ferramentas de bibliometria aplicadas para mineração de dados e mineração de textos (*data mining e text mining*) para busca, recuperação, processamento e análise de dados de publicações científicas na área de química do estado sólido na indústria farmacêutica. Nesta etapa, o cenário atual da pesquisa na área é analisado a partir da elaboração de indicadores de C&T a partir de publicações científicas e

depósitos de patentes, permitindo a investigação de tendências e desafios científicos e tecnológicos a nível mundial e também no Brasil.

Na segunda etapa, um modelo baseado no artigo submetido a publicação para revista *Expert Systems with Applications* e intitulada “*Towards a Systematic Methodology to Associate Research-Promoting Entities with Industrial Sectors: Application to the Nanotechnology Area*”, tech.memo, No02/18, DEI/PUC-Rio, Rio de Janeiro, 2018 e aprovado no exame de qualificação de doutorado do autor dessa tese (**APENDICE B**), é aplicado para medir a correlação entre as principais palavras-chave identificadas nas publicações científicas analisadas e grandes tópicos dentro da área de química do estado sólido. A avaliação da relação entre as palavras-chave e os tópicos é realizada por um grupo de especialistas. Esta etapa tem como objetivo permitir a classificação dos pesquisadores e instituições nos grandes tópicos de química do estado sólido a partir da utilização de ferramentas de *text mining* no conjunto de publicações científicas.

Por fim, a terceira etapa envolve a aplicação de um modelo multivariado e não genérico para avaliação de liderança e competências em áreas de aplicação específica dentro da química do estado sólido definidas nos grandes tópicos da etapa anterior. O modelo permite a definição de pesos para cada uma das variáveis incorporadas, permitindo sua adaptação aos interesses específicos de sua aplicação.

4.1

Etapa 1: Procedimento geral para Elaboração de indicadores de C&T

Esta seção apresenta os procedimentos metodológicos envolvidos na elaboração de indicadores de C&T de publicações científicas e depósitos de patentes na área de química do estado sólido, tendo como objetivo mapear os avanços recentes e investigar tendências e desafios científicos e tecnológicos a nível mundial e também no Brasil. Para tanto, são utilizadas ferramentas de bibliometria com aplicações voltadas para mineração de dados e de textos (*data mining e text mining*). O detalhamento da Etapa 1 é apresentado em três subseções, sendo a primeira dedicada a definição da estratégia de busca de dados empregada, a segunda ao processamento dos dados bibliográficos e a terceira ao processamento e cálculos aplicados.

A estratégia proposta para o desenvolvimento deste estudo envolve as etapas de busca, processamento e elaboração de indicadores de publicações científicas e depósitos de patentes relacionadas à química do estado sólido na área farmacêutica. O recorte temporal da pesquisa composto por 20 anos foi definido em função do expressivo avanço desse campo de pesquisa no período, permitindo a análise de um amplo conjunto de dados.

Para fornecimento dos dados bibliográficos de publicações científicas, foi utilizada a base *Web of Science* (WoS) disponível no Portal de Periódicos da Capes. Além de sua amplitude e da possibilidade de realização de buscas complexas, com diversos operadores booleanos e a partir de uma abordagem modular, a *Web of Science* fornece ainda uma estrutura de classificação própria de revistas e áreas do conhecimento com contribuições importantes para o desenvolvimento desta tese. Avaliando-se outras bases disponíveis como a *Scopus*, concluiu-se que a base de dados multidisciplinar *Web of Science* é suficiente e apropriada para a presente pesquisa.

Ainda que a base *Scopus* possa apresentar maior abrangência em determinadas áreas (MONGEON & PAUL-HUS, 2015), a utilização da *Web of Science* mostrou-se mais adequada em função de diversos fatores, sobretudo em função de ser utilizada no modelo de identificação de competência (Etapa 3 da metodologia) indicadores como o Journal Citation Reports (JCR) publicado pela Clarivate Analytics, a mesma responsável pela base de dados *Web of Science*, sendo um dos indicadores principais adotados pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) para avaliação de fator de impacto de revistas no Brasil.

Diversos autores já abordaram a problemática do impacto da seleção de bases de dados específicas sobre estudos envolvendo indicadores de C&T, sobretudo envolvendo publicações científicas (MONGEON & PAUL-HUS, 2015; ARCHAMBAULT et al., 2009; GAVEL & ISELID, 2008). De modo geral, embora se reconheça potenciais influências nos resultados em função das bases de dados escolhidas, a utilização tanto da base *Web of Science* quanto da base *Scopus* tem sido considerada adequadas para estudos do gênero (ARCHAMBAULT et al., 2009). Assim, ainda que ambas as bases pudessem ser utilizadas, optou-se nesta pesquisa pela seleção da *Web of Science* tendo em vista os fatores já citados anteriormente.

Por fim, ainda que se possa argumentar em relação a complementaridade das bases *Web of Science* e *Scopus* e a necessidade de utilização de ambas em um mesmo estudo, as diferentes estruturas de indexação de dados encontradas em cada uma torna o processo de fusão de dados complexo e pouco efetivo. Em especial, destaca-se a diferença nos formatos de padronização de nomes de autores e instituições, dificultando o cruzamento das informações das bases para que se obtenha um conjunto de dados conciso e robusto para análise. Desse modo, embora a junção de dados de diferentes bases seja possível para análises mais superficiais, tal abordagem não se mostra efetiva para estudos mais complexos como os apresentados nesta tese justamente pelas inconsistências entre as bases na estruturação dos dados apresentados. Ainda que no futuro tais dificuldades possam ser superadas, destaca-se o fato de que tal problemática não constitui objeto do escopo desta pesquisa.

Também foi feita a seleção de bases para recuperação dos dados bibliográficos de documentos de patentes, tendo sido selecionada a base *Derwent Innovations Index* em função principalmente de sua cobertura mundial e da possibilidade de exportação de grandes conjuntos de registros simultâneos. Tendo em vista que muitas bases de patentes abertas como a do USPTO (Estados Unidos), Espacenet (EPO) e Patentscope (WIPO) não possibilitam a exportação de grandes conjuntos estruturados de dados, a base *Derwent Innovations Index*, única base comercial disponível no Portal de Periódicos da Capes, foi considerada adequada. Neste aspecto, ressalta-se o fato de haver poucas diferenças de cobertura entre as bases de patente citadas, sendo o critério mais relevante para a seleção da *Derwent Innovations Index* a possibilidade de exportação de grandes conjuntos de dados bibliográficos de maneira estruturada.

4.1.1

Estratégia de busca

Para recuperação dos dados das bases de publicações científicas e depósitos de patentes, foi desenvolvida uma expressão de busca específica para química do estado sólido na indústria farmacêutica. Tendo em vista as particularidades de cada tipo de busca, foram realizadas adaptações nas estratégias para as bases de dados de publicações científicas e de documentos de patente.

A busca e recuperação de registros bibliográficos de publicações científicas na base de dados *Web of Science* envolveu inicialmente uma ampla busca por publicações científicas a partir de palavras-chave relacionadas à química do estado sólido na área farmacêutica. Tendo sido identificado o conjunto de publicações para a estratégia empregada, foi realizada em um segundo momento uma etapa de exclusão de publicações identificadas na primeira etapa, mas que não tinham relação direta com o escopo desta pesquisa.

Para tanto, foram excluídas publicações que continham palavras-chave específicas como “*polymorphisms*”, “*genetic*” e DNA, as quais estão geralmente associadas ao fenômeno do polimorfismo genético e, portanto, não relacionadas ao tema deste estudo. Além disso, foi feito um refinamento excluindo-se publicações indexadas na base de dados relacionadas a áreas de pesquisa não contempladas no escopo da tese. Este refinamento teve como base as informações do campo “*Categoria Web of Science*”.

A definição das palavras-chave empregadas na estratégia de busca, bem como a seleção das categorias *Web of Science* na etapa de exclusão, foram realizadas a partir do levantamento e discussão com 20 especialistas em química do estado sólido. Além disso, foi empregado um processo iterativo para avaliação da estratégia de busca caracterizado por rodadas de busca e avaliação crítica de resultados com o auxílio dos especialistas até a definição da estratégia final. A título de exemplificação, pode-se citar o caso do termo “*Crystal structure*” que, embora estivesse inicialmente contido na estratégia, foi retirado posteriormente por apresentar uma contribuição expressiva para recuperação de registros bibliográficos destoantes do escopo do objeto de estudo da pesquisa. Desse modo, mesmo sendo um termo relevante para a química do estado sólido, sua inclusão traria mais ruído do que benefícios para a análise pretendida.

A partir da aplicação da estratégia de busca, foram identificados 7.200 registros bibliográficos de publicações científicas na área de química do estado sólido no período de 1997 a 2016. O detalhamento da estratégia empregada é apresentado abaixo:

#	Query de busca para publicações científicas
1	TS=(pharmaceutic* OR Drug* OR "Active pharmaceutical ingredient*")
2	TS=("crystal* form\$" OR "crystal* phase\$" OR polymorph* OR Cocrystal\$ OR Co-crystal\$ OR "solid form\$" OR "solid-state" or "amorph* phase\$" OR "amorph* form\$" OR "amorph* solid")
3	#2 AND #1
4	#3 NOT TS=(gene\$ or genetic\$ or dna or nucleotide or transcriptase or polymorpha or polymorphonuclear or "NATURAL POLYMORPHISM" or mRNA or rna or SNPs or virus or "polymorphisms")
5	#4 NOT WC=(BIOCHEMISTRY MOLECULAR BIOLOGY OR CHEMISTRY MEDICINAL OR BIOPHYSICS OR CHEMISTRY INORGANIC NUCLEAR OR BIOCHEMICAL RESEARCH METHODS OR CELL BIOLOGY OR BIOTECHNOLOGY APPLIED MICROBIOLOGY OR MEDICINE RESEARCH EXPERIMENTAL OR MATERIALS SCIENCE BIOMATERIALS OR CARDIAC CARDIOVASCULAR SYSTEMS OR MICROBIOLOGY OR TOXICOLOGY OR FOOD SCIENCE TECHNOLOGY OR MATHEMATICAL COMPUTATIONAL BIOLOGY OR ENGINEERING BIOMEDICAL OR NEUROSCIENCES OR ONCOLOGY OR IMMUNOLOGY OR PSYCHIATRY OR INFECTIOUS DISEASES OR MEDICINE GENERAL INTERNAL OR ENDOCRINOLOGY METABOLISM OR BIOLOGY OR CLINICAL NEUROLOGY OR DERMATOLOGY OR PARASITOLOGY OR PERIPHERAL VASCULAR DISEASE OR GENETICS HEREDITY OR ENVIRONMENTAL SCIENCES OR HEMATOLOGY OR MATERIALS SCIENCE CERAMICS OR ENGINEERING ENVIRONMENTAL OR POLYMER SCIENCE OR GASTROENTEROLOGY HEPATOLOGY OR RESPIRATORY SYSTEM OR ELECTROCHEMISTRY OR PHYSIOLOGY OR PLANT SCIENCES OR ENERGY FUELS OR MEDICAL LABORATORY TECHNOLOGY OR PUBLIC ENVIRONMENTAL OCCUPATIONAL HEALTH OR NUTRITION DIETETICS OR TROPICAL MEDICINE OR UROLOGY NEPHROLOGY OR ENGINEERING MECHANICAL OR PATHOLOGY OR RADIOLOGY NUCLEAR MEDICINE MEDICAL IMAGING OR METALLURGY METALLURGICAL ENGINEERING OR MATERIALS SCIENCE COATINGS FILMS OR VIROLOGY OR EDUCATION SCIENTIFIC DISCIPLINES OR SUBSTANCE ABUSE OR ALLERGY OR NUCLEAR SCIENCE TECHNOLOGY OR PEDIATRICS OR SURGERY OR AGRICULTURE MULTIDISCIPLINARY OR VETERINARY SCIENCES OR MATERIALS SCIENCE COMPOSITES OR MATERIALS SCIENCE TEXTILES OR MATERIALS SCIENCE PAPER WOOD OR ANESTHESIOLOGY OR GREEN SUSTAINABLE SCIENCE TECHNOLOGY OR GERIATRICS GERONTOLOGY OR AGRICULTURAL ENGINEERING OR MEDICINE LEGAL OR TRANSPLANTATION OR WATER RESOURCES OR AGRONOMY OR BEHAVIORAL SCIENCES OR MYCOLOGY OR CRITICAL CARE MEDICINE OR MECHANICS OR DENTISTRY ORAL SURGERY MEDICINE OR EMERGENCY MEDICINE OR HEALTH CARE SCIENCES SERVICES OR METEOROLOGY ATMOSPHERIC SCIENCES OR OPHTHALMOLOGY OR ACOUSTICS OR CELL TISSUE ENGINEERING OR ENGINEERING CIVIL OR OBSTETRICS GYNECOLOGY OR PSYCHOLOGY OR ZOOLOGY OR REPRODUCTIVE BIOLOGY OR DEVELOPMENTAL BIOLOGY OR MEDICAL INFORMATICS OR PSYCHOLOGY CLINICAL OR AGRICULTURE DAIRY ANIMAL SCIENCE OR ANATOMY MORPHOLOGY OR ENTOMOLOGY OR HISTORY PHILOSOPHY OF SCIENCE OR MARINE FRESHWATER BIOLOGY OR NURSING OR TELECOMMUNICATIONS OR MINING MINERAL PROCESSING OR ORTHOPEDICS OR ECOLOGY OR ENGINEERING AEROSPACE OR EVOLUTIONARY BIOLOGY OR GEOSCIENCES MULTIDISCIPLINARY OR NEUROIMAGING OR PSYCHOLOGY BIOLOGICAL OR REMOTE SENSING OR ROBOTICS OR SOCIAL SCIENCES BIOMEDICAL OR CONSTRUCTION BUILDING TECHNOLOGY OR ECONOMICS OR LAW OR LIMNOLOGY OR MANAGEMENT OR PRIMARY HEALTH CARE OR PSYCHOLOGY EXPERIMENTAL OR PSYCHOLOGY MULTIDISCIPLINARY OR PSYCHOLOGY SOCIAL OR SOCIAL WORK OR SOIL SCIENCE OR RHEUMATOLOGY OR MINERALOGY OR INTEGRATIVE COMPLEMENTARY MEDICINE OR ENGINEERING MANUFACTURING OR PHYSICS FLUIDS PLASMAS)

Para a busca e recuperação de registros bibliográficos de documentos de patentes, foi utilizada também uma estratégia modular composta de uma busca inicial ampla, seguida da exclusão de registros com os mesmos termos específicos. Além disso, para maior acurácia da estratégia, foi utilizada ainda a Classificação Internacional de Patentes (CIP) tendo sido selecionadas as classificações mais aderentes ao objeto de interesse da pesquisa após levantamento e discussão com especialistas em química do estado sólido. As classificações selecionadas são apresentadas abaixo:

A61K*	Preparações para finalidades médicas, odontológicas ou higiênicas
A61P*	Atividade terapêutica específica de compostos químicos ou preparações medicinais
C30B-029/54	Crescimento de cristais > Monocristais ou material policristalino homogêneo com estrutura definida caracterizados pelo material ou por seus formatos > Compostos orgânicos
B01D-009	Processos ou aparelhos físicos ou químicos em geral > Separação > Cristalização

* Caractere utilizado para truncagem do termo de busca, permitindo a recuperação de todos os grupos e subgrupos abaixo das subclasses da CIP selecionadas (ex.: A61k-009/14)

A partir da aplicação da estratégia de busca, foram identificados 4.665 registros bibliográficos de documentos de patente na área de química do estado sólido na indústria farmacêutica entre 1997 e 2015. O detalhamento da estratégia empregada é apresentado abaixo:

#	Query de busca para documentos de patente
1	TI=("crystal* form\$" OR "crystal* phase\$" OR polymorph* OR Cocrystal\$ OR Co-crystal\$ OR "solid form\$" OR "solid-state" or "amorph* phase\$" OR "amorph* form\$" OR "amorph* solid")
2	IP=(A61K* OR A61P* OR C30B-029/54 OR B01D-009*)
3	#3 NOT TS=(gene\$ or genetic\$ or dna or nucleotide or transcriptase or polymorpha or polymorphonuclear or "NATURAL POLYMORPHISM" or mRNA or rna or SNPs or virus or "polymorphisms")

4.1.2

Processamento dos dados

Tendo sido concluída a etapa de busca das publicações na base de dados *Web of Science* e dos documentos de patentes na base *Derwent Innovations Index*,

foi realizada a recuperação (download) dos registros bibliográficos completos disponíveis para exportação em cada uma das bases de dados. Em seguida, os dados foram importados para o software *VantagePoint* versão 9.5. Esta ferramenta auxilia no processamento dos dados bibliográficos e no processo de mineração de dados e texto (*data mining* e *text mining*), permitindo a análise de diversos indicadores bibliométricos em grandes conjuntos de dados. A utilização do *VantagePoint* auxilia na identificação de padrões no conjunto de dados, identificando relacionamentos entre variáveis. Além disso, as ferramentas disponíveis no software auxiliam a padronização de nomes de autores e instituições por meio de algoritmos de lógica fuzzy, tornado mais eficiente o processo de tratamento de dados.

Após a importação dos registros, foi realizado o processamento dos dados permitindo a limpeza e padronização dos dados. Para os documentos de patente, foram realizados procedimentos para padronização dos nomes dos depositantes de modo a permitir uma análise mais consistente dos resultados. Para as publicações científicas, inicialmente foram excluídos registros duplicados, o que reduziu o conjunto final de análise à 7.191 publicações. Em seguida, foi realizada a limpeza e padronização de nomes de autores e de instituições. Essas etapas são consideradas as mais demoradas no procedimento de mineração de dados e possuem grande relevância para a qualidade final do conjunto de dados a ser analisado (LEOPOLD et al., 2004).

Em especial, para uma análise mais detalhada dos principais tópicos de pesquisa na área de química do estado sólido no período de 1997 a 2016, foi realizado o processamento do conteúdo de dois campos extraídos da base de dados *Web of Science*: “*Keyword (Author’s)*” e “*Keyword Plus*”. Enquanto o primeiro contém as palavras-chave indicadas pelos autores das publicações, o segundo campo contém palavras-chave atribuídas pela *Web of Science* às publicações com base na extração de palavras e termos contidos nos títulos e resumos dos documentos. Inicialmente, foi realizada a fusão dos campos de modo a criar um novo campo único³. Em seguida, foi realizado um procedimento para

³ Mais informações sobre a relevância da utilização das *Keywords Plus* em completo as palavras-chave do autor podem ser encontradas em Zhang et al. (2016).

padronização de termos no singular e plural por meio de algoritmos disponibilizados pelo software ou por processo manual⁴.

4.1.3

Processamento e cálculos aplicados às publicações científicas

O indicador de índice de impacto h-index para análise dos países foi calculado utilizando-se ferramentas da própria base de dados *Web of Science*. Tendo sido aplicada a estratégia de busca apresentada anteriormente para publicações relacionadas a química do estado sólido na indústria farmacêutica, foi realizado o refinamento a partir da ferramenta de Filtro para “Países/Territórios”, permitindo a análise específica das publicações científicas de cada um dos principais países identificados. Em seguida, foi utilizada a opção para criação de relatório de citações, a qual gerou o h-index para cada um dos países selecionados.

O índice de colaboração para os países foi obtido utilizando-se a equação abaixo e o resultado é percentual.

$$\text{Índice de Colaboração} = \frac{\text{Total de Publicações em Colaboração}}{\text{Total de Publicações}} \times 100$$

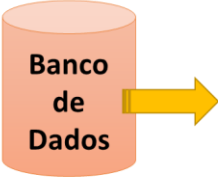
4.1.4

Processamento das publicações científicas para subsídios ao modelo de identificação de competências específicas

A partir da extração de informações contidas em cada publicação científica recuperada na Etapa 1, essas podem ser rearranjadas em uma Tabela Base conforme exemplo demonstrado na Figura 7 tendo como objetivo fornecer dados para elaboração do modelo de identificação de competências científicas descrito na Etapa 3. Essa Tabela Base contém grande parte das informações que são usadas para o ranqueamento dos pesquisadores.

⁴ Procedimento realizado utilizando-se a função List Cleanup do software VantagePoint com algoritmo de Lógica Fuzzy. Definição dos termos baseada naquele de maior ocorrência.

- Tabela Base:



Artigo	Revista	Ano	Autores	Instituição	# de "Palavras-Importantes"					# Citações do Artigo	Fator de Impacto da Revista
					PVA	PE	...	PP	MN T		
Characterization of ...	Journal of Food Engineering	2008	A; B; C	UCLA; UCB; UCLA	5	4	...	0	0	7	3,4
A novel composite ...	Journal of Condensed Material	2002	A	UCLA	4	2	...	1	0	22	0,8
...

Figura 7 – Tabela Base que contém as principais informações extraídas da publicação.

É possível verificar que as colunas são compostas por:

- Título do artigo;
- Título do periódico em que foi publicado;
- Nome dos autores;
- Instituição dos autores;
- Contagem das “palavras-importantes” contidas no resumo da publicação;
- Contagem das citações do artigo ao longo do tempo;
- Fator de impacto da revista.

Cada registro da tabela são as informações contidas no trabalho analisado. O campo “palavras importantes” é usado para realizar a composição do perfil de cada pesquisador. A definição desse conjunto de termos é mais bem descrita na subseção 4.2. O restante das informações é usado para o cálculo das métricas para o ordenamento dos pesquisadores.

4.2

Etapa 2: Modelo de Correlação entre palavras-chave e tópicos

Na segunda etapa, um modelo baseado em relações fuzzy é aplicado para definir o grau de pertinência ou afinidade das principais palavras-chave citadas nas 7.191 publicações científicas analisadas e grandes tópicos dentro da área de

química do estado sólido previamente definidos. Os resultados obtidos foram analisados usando também o método Fuzzy C-means. A definição dos tópicos foi realizada a partir de intensa busca na literatura (BYRN *et al.*, 2017; PRADO & ROCHA, 2015; ABRIMOV, 2016) além da consulta a grupo de especialistas, tendo sido selecionados 12 tópicos que representam áreas de conhecimento e propriedades específicas dentro da química do estado sólido. Os tópicos selecionados são apresentados abaixo:

Quadro 3 – Tópicos selecionados em química do estado sólido na indústria farmacêutica.

Polimorfismo
Sais
Cocristais
Sólidos amorfos
Propriedades Higroscópicas
Propriedades Mecânicas
Solubilidade e Dissolução
Propriedades Físicas
Propriedades termodinâmicas
Seleção de formas sólidas
Propriedades Químicas
Propriedades de Superfície

O método proposto para definir o grau de pertinência ou afinidade das principais palavras-chave citadas e os tópicos dentro da área de química do estado sólido indicados acima é baseado em Relações Fuzzy. Uma Relação Fuzzy (KLIR e YUAN, 1995; PEDRYCZ e GOMIDE, 1998) representa um grau de presença/ausência de associação, interação ou interconexão entre elementos de dois ou mais conjuntos. Uma Relação Fuzzy estende a noção clássica de relação entre elementos, propondo o uso de ferramentas da Lógica Fuzzy (ZADEH, 1965) para medir graus de associação entre elementos e efetuar o processo de inferência e composição de relações.

Não obstante, ressalta-se que a alternativa via métodos baseados na teoria das probabilidades não se aplica nesse caso. Basta ter em consideração que não se está interessado em medir a probabilidade de uma palavra-chave estar relacionada a um determinado tópico ou não, mas sim qual o seu grau de afinidade ou

associação a esse tópico. Esse tipo de incerteza, ocasionada pela dificuldade de categorizar uma palavra-chave genérica que pode estar relacionada a diversos tópicos, cada um com um grau de pertinência, pode ser mensurado a partir de ferramentas da Lógica Fuzzy.

4.2.1

Matriz de pertinência

Objetivamente, o método proposto nessa tese faz uso de duas matrizes:

- Matriz de Pertinência: Autores \times Palavras-chaves
- Matriz de Pertinência: Palavras-Chaves \times Tópicos

O método proposto usa duas matrizes de relevância – a matriz R (Autores \times Palavras-chaves) e a matriz A (Palavras-Chaves \times Tópicos) – cujas informações possibilitam que se obtenha a matriz de relevância S (Autores \times Tópicos). Sendo assim, considere-se a existência de n Palavras-Chaves ($i = 1, \dots, n$) e K Autores ($k = 1, \dots, K$). Seja r_{ki} o grau de afinidade entre a k -ésima Autores (professores e pesquisadores) e a i -ésima palavras-chaves (como *solubility*, *cocrystals*, *polymorphism* etc.). Cada r_{ki} compõe a matriz de adesão R , ou seja, a matriz de afinidade entre Autores e Palavras-chaves:

$$R = \begin{bmatrix} r_{11} & \cdots & r_{1n} \\ \vdots & \ddots & \vdots \\ r_{K1} & \cdots & r_{Kn} \end{bmatrix}_{K \times n}, \text{ de modo que } r_{ki} \in [0,1]$$

Seja a_{ij} o grau de associação entre a Palavras-chaves i e o Tópicos j ($j = 1, \dots, J$). A matriz de adesão A , ou seja, a matriz de associação entre Palavras-chaves e Tópicos é dada por:

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1J} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nJ} \end{bmatrix}_{n \times J}, \text{ de modo que } a_{ij} \in [0,1]$$

Da composição das matrizes R e A , a matriz de adesão S é obtida, de modo que cada posição representa o grau de afinidade entre a k -ésima Autores e o j -ésimo Tópicos:

$$S = \begin{bmatrix} s_{11} & \cdots & s_{1j} \\ \vdots & \ddots & \vdots \\ s_{k1} & \cdots & s_{kj} \\ \vdots & \ddots & \vdots \\ s_{K1} & \cdots & s_{Kj} \end{bmatrix}_{K \times J}, \text{ de modo que } s_{kj} \in [0,1]$$

Onde $S = R \circ A$, em que o operador \circ representa a composição máx-mín (Klir & Yuan, 1995), e, portanto, cada $s_{kj} = \max(\min(r_{k1}, a_{1j}), \dots, \min(r_{kn}, a_{nj}))$. Logo, o tomador de decisão pode avaliar o grau de afinidade entre as Autores e os mais diversos Tópicos e, portanto, obter informações estratificadas das competências de cada Autor atividades específicas dentro de química do estado sólido (sólidos amorfos, polimorfismo, solubilidade e dissolução etc.).

Contudo, é necessário estabelecer critérios que dizem respeito à estimativa dos elementos das matrizes de pertinência R e A , ou seja, os valores r_{ki} e a_{ij} . A próxima seção discute com mais detalhes a metodologia estabelecida para computar tais quantidades.

4.2.1.1

Estimativa dos elementos da matriz Autores x Palavras-Chaves

Esta seção apresenta métodos para fazer a estimativa dos itens da matriz de relevância entre *Autores* e cada Palavra-chave. O problema original é determinar o grau de associação entre as k -ésimas Autores e a Palavra-chave i (exemplos: *solubility*, *cocrystals*, *polymorphism* etc.), que será denotado por r_{ki} . Esse grau de associação considera valores entre 0 e 1, de modo que quanto mais próximo r_{ki} estiver de 1, maior será a associação entre as Autores k com a Palavra-chave i .

Para conseguir isso, consideremos $\#P_{k \Rightarrow i}$ o número de publicações do k -ésimo Autor na i -ésima palavra-chave, enquanto $\#P_k$ representa o número total de publicações do k -ésimo Autor. Para computar essas quantidades, as publicações de pesquisadores ligados a essa entidade precisam ser extraídas, considerando o período de publicação de 1997 a 2016, categorizadas por área do conhecimento, o

que é bastante comum na academia. A partir dessas definições, propõe-se o seguinte estimador para r_{ki} :

$$\hat{r}_{ki} = \frac{\#P_{k \Rightarrow i}}{\#_k} \in [0,1]$$

onde \hat{r}_{ki} é a frequência de publicação do k -ésimo Autor na Palavra-chave i . Quanto mais Autores publicar numa dada Palavra-Chave, maior será sua afinidade com ela. Após computar as quantidades presentes na matriz de adesão R , se torna necessário computar os elementos da matriz de adesão A para finalizar o processo de determinar a matriz de adesão S .

4.2.1.2

Estimativa dos elementos da matriz Palavras-Chaves x Tópicos

Esta seção apresenta um mecanismo para estimar os elementos da matriz de pertinência entre Palavras-Chaves e Tópicos. O problema original é determinar o grau de associação entre a Palavra-Chave i (exemplos: *solubility*, *cocrystals*, *polymorphism* etc.) e o Tópico j (exemplos: sólidos amorfos, polimorfismo, solubilidade e dissolução etc.) que será denotado por a_{ij} . Esse grau de associação considera valores entre 0 e 1, de modo que quanto mais próximo a_{ij} estiver de 1, maior será a associação entre a Palavra-Chave i e o Tópico j .

Há duas maneiras de se fazer a estimativa de a_{ij} . A primeira, uma abordagem objetiva, em que se verifica o número de Palavras-Chaves i que compartilha a mesma aplicação com (ou a propõe como campo de implementação do) Tópico j . Denotemos por $\#P_{i \Rightarrow j}$ o número de publicações associadas com a Palavra-Chave i com o Tópico j , enquanto que $\#P_i$ é o número de publicações examinadas na Palavra-Chave i . Logo, o estimador é:

$$\hat{a}_{ij} = \frac{\#P_{i \Rightarrow j}}{\#P_i} \in [0,1]$$

É possível ver que \hat{a}_{ij} é a frequência relativa de publicações que associam a palavra-chave i com o tópico j . Portanto, se o espectro completo de publicações for pesquisado, \hat{a}_{ij} converge para a probabilidade de se aplicar a publicação na

palavra-chave i no t3pico j . Note-se que, nesse caso, a frequ3ncia relativa est3 sendo considerada como uma associa33o entre a palavra-chave e o t3pico.

Uma segunda abordagem seria mais subjetiva, em que propomos que haja um comit3 de especialistas, cujos pontos-de-vista e experi3ncias combinados podem formar uma estimativa para a_{ij} . O especialista servir3, portanto, como ponte entre o conhecimento criado na palavra-chave i e seu potencial uso no t3pico j . Consideremos $x_{i \Rightarrow j}^{(d)}$ como a opini3o do d -3simo especialista ($d = 1, \dots, D$) sobre a associa33o entre a palavra-chave i e o t3pico j . Sugerimos o seguinte estimador para a_{ij} :

$$\ddot{a}_{ij} = \sum_{d=1}^D w^{(d)} x_{i \Rightarrow j}^{(d)} \in [0,1], \text{ com } w^{(d)} \geq 0 \text{ e } \sum_{d=1}^D w^{(d)} = 1$$

onde $w^{(d)}$ representa o peso ou import3ncia do especialista d no corpo de opini3es dos especialistas (definido atrav3s de um desempenho maior na 3rea, titula33o, publica33es, etc.). Se todos os especialistas tiverem a mesma relev3ncia, podemos escolher $w^{(d)} = \frac{1}{D}$ para todos os d , ou seja, temos que \ddot{a}_{ij} 3 a m3dia aritm3tica das opini3es de cada especialista.

Outro ponto de an3lise 3 como coletar a informa33o $x_{i \Rightarrow j}^{(d)}$ de cada especialista. 3 f3cil perceber o qu3o oneroso e amb3guo seria pedir a um especialista um valor entre 0 e 1 para $x_{i \Rightarrow j}^{(d)}$. Visto que esse procedimento n3o 3 vi3vel, foi utilizado o procedimento da Escala de Likert, muito usado em estudos psicom3tricos (interesse nos produtos, grau de afinidade entre ideias, etc.). Um item de Likert 3 simplesmente uma afirma33o 3 qual o entrevistado responde com o n3vel de concord3ncia ou discord3ncia com a afirma33o. Normalmente s3o usados cinco n3veis de resposta, embora alguns pesquisadores gostem de usar sete ou at3 nove n3veis. O formato t3pico de um item de Likert 3:

- Discordo totalmente – $x_{i \Rightarrow j}^{(d)} = 0$
- Discordo parcialmente – $x_{i \Rightarrow j}^{(d)} = 0,25$
- Sou indiferente – $x_{i \Rightarrow j}^{(d)} = 0,50$
- Concordo parcialmente – $x_{i \Rightarrow j}^{(d)} = 0,75$
- Concordo totalmente – $x_{i \Rightarrow j}^{(d)} = 1,00$

Em que a afirmação feita ao especialista é: "**Essa palavra está associada a esse tópico**". A partir dessa afirmação, o especialista indica seu nível de concordância. Se houver discordância total, atribuímos $x_{i \Rightarrow j}^{(d)}$ ao valor nulo, ou seja, grau de afinidade 0, do contrário, um grau positivo de afinidade é estabelecido para $x_{i \Rightarrow j}^{(d)}$. Após apresentar todas as afirmações (ao todo, há n para cada j -ésimo setor produtivo) para cada um dos especialistas, o estimador apresentado previamente é computado. Esse processo ocorre para todos os Tópicos, de modo que a matriz A é formada pelos elementos \ddot{a}_{ij} .

Por fim, destaca-se que a abordagem ideal seria combinar informações objetivas e subjetivas, visto que se adicionaria tanto um diagnóstico do passado e do presente (forma objetiva) e novas fronteiras (percepção subjetiva) na relação entre Autores e Tópicos em química do estado sólido. Contudo, a abordagem objetiva é muito complexa para se conduzir normalmente, dado o número de publicações nas diversas palavras-chaves e tópicos que essa pesquisa engloba. Assim, o caminho subjetivo de estimativas será usado para determinar os valores que compõem a matriz da relação Palavras-Chaves x Tópicos. A próxima seção a apresentação do algoritmo usado.

4.2.1.3

Algoritmo do método Fuzzy C-Means

O algoritmo do método Fuzzy C-Means (Bezdek e Pal, 1992) funciona de forma bastante similar com o método tradicional K-Means, mas, em vez de fornecer partições rígidas, ele gera partições maleáveis – ou seja, conjuntos que se sobrepõem. Logo, em vez de um elemento pertencer a um único grupo, ele pertence a vários grupos, mas com graus diferentes de adesão. Mais especificamente, consideremos uma base de dados $X = \{x_1, \dots, x_N\}$, com N objetos, o resultado final é uma matriz de partição difusa (*fuzzy*) desses objetos num dado número k de grupos. Essa matriz de partição é representada pela matriz U ($k \times N$), cujos elementos U_{ij} representam o grau de adesão do j -ésimo objeto para com o i -ésimo grupo difuso.

Em geral, o algoritmo C-Means busca minimizar alguma variante ou extensão da seguinte função objetiva:

$$\text{Min} : J(v_j) = \sum_{i=1}^N \sum_{j=1}^k (U_{ij})^m D_{ij}^2$$

s.a.

$$\sum_{j=1}^k U_{ij} = 1$$

$$1 \leq i \leq N$$

O expoente m é um parâmetro de “fuzzificação” (normalmente igual a 2), D_{ij} é uma determinada distância entre o j -ésimo objeto e o i -ésimo grupo e J é uma medida da dissimilaridade intragrupo. A escolha do número de grupos é subjetiva e específica para o problema, mas é possível considerar métricas de qualidade do agrupamento (Soma dos Desvios Quadráticos, Centralidade e Silhueta).

4.3

Etapas 3: Modelo de identificação de competências específicas – NB-RQ

Para identificar competências e lideranças de pesquisadores ou instituições, foi utilizada nesta tese o método NanoBusiness Researcher Quality Index (NB-RQ) para identificação de competências específicas. Esse modelo foi desenvolvido pela equipe NanoBusiness para atender demandas na identificação de talentos, empresas e instituições de organizações com e sem fins lucrativos como, por exemplo, a Braskem e a UNESCO. O NB-RQ é composto por 5 escores que avaliam o pesquisador ou instituição sob cinco pontos de vista diferentes (Perspectiva 5 V's):

- **Volume** de publicações - V_{pub}
- **Validade** das publicações - V_{val}
- **Valor acadêmico** - V_{acad}
- **Valor industrial** - V_{ind}
- **Valor ao projeto** - V_{proj}

Quanto maior for o escore do pesquisador ou da instituição em cada quesito, maior é a sua relevância dentro da comunidade, assim como o uso da sua pesquisa pela indústria. Portanto, o NB-RQ é calculado da seguinte forma:

$$NB - RQ = w_1 V_{pub}^* + w_2 V_{val}^* + w_3 V_{acad}^* + w_4 V_{ind}^* + w_5 V_{proj}^*$$

tal que: $\sum_{i=1}^5 w_i = 1$ e cada $w_i \geq 0$. Cada w_i representa o peso ou o grau de importância de cada escore para compor o NB-RQ. Caso todos os w_i 's sejam iguais, o NB-RQ é a média aritmética dos escores. A escolha de cada w_i deve ser realizada levando-se em conta os objetivos desejados, de maneira que se possa definir quais critérios (volume de publicações, valor acadêmico, etc.) são os mais representativos para o ordenamento entre os pesquisadores ou instituições. Cabe ressaltar que para agregar os escores é necessário normaliza-los (por isso a notação V se tornou V^*), pois cada um desses possui escalas distintas. A seguir são apresentados os estimadores para cada escore.

- **Volume de publicações:** o volume de publicações do pesquisador ou instituição é computado pela média de publicações durante o período em estudo, dado por:

$$V_{pub} = \frac{\sum_{t=1}^T \#P_t * a_t}{T}$$

onde $\#P_t$ é o total de publicações do pesquisador ou instituição no ano t ($t=1, \dots, T$), $a_t \in [0,1]$ é um fator que leva em conta a “idade” da publicação. Para períodos curtos de estudo, e dependendo da área de pesquisa sob análise, a_t pode ser escolhido constante, não impactando, portanto, no resultado final.

- **Validade das publicações:** a validade das publicações é computada a partir da média do fator de impacto dos periódicos em que o pesquisador ou instituição publicou seus artigos. O fator de impacto de um periódico, calculado pelo Journal Citation Reports (JCR) publicado pela Clarivate Analytics, é definido pelo o número médio de citações dos artigos que

foram publicados durante o biênio anterior. A partir dessa informação, computa-se a seguinte quantidade para cada pesquisador ou instituição:

$$V_{val} = \frac{\sum_{t=1}^T FI_t * a_t}{T}$$

onde FI_t é o média do fator de impacto do pesquisador ou instituição no ano t.

- **Valor acadêmico:** o valor acadêmico do pesquisador ou instituição está intimamente ligado em o quanto a sua produção é citada pelos seus demais pares da comunidade. Portanto, é possível estimar o valor acadêmico do pesquisador por:

$$V_{acad} = \frac{\sum_{t=1}^T C_t * a_t}{T}$$

onde C_t é a média das citações recebidas pelo pesquisador no ano t.

- **Valor industrial do pesquisador:** o valor industrial do pesquisador ou instituição é baseado no número de patentes depositadas ou em fase de depósito, que este produziu ao longo do tempo. Ainda, as patentes consideradas precisam ter algum nível de ligação com o projeto. Assim, calcula-se o valor industrial do pesquisador como sendo:

$$V_{ind} = \frac{\sum_{t=1}^T Pat_t * a_t}{T}$$

onde Pat_t é o total de patentes depositadas no ano t. Cabe ressaltar que essa informação é a única não contida nos registros de publicações científicas identificadas na etapa de 3.1 de Elaboração de indicadores de C&T em química do estado sólido. Sendo assim, foi empregada uma estratégia semelhante a apresentada para identificação das publicações científicas, tendo sido adaptada para busca por patentes. A base de dados utilizada para essa atividade foi a *Derwent Innovations Index*.

- **Valor ao projeto:** o valor do pesquisador ou instituição é estimado a partir do grau de afinidade da sua produção científica aos tópicos específicos de interesse definidos dentro do tema de estudo. Para este critério, utiliza-se o modelo de correlação entre palavras-chave e tópicos apresentado anteriormente. Este critério é então aferido com base na avaliação de um conjunto de especialistas quanto ao grau de pertinência ou afinidade das principais palavras-chave das publicações científicas analisadas e os tópicos em química do estado sólido apresentados no item 3.2 e que incluem polimorfismo, sais, cocristais, dentre outros. Sendo assim, as publicações de cada pesquisador ou instituição podem ser correlacionadas aos tópicos de interesse, permitindo uma análise específica de liderança e competência em cada tópico. Então, o valor do pesquisador ou instituição é dado por:

$$V_{proj} = S = \begin{bmatrix} s_{11} & \cdots & s_{1J} \\ \vdots & \ddots & \vdots \\ s_{K1} & \cdots & s_{KJ} \end{bmatrix}_{K \times J}, \text{ de modo que } s_{kj} \in [0,1]$$

onde s_{kj} é o grau de afinidade das publicações do pesquisador ou instituição no ano t . O ponto de análise a se enunciar é como coletar a informação $a_i^{(d)}$ de cada especialista. É fácil ver como é custoso e ambíguo pedir ao especialista um valor entre 0 e 1 para $a_i^{(d)}$. Como tal procedimento é inviável, utilizou-se o procedimento da Escala de Likert, comumente usada em estudos psicométricos (interesse em produtos, grau de afinidade entre ideias, etc.). Um item Likert é uma afirmação à qual o sujeito pesquisado responde com o nível de concordância ou não concordância à afirmação. Usualmente são usados cinco níveis de respostas, cujo formato de um item Likert é:

- Não concordo totalmente – $a_i^{(d)} = 0$
- Não concordo parcialmente – $a_i^{(d)} = 0,25$
- Não concordo nem discordo – $a_i^{(d)} = 0,50$
- Concordo parcialmente – $a_i^{(d)} = 0,75$
- Concordo totalmente – $a_i^{(d)} = 1,00$

onde a afirmação feita ao especialista é: “**A Publicação i é útil/relevante/aplicável ao tópico.**”. A partir dessa afirmação, o especialista atribui o seu nível de concordância. Caso esse não concorde totalmente, atribuímos a $a_i^{(d)}$ o valor nulo, isto é, grau de afinidade 0, caso contrário, um grau de afinidade positivo é estabelecido a $a_i^{(d)}$.

Com o cálculo de cada uma dessas métricas é possível o ranqueamento dos pesquisadores e instituições segundo o seu nível de publicações, desempenho acadêmico e afinidade com o projeto.

5

Resultados e discussão

5.1

Indicadores científicos e tecnológicos em química do estado sólido no mundo

5.1.1

Indicadores científicos em química do estado sólido

Esta subseção apresenta a análise dos indicadores científicos elaborados a partir das 7.191 publicações científicas em química do estado sólido recuperadas a partir da estratégia de busca empregada nesta pesquisa para o período de 1997 a 2016. Como pode ser observado na Figura 8, o número de publicações científicas em química do estado sólido no mundo cresceu substancialmente a partir de meados dos anos 2000. Entre 1997 e 2016, o número de publicações científicas passou de 82 para 896 com um crescimento médio anual de 13,41% no período.

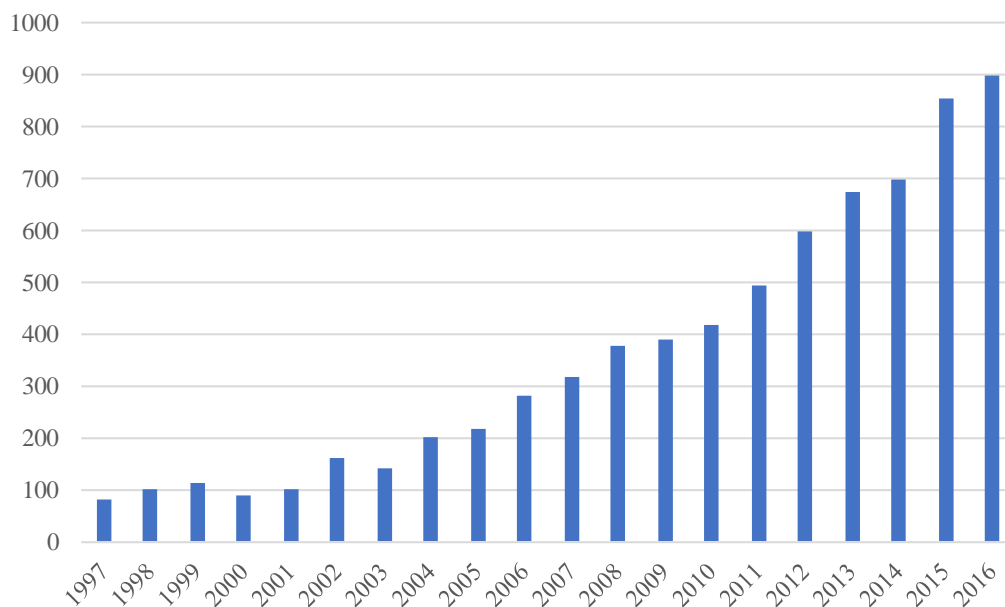


Figura 8 – Evolução das publicações científicas em química do estado sólido (1997-2016).

Fonte: Web of Science. Elaborado pelo autor.

Este crescimento, sobretudo a partir de 2004, reflete o ganho de importância da área para a indústria farmacêutica seguindo a esteira de diversos acontecimentos relevantes nos anos anteriores como o emblemático caso do IFA ritonavir em 1998 e o lançamento do guia do *International Committee on Harmonization* (ICH) Q6A em 1999.

Entre os principais países com maior número de publicações científicas em química do estado sólido destacam-se EUA, Reino Unido, Índia, China e Alemanha que, em conjunto, representaram 57,8% das publicações científicas do período analisado. Além de apresentar a maior contribuição no número total de publicações científicas, os EUA também tiveram o maior índice de impacto (h-index) dos países listados, seguido do Reino Unido e da Alemanha.

Quadro 4 – Indicadores de perfil dos principais países em pesquisa em química do estado sólido.

País	Total de publicações	h-index	Índice de colaboração (%)
EUA	1617	115	29,07
Reino Unido	883	73	50,40
Índia	816	52	16,42
China	736	45	21,60
Alemanha	510	61	50,39
Itália	408	46	35,29
França	391	46	53,20
Japão	375	41	19,20
Brasil	287	25	28,22
Polônia	235	20	32,77

Fonte: Web of Science. Elaborado pelo autor.

A Índia também apresentou um índice de impacto (h-index) de destaque frente aos demais países, evidenciando sua importância para a área de química do estado sólido. O país asiático é hoje o maior fornecedor de medicamentos genéricos do mundo, sendo responsável por aproximadamente 20% das exportações globais em volume. Inicialmente, a Índia teve como pilar de atuação o mercado farmacêutico de medicamentos genéricos de baixo valor agregado. No entanto, o posicionamento indiano vem mudando ao longo dos últimos anos. Com o avanço de sua legislação de proteção de propriedade intelectual e baixos custos de produção, diversas farmacêuticas multinacionais passaram a terceirizar atividades de produção no país. Assim como a Índia, a China também passou a

ocupar uma posição relevante no mercado farmacêutico mundial, sobretudo por meio do fornecimento de insumos farmacêuticos ativos de menor valor agregado.

Ainda que Índia e China estejam entre os cinco principais países em número total de publicações, os índices de colaboração internacional desses dois países emergentes mostram-se relativamente baixos, destoando do restante dos líderes em publicações. Brasil e Polônia possuem índices de colaboração internacional pouco maiores (28,22% e 32,77%, respectivamente) do que Índia e China, mas os índices de impacto (h-index) de suas publicações são muito inferiores ao restante dos países listados, o que sugere que as pesquisas nesses países ainda estão em estágio inicial.

No que se refere a colaboração internacional, destacam-se França (53,20%), Reino Unido (50,40%) e Alemanha (50,39%), os quais apresentaram grande colaboração entre si conforme apresentado na Figura 9. Por outro lado, o Japão apresentou Índice de Colaboração relativamente baixo (19,20%). Os EUA foram o principal parceiro do Reino Unido, Índia, China, Alemanha, Japão e Brasil no período, conforme Figura 9.

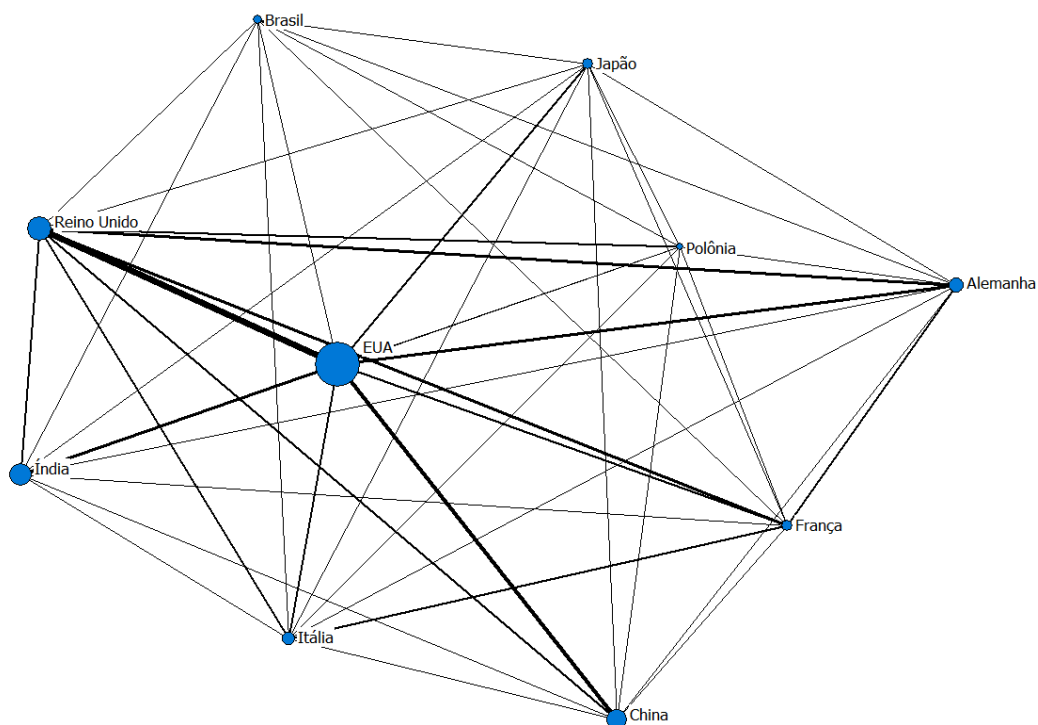


Figura 9 – Rede de colaboração dos principais países com publicações em química do estado sólido.

Fonte: Web of Science. Elaborado pelo autor.

Para o Brasil, os países que mais colaboraram com suas pesquisas além dos EUA foram Reino Unido, França, Itália, Alemanha e, por fim, Índia. O índice de colaboração relativamente baixo do Brasil (28,22%), bem como o seu baixo nível de impacto (4,21), evidenciam a necessidade de se criar estratégias de longo prazo que viabilizem a cooperação de cientistas brasileiros com pesquisadores estrangeiros, bem como condições que favoreçam o aumento na formação e titulação do número de doutores brasileiros com foco na qualidade de pesquisa, com capacidade de realizar interações com redes internacionais de colaborações.

Ao considerarmos a evolução do número de publicações ao longo do período analisado, observa-se que os esforços de Índia e China são mais recentes tendo ganhado força a partir de 2007. Além disso, suas publicações por ano já superaram as do Reino Unido e devem ultrapassar as dos EUA em breve.

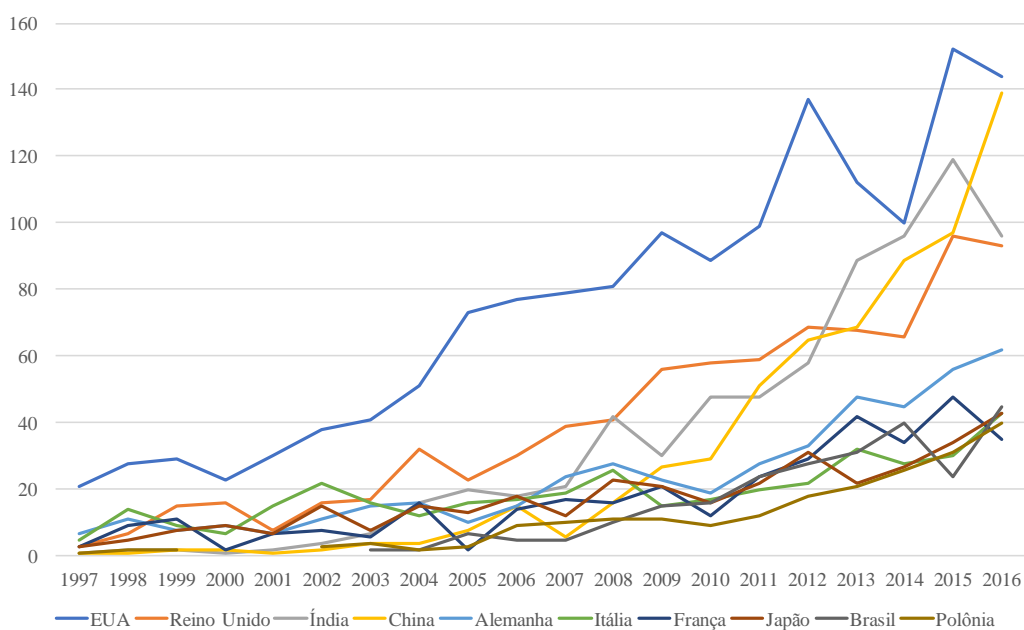


Figura 10 – Evolução temporal das publicações em química do estado sólido por país. Fonte: Web of Science. Elaborado pelo autor.

5.1.1.1

Principais áreas de pesquisa em química do estado sólido

Esta subseção tem como objetivo analisar as principais áreas de pesquisa em química do estado sólido. Inicialmente, são avaliadas as categorias da base de dados *Web of Science*, as quais são baseadas na classificação das revistas indexadas nas quais as publicações foram realizadas. Em seguida, são analisados

os termos mais frequentes nas publicações a partir da análise de conteúdo textual (*text mining*).

Como pode ser observado na Figura 11, as categorias de “Pharmacology & Pharmacy” e “Chemistry, Multidisciplinary” foram as duas principais ao longo do período. Um aspecto interessante é o forte crescimento das categorias “Crystallography” e “Materials Science, Multidisciplinary” a partir de 2007, refletindo uma maior importância do estudo de formas cristalinas em relação às propriedades do estado sólido nos últimos anos. Do mesmo modo, destaca-se a categoria “Chemistry, Physical”, a qual também compõem o grupo das principais categorias para o período juntamente com aquelas supracitadas.

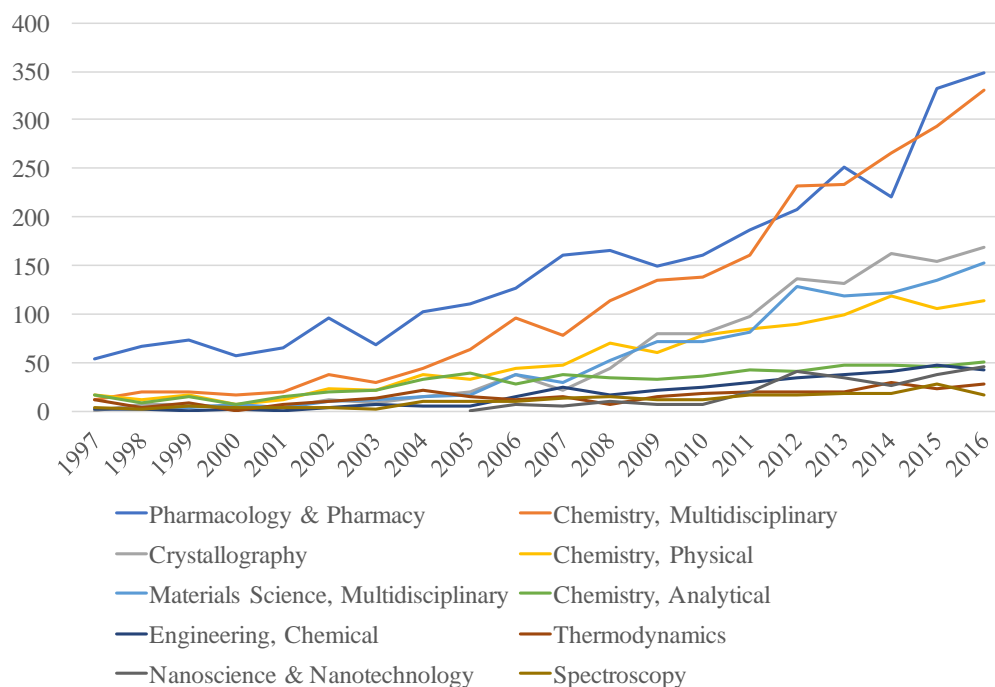


Figura 11 – Evolução temporal das categorias *Web of Science*.

Fonte: Web of Science. Elaborado pelo autor.

Para uma análise mais detalhada dos principais tópicos de pesquisa na área de química do estado sólido no período de 1997 a 2016, foi realizada a análise do conteúdo de dois campos extraídos da base de dados *Web of Science*: “*Keyword (Author’s)*” e “*Keyword Plus*”, conforme procedimento indicado na metodologia. Como pode ser observado na Figura abaixo, as palavras-chave mais frequentes nas publicações foram Solubility, Cocrystals e Polymorphism, seguidas de Drugs e Solid State.

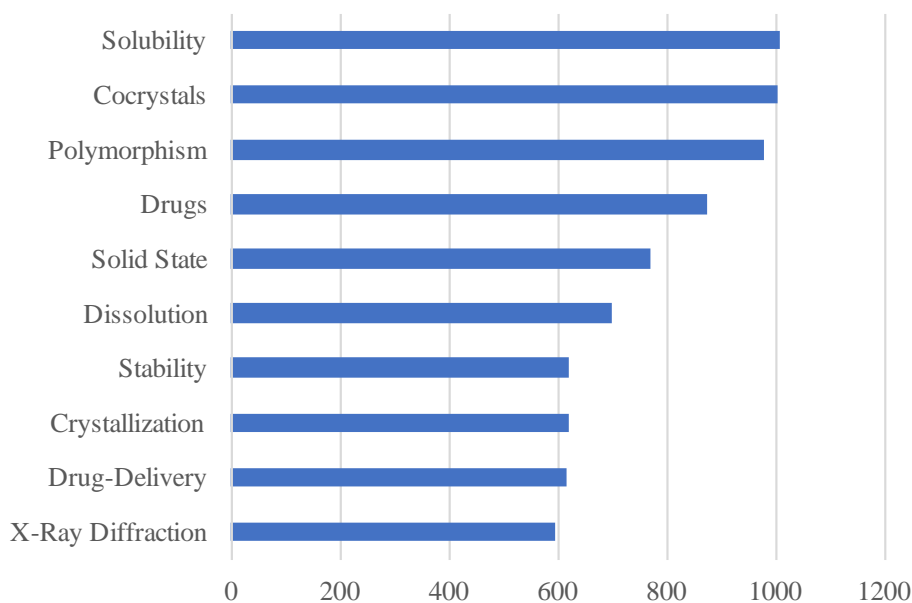


Figura 12 – Palavras-chave mais frequentes nas publicações em química do estado sólido.

Fonte: Web of Science. Elaborado pelo autor.

Ao analisarmos a frequência das palavras-chave ao longo do período (Figura 12), observa-se que o termo “Cocrystals” surge pela primeira vez somente em 2003 apresentando um forte crescimento nos anos seguintes até ocupar a segunda posição entre as palavras-chave mais frequentes. Segundo definição do FDA (2016), os cocristais são materiais cristalinos compostos de duas ou mais moléculas diferentes, tipicamente fármacos e formadores de cocrystal (“coformadores”), na mesma estrutura cristalina. Desse modo, o conceito retrata IFAs cristalizados com outras substâncias sólidas e neutras em condições ambientais, inclusive cristalizados com outros IFAs. O crescente interesse pelo tema nos últimos anos está relacionado com a possibilidade da utilização de cocristais para melhoria de características físico-químicas essenciais de um medicamento incluindo solubilidade, dissolução, biodisponibilidade e estabilidade de compostos farmacêuticos mantendo sua atividade terapêutica (BYRN et al, 2017).

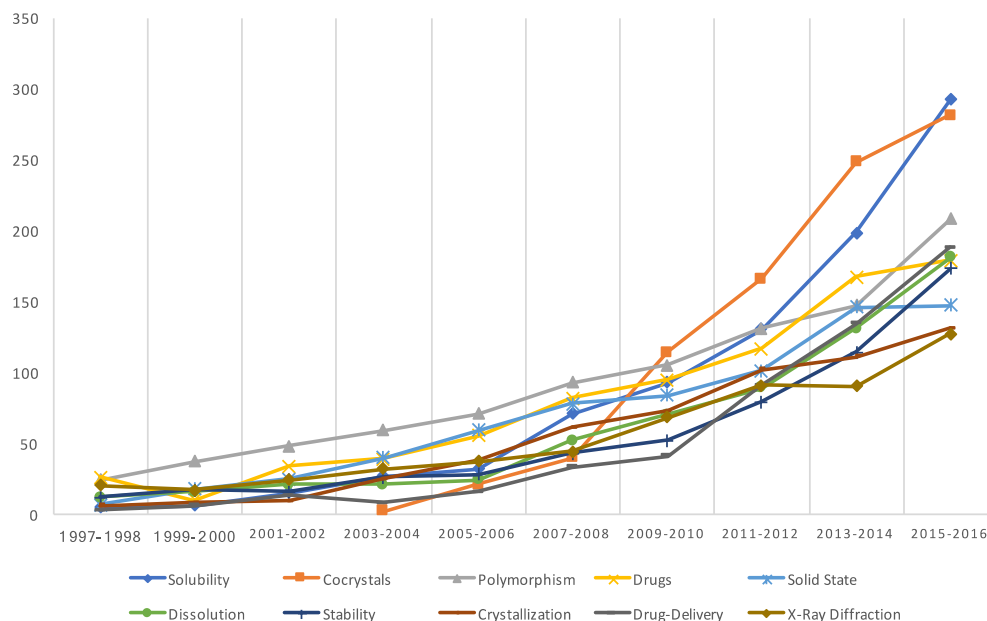


Figura 13 – Evolução temporal das palavras-chave mais frequentes nas publicações em química do estado sólido (1997-2016).

Fonte: Web of Science. Elaborado pelo autor.

Embora a maior parte das palavras-chave mais frequentes esteja relacionada a conceitos (ex.: “Cocrystals”) e propriedades do estado sólido (ex.: “Solubility”), destaca-se a presença de somente um termo relacionado a técnicas analíticas: “X-Ray Diffraction”. Diversas técnicas analíticas são empregadas para que se possa caracterizar e conhecer em detalhes as propriedades do estado sólido de um IFA. Em 2007, o Food and Drug Administration (FDA) lançou um guia contendo os principais aspectos relacionados ao polimorfismo sobre a qualidade dos medicamentos. Este guia destaca a responsabilidade do fabricante em controlar a forma cristalina caso a biodisponibilidade do medicamento possa ser afetada e recomenda que procedimentos analíticos apropriados sejam utilizados para detectar as formas cristalinas. Entre as técnicas preconizadas para o controle de polimorfismo em fármacos e medicamentos, a difração de raios-X por policristais se destaca por ser considerada uma prova inequívoca de existência de polimorfismo, sendo reconhecidamente, uma das referências mundiais para a identificação de polimorfismo em fármacos sólidos.

Embora a difração de raios-X por policristais seja a mais indicada para determinação de formas sólidas em materiais farmacêuticos sólidos, outras técnicas correlatas como análise térmica podem ser fundamentais para solução de problemas específicos. Ao analisarmos as técnicas analíticas mais citadas nas

publicações científicas em química do estado sólido, destacam-se, além da Difração de Raios-X (DRX), a Calorimetria Diferencial de Varredura (DSC), Infravermelho (FT-IR) e Ressonância Magnética Nuclear (RMN). Como pode ser observado na Figura 14, a técnica de DRX ganha maior relevância a partir de 2007, acompanhando as diretrizes de órgãos reguladores como o FDA quanto a centralidade dessa técnica para estudo do estado sólido de fármacos e medicamentos.

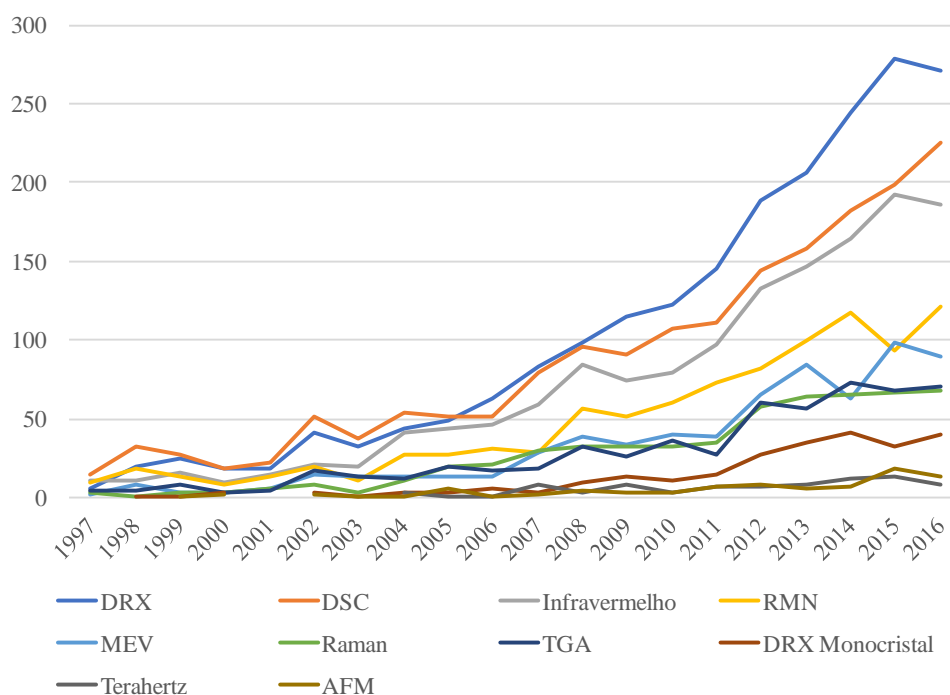


Figura 14 – Evolução temporal das principais técnicas analíticas nas publicações em química do estado sólido (1997-2016).

Fonte: Web of Science. Elaborado pelo autor.

Outro aspecto interessante da análise das palavras-chave mais frequentes nas publicações científicas é a avaliação de sua co-ocorrência, ou seja, a relação existente entre as mesmas dentro do conjunto das publicações. Desse modo, pode-se obter uma representação visual das áreas mais intensivas nas pesquisas em química do estado sólido, fornecendo indicativos do perfil da geração de conhecimento. A Figura 15 apresenta a visualização da densidade da rede, onde a força dos nós mais conectados é representada por uma medida de cor. Assim, o mapa de densidade apresenta a rede de palavras-chave e seus principais clusters,

levando-se em conta sua co-ocorrência, bem como a intensidade de cada palavra-chave no conjunto das publicações⁵.

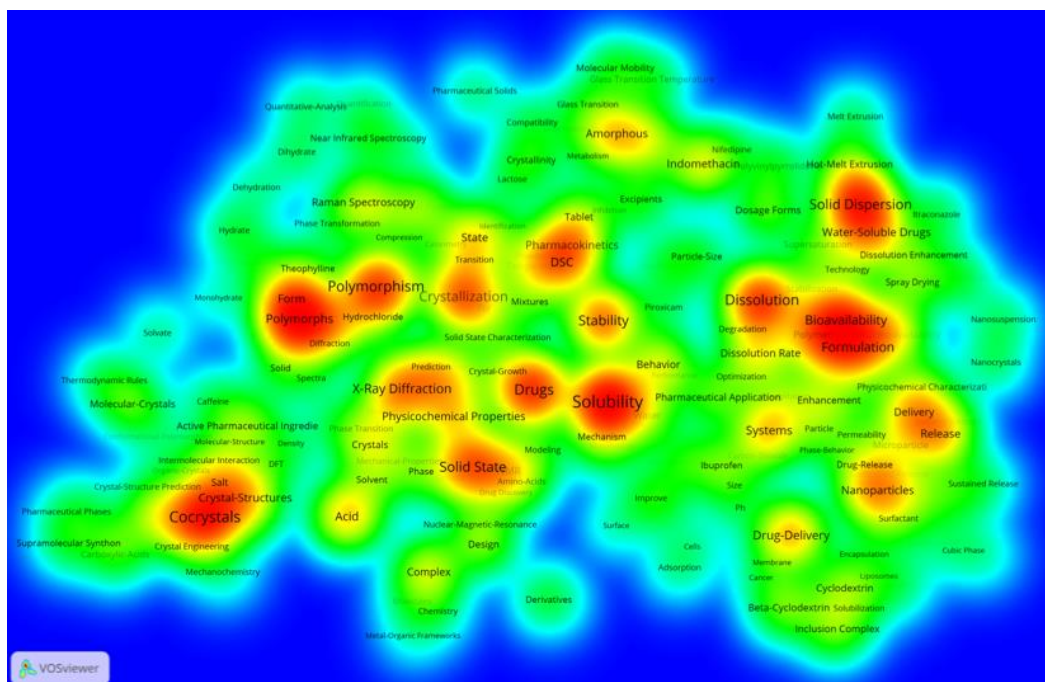


Figura 15 – Mapa de densidade da rede de palavras-chave mais freqüentes.
Fonte: Web of Science. Elaborado pelo autor.

Como pode ser observado, a distribuição das palavras-chave no mapa de densidade apresenta uma visão ampla das pesquisas na área de química do estado sólido. Ao analisarmos mais detalhadamente o cluster formado em torno da área da palavra “Cocrystals” pode-se observar sua próxima relação com os conceitos de “Crystal-Structure”, “Salt”, “Crystal Engineering” e “Crystal-Structure Prediction”.

Além disso, a partir da análise dos dados de co-ocorrência entre as palavras-chave, pode-se extrair ainda uma visualização consistente de conceitos específicos. No que se refere ao conceito de “Cocrystals”, observa-se uma forte relação de co-ocorrência com “Acid”, “Hydrogen Bonding”, “Crystal-Structures”, “Drugs” e “Physicochemical Properties”. Estas relações mais detalhadas podem ser observadas na figura a seguir.

⁵ Mais informações sobre a elaboração de mapas de densidade utilizando o software VOSviewer podem ser encontradas em Van Eck & Waltman (2011).

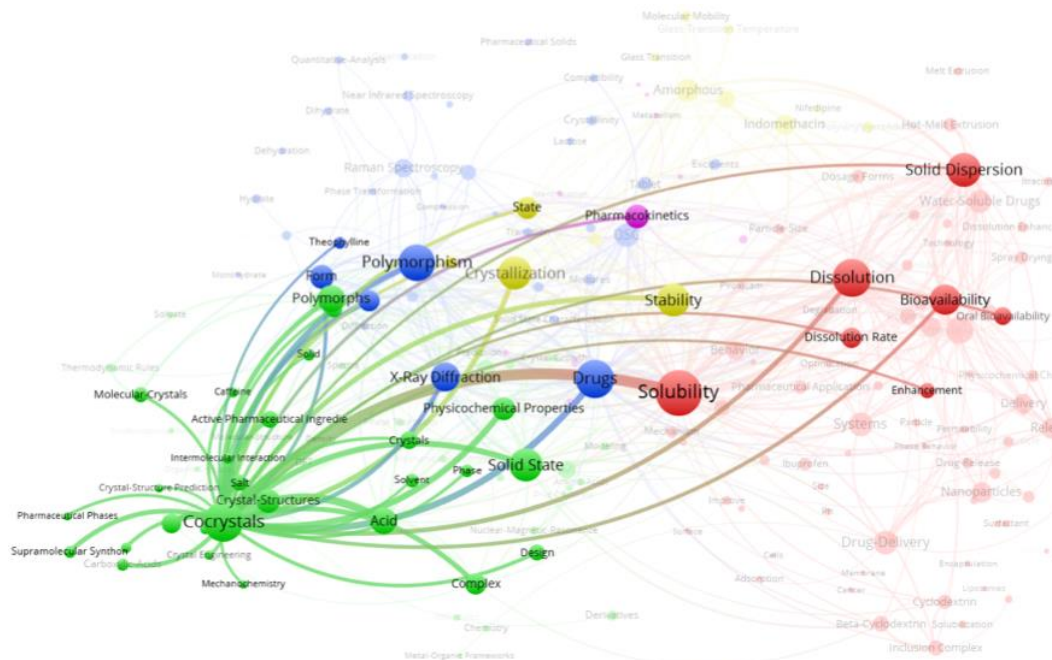


Figura 16 – Mapa de rede de co-ocorrência de palavras-chave com o termo cocrystals. Fonte: Web of Science. Elaborado pelo autor.

5.1.1.2

Principais organizações em publicações científicas na área de química do estado sólido na indústria farmacêutica

Nesta subseção são apresentados os indicadores referentes as quinze principais organizações em número de publicações científicas na área de química do estado sólido na indústria farmacêutica no período de 1997 a 2016. Como pode ser observado na Figura 17, a University of Copenhagen aparece no topo da lista com 130 publicações, seguida da Chinese Academy of Sciences (121), da Purdue University (118) e da University of Cambridge (102). Deve-se destacar ainda a presença da Universidade de São Paulo com 76 publicações, sendo esta a única representante brasileira no ranking.

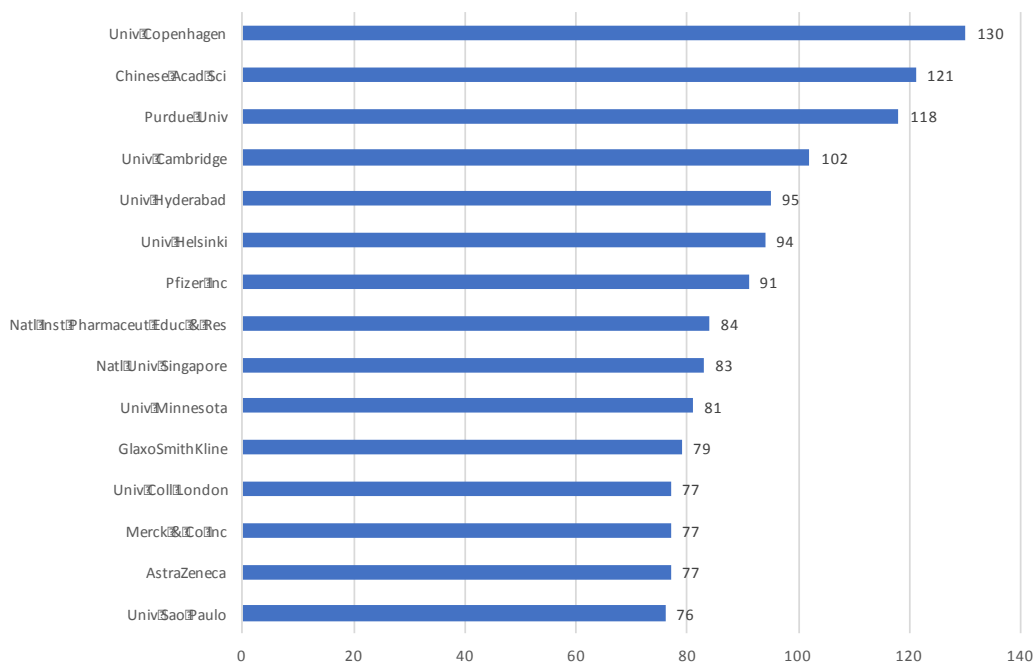


Figura 17 – Principais organizações em publicações científicas na área de química do estado sólido.

Fonte: Web of Science. Elaborado pelo autor.

O desempenho da University of Copenhagen pode ser atribuído a atuação de sua Faculdade de Saúde e Ciências Médicas, a qual inclui em sua estrutura o Departamento de Farmácia e o Departamento de Drug Design e Farmacologia, sendo estes os principais responsáveis pelas publicações da universidade. Entre as principais áreas de atuação da universidade destacam-se os sólidos amorfos e as dispersões sólidas, tendo como objetivo a melhoria da estabilidade de formulações e da solubilidade de fármacos pouco solúveis. Além disso, a universidade possui projetos com grandes empresas farmacêuticas, incluindo a dinamarquesa Novo Nordisk e a norte-americana Bristol-Myers Squibb, sendo esta última responsável por uma parceria para incentivar o desenvolvimento de competências em empreendedorismo e inovação aos alunos da Faculdade de Saúde e Ciências Médicas.

Embora os EUA tenham sido responsáveis pelo maior número de publicações no período de estudo como demonstrado anteriormente, somente duas universidades americanas aparecem entre as quinze principais organizações listadas, sendo estas a Purdue University e a University of Minnesota. Por outro lado, embora a Dinamarca não tenha sido identificada dentre os principais países em publicações, a University of Copenhagen lidera o ranking das organizações.

Dentre as organizações listadas, destaca-se a presença de quatro empresas farmacêuticas multinacionais: Pfizer Inc (91), GlaxoSmithKline (79), Merck & Co Inc (77) e AstraZeneca (77). A identificação desses atores dentre as principais organizações com publicações científicas demonstra que as empresas farmacêuticas desempenham um papel importante no avanço da geração de conhecimento na área da química do estado sólido, não estando esse movimento restrito ao âmbito acadêmico. Tal fato relaciona-se com a relevância da área para as atividades de P&D dessas empresas e seu posicionamento estratégico competitivo no mercado.

De fato, ao avaliarmos todo o conjunto das 7191 publicações científicas identificadas no período do estudo, diversas empresas farmacêuticas aparecem entre as organizações, como pode ser observado na Figura 18 que apresenta o ranking somente para este tipo de organização. A multinacional americana Pfizer Inc ocupa a primeira colocação com 91 publicações, seguida da multinacional inglesa GlaxoSmithKline com 79. Entre as principais empresas destacam-se ainda Merck & Co Inc (77), AstraZeneca (77) e Bristol Myers Squibb Co (67).

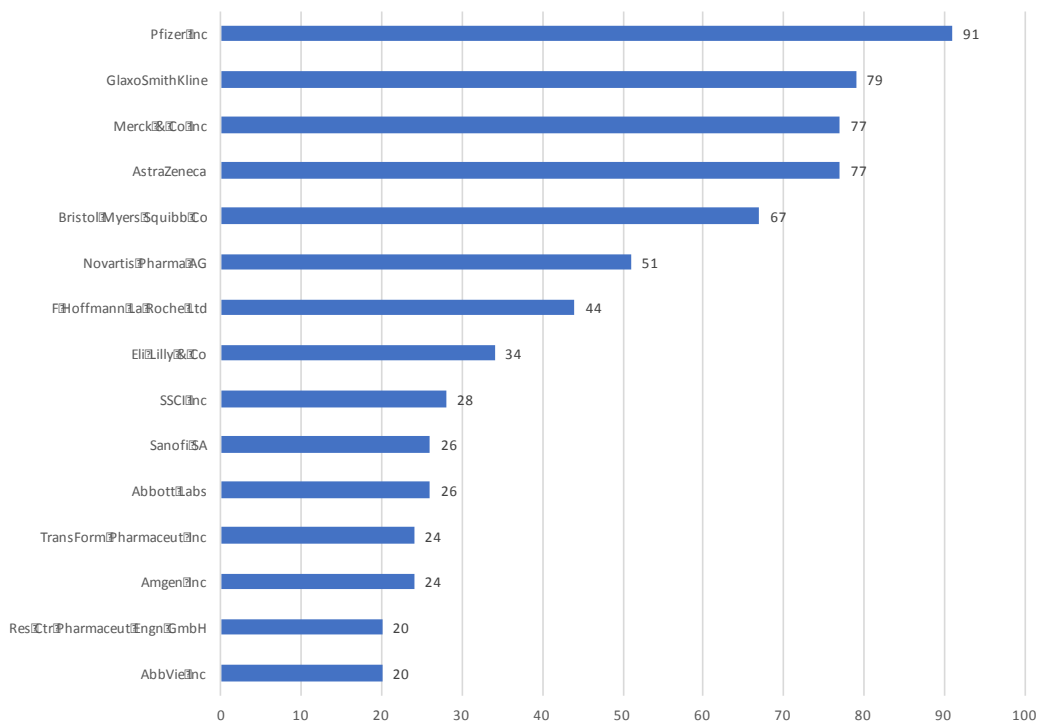


Figura 18 – Principais empresas em publicações científicas na área de química do estado sólido.

Fonte: Web of Science. Elaborado pelo autor.

A intensidade na geração de conhecimentos em química do estado sólido evidencia que a área tem ocupado uma posição cada vez mais relevante nas atividades de P&D farmacêuticas. Neste aspecto, deve-se ressaltar ainda que para as empresas farmacêuticas as publicações científicas são quase sempre acompanhadas da intensa utilização de mecanismos de proteção de propriedade intelectual, sobretudo patentes (WIPO, 2017).

Ao analisarmos mais detalhadamente o perfil das principais empresas em publicações científicas listadas na Figura 18, destaca-se a presença da norte-americana SSCI Inc. com 28 publicações. Enquanto todas as outras empresas listadas são farmacêuticas e atuam na produção e comercialização de medicamentos, a SSCI – acrônimo para Solid State Chemical Information – diferencia-se por atuar na prestação de serviços analíticos e de pesquisa para a indústria farmacêutica.

Fundada em 1991, a SSCI possui como foco de atuação acelerar o ritmo de produção de novos produtos e sua regulamentação, melhorando também a qualidade dos medicamentos oferecidos ao mercado. A empresa está localizada no Parque Tecnológico de Purdue, no estado de Indiana, próxima da Purdue University com a qual possui intensa parceria. Em fevereiro de 2015, a SSCI foi adquirida pela Albany Molecular Research Inc. (AMRI) por 36 milhões de dólares. A AMRI é uma das principais fornecedoras de serviços de pesquisa contratuais (CRO) e de terceirização de fabricação para indústrias farmacêutica e de biotecnologia. Atualmente, a SSCI é reconhecida como o centro de excelência da AMRI para química de estado sólido, auxiliando diversas empresas farmacêuticas em suas demandas relacionadas à química do estado sólido.

5.1.2

Indicadores tecnológicos em química do estado sólido

5.1.2.1

Evolução do patenteamento em química do estado sólido

Esta subseção apresenta a análise dos indicadores tecnológicos elaborados a partir dos 4.665 documentos de patente em química do estado sólido recuperados a partir da estratégia de busca empregada nesta pesquisa para o período de 1997 a

2015. Nesta análise são considerados apenas o número anual de patentes até 2015, pois os dados referentes a 2016 estão incompletos como consequência do período de sigilo no processo de patenteamento, normalmente de 18 meses, e do tempo de indexação do documento pela base.

Como pode ser observado na Figura 19, o patenteamento em química do estado sólido apresentou um forte crescimento entre 1997 e 2007, seguido de um período de queda entre 2008 e 2011. A partir de 2012, os depósitos de patente voltaram a crescer indicando uma tendência de reversão em relação ao período anterior.

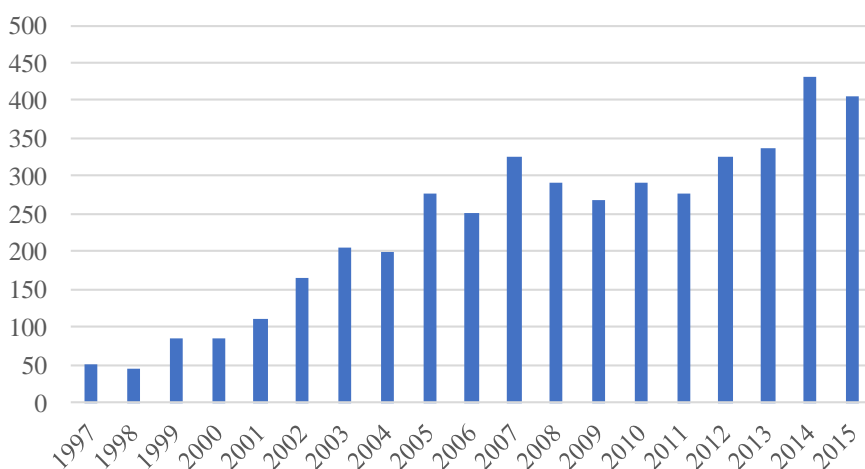


Figura 19 – Patenteamento em química do estado sólido por ano de prioridade (1997-2015).

Fonte: Derwent Innovations Index. Elaborado pelo autor.

Do total de patentes recuperadas mundialmente, aproximadamente 31% tiveram os Estados Unidos como país de origem do depósito, seguido de China (18%) e Índia (12%). Embora o país de origem do primeiro depósito seja uma importante referência do local onde a invenção foi gerada, deve-se levar em conta que a indústria farmacêutica mundial é altamente internacionalizada, sendo os Estados Unidos um dos principais mercados farmacêuticos mundiais.

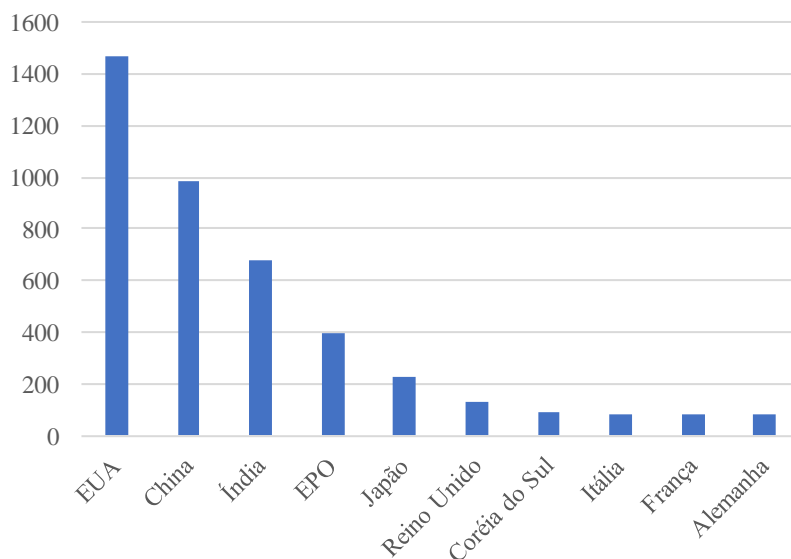


Figura 20 – Principais países de prioridade dos documentos de patente.
Fonte: Derwent Innovations Index. Elaborado pelo autor.

Dentre os países listados, a China destaca-se pelo forte crescimento no número de depósitos de patentes no período. Entre 1997 e 2015, os depósitos de patente chineses apresentaram um crescimento anual médio de 33,73%, número muito superior àquele apresentado pelo restante dos países. Além disso, deve-se ressaltar que a China apresentou um número de depósitos de patentes (1028) superior ao número de publicações científicas no período do estudo (736), o que indicaria forte capacidade para converter o avanço científico em desenvolvimento tecnológico do país. Entretanto, apesar do expressivo crescimento, poucos pedidos de patente em química do estado sólido da China foram depositados para mercados importantes como o norte-americano e o europeu. Dos 1028 depósitos, somente 69 possuem extensão para os EUA e 52 para Europa.

Por outro lado, ao analisarmos o perfil dos depósitos que possuem a Índia como país de origem, observa-se que grande parte desses possuem extensão para mercados importantes. Dos 685 depósitos indianos, 241 possuem extensão para os EUA e 155 para países europeus. Tal fato demonstra o grande potencial da Índia retratado tanto por seu desempenho científico (3º país em número de publicações), quanto pelo seu desempenho tecnológico com extensão dos depósitos nos mercados farmacêuticos mais importantes.

5.1.2.2

Principais depositantes de patentes em química do estado sólido

O

Quadro 5 apresenta os principais depositantes de patentes em química do estado sólido no período de 1997 a 2015. A israelense Teva Pharmaceutical ocupa a primeira posição com 188 depósitos de patentes, seguida da indiana Dr. Reddy's com 110 e a norte-americana Pfizer com 83.

Quadro 5 – Principais depositantes de patente em química do estado sólido e país de origem.

Depositante	Nº Depósitos de patente	País (Sede)
TEVA PHARM USA INC	188	Israel
REDDY'S LAB LTD	110	Índia
PFIZER INC	83	EUA
ASTRAZENECA AB	65	Reino Unido
RANBAXY LAB LTD	65	Índia
NOVARTIS AG	59	Suíça
SANDOZ AG	57	Alemanha
CADILA HEALTHCARE LTD	53	Índia
ABBOTT LAB	49	EUA
ZENTIVA KS	46	República Tcheca
WYETH	43	EUA
GENERICS UK LTD	42	Reino Unido
GLAXO GROUP LTD	41	Reino Unido
MERCK & CO INC	41	EUA
HETERO RES FOUND	40	Índia

Fonte: Derwent Innovations Index. Elaborado pelo autor.

Ao analisarmos o perfil das empresas listadas, destaca-se que grande parte delas têm forte atuação no mercado de medicamentos genéricos. Além da Teva Pharmaceuticals que ocupa atualmente a posição de maior fornecedora de medicamentos genéricos do mundo, podem-se citar ainda Dr. Reddy's, Pfizer e Novartis que aparecem entre as principais empresas produtoras de genéricos em faturamento. Embora algumas dessas empresas também atuem no desenvolvimento fármacos inovadores, tal fato demonstra a relevância da química

do estado sólido no segmento de medicamentos genéricos, não estando limitada ao descobrimento de novos fármacos.

5.2

Indicadores científicos e tecnológicos em química do estado sólido no Brasil

5.2.1

Indicadores científicos em química do estado sólido no Brasil

Nesta seção são apresentados os indicadores científicos relacionados às publicações na área de química do estado sólido com pelo menos um autor brasileiro no período de 1997 a 2016, totalizando 287 publicações. Ao analisarmos a evolução do Brasil, pode-se observar que o número de publicações do país na área ganhou força somente a partir de 2008, tendo apresentado uma baixa atividade no período de 1997 a 2007. Padrão semelhante também pôde ser observado nas publicações de outros países como apresentado anteriormente, refletindo o progressivo ganho de importância da área de química do estado sólido a partir dos últimos anos (Figura 21).

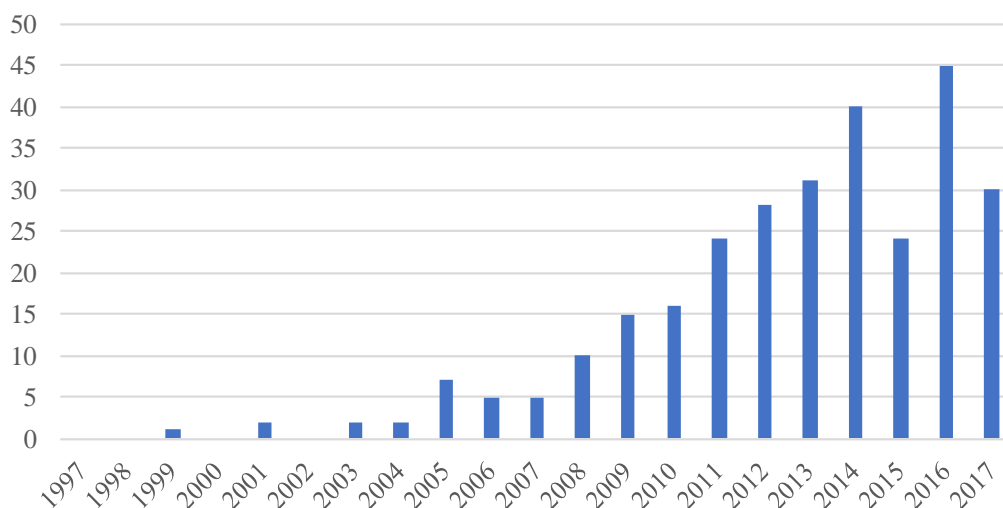


Figura 21 – Evolução das publicações científicas de autores brasileiros em química do estado sólido (1997- 2016).

Fonte: Web of Science. Elaborado pelo autor.

Entre 2008 e 2014, as publicações de autores brasileiros apresentaram um crescimento médio anual de 25,99%. Embora tenha sido observada uma forte

queda em 2015, o número de publicações voltou a crescer em 2016 atingindo o maior número anual de publicações do período analisado (45). Além disso, apesar dos resultados modestos, o Brasil alcançou a 9ª posição entre os países com maior número de publicações científicas no mundo como demonstrado anteriormente.

Com base na classificação de categorias de assunto atribuídas pela base de dados *Web of Science*, as publicações brasileiras em química do estado sólido foram distribuídas em 23 categorias. As categorias de assunto com mais de dez publicações são apresentadas na Figura 22. A distribuição indica uma forte participação nos campos de “Pharmacology & Pharmacy”, “Chemistry, Analytical”, “Chemistry, Physical” e “Chemistry, Multidisciplinary”. Ao compararmos o perfil das publicações brasileiras com o do resto do mundo observa-se uma maior relevância da categoria “Chemistry, Analytical” que, embora seja a segunda maior para as publicações do Brasil, ocupa somente a sexta posição no conjunto de publicações para o restante do mundo.

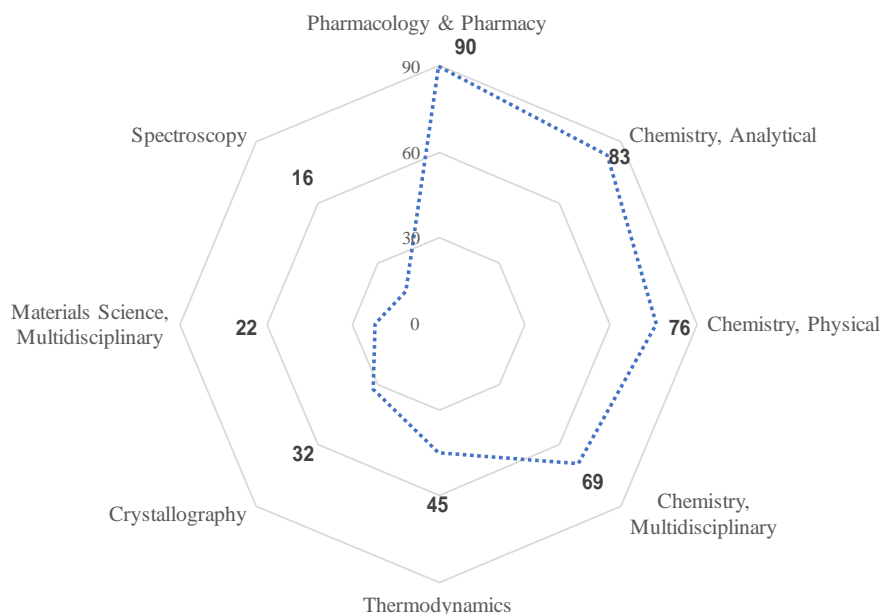


Figura 22 – Distribuição das principais categorias de assunto das publicações em química do estado sólido.

Fonte: Web of Science. Elaborado pelo autor.

Ainda que as categorias forneçam indicativos do perfil da geração de conhecimentos em química do estado sólido por pesquisadores brasileiros, uma visão mais detalhada desse perfil pode ser obtida a partir da análise textual dos registros bibliográficos das publicações a partir das palavras-chave mais

frequentes nas publicações. A Figura 22 apresenta as palavras-chave mais frequentes no conjunto de publicações de autores brasileiros.

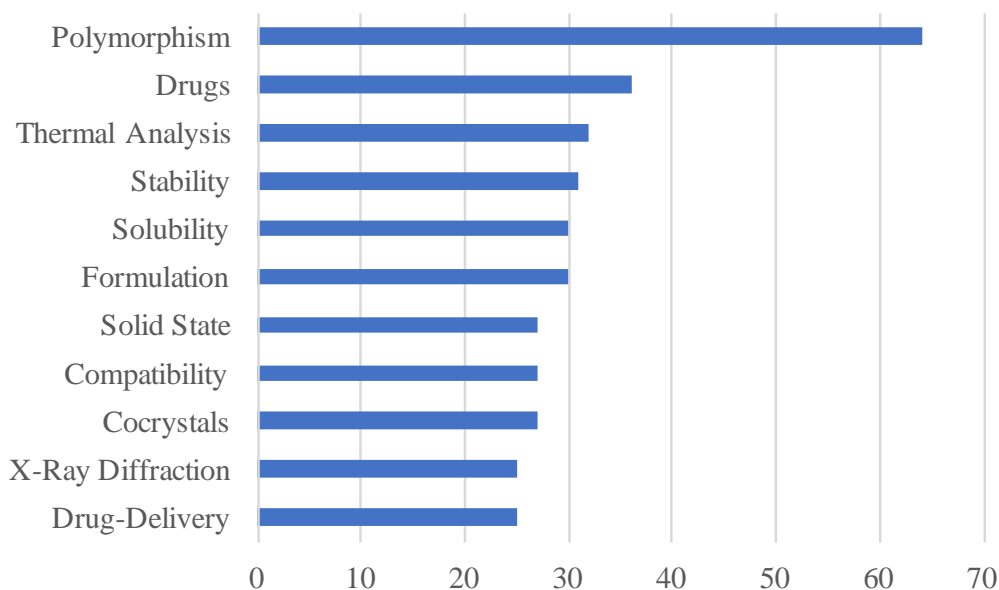


Figura 23 – Palavras-chave mais frequentes nas publicações em química do estado sólido de autores brasileiros.

Fonte: Web of Science. Elaborado pelo autor.

Como pode ser observado, a palavra-chave mais frequente nas publicações de autores brasileiros foi “Polymorphism”, seguido de “Drugs”. Outro termo frequente nas publicações foi “Thermal Analysis”, sendo este termo relacionado às técnicas analíticas baseadas em análise térmica para estudo de propriedades do estado sólido de materiais farmacêuticos. Ainda que se possa observar semelhanças entre o perfil das publicações brasileiras e o restante do mundo no que se refere às palavras-chave mais frequentes, algumas particularidades podem ser destacadas.

De modo geral, a identificação das palavras-chave “Stability”, “Formulation” e “Compatibility” dentre as mais frequentes nas publicações brasileiras indica uma atuação mais intensa de pesquisadores brasileiros no estudo de formulações de produtos farmacêuticos. Em especial, destaca-se a crescente importância da avaliação da compatibilidade entre IFAs e excipientes presentes em formulações com o objetivo de identificar possíveis interações e garantir a estabilidade da formulação final. Nesse aspecto, pode-se ressaltar o recente avanço da legislação brasileira que tornou obrigatório a realização de estudos de compatibilidade entre fármacos e excipientes para o registro de medicamentos por

meio da Resolução da Diretoria Colegiada (RDC) da Anvisa nº 60/2014, reforçando a relação próxima entre a geração de conhecimentos em química do estado sólido e as demandas da indústria farmacêutica como observado anteriormente.

No entanto, a frequência dessas palavras-chave nas publicações brasileiras indica um possível foco maior na avaliação de propriedades de formas sólidas já conhecidas ao invés do estudo de propriedades do estado sólido de novos fármacos descobertos ou processos para obtenção de novos tipos de sólidos para fármacos já conhecidos. Essas atividades apresentam maior complexidade e podem ser identificadas no perfil das publicações no resto do mundo, nas quais palavras-chave como “Solubility”, “Dissolution”, “Crystallization” e “Cocrystals” aparecem entre as mais frequentes.

5.2.1.1

Colaboração internacional da pesquisa brasileira em química do estado sólido

Do total de publicações de autores brasileiros em química do estado sólido, somente 28% tiveram colaboração com pesquisadores de outros países, evidenciando a baixa internacionalização da pesquisa brasileira. Como pode ser observado na Figura 24, entre os principais países com os quais o Brasil possui colaboração internacional destaca-se a Argentina, seguida dos EUA, Reino Unido e Espanha.

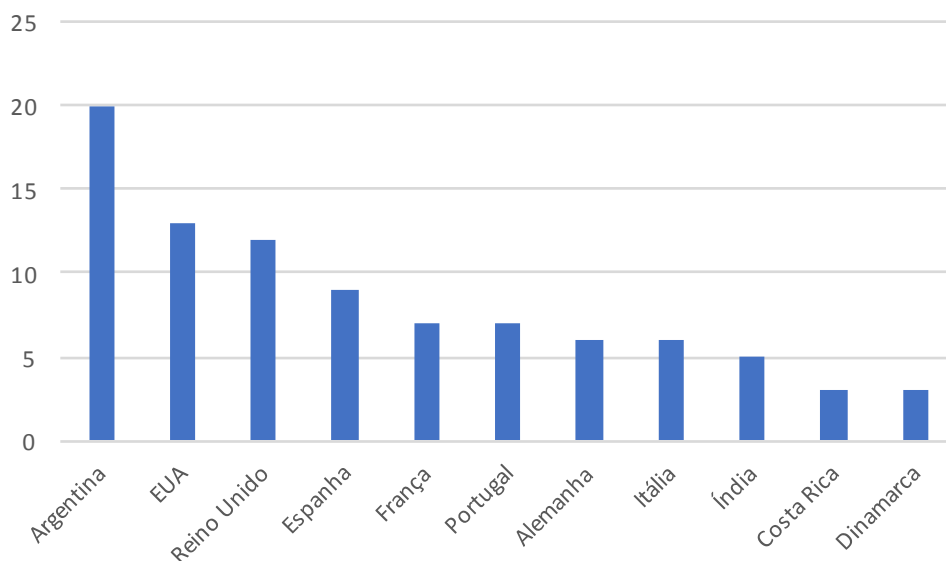


Figura 24 – Principais países com os quais o Brasil possui colaboração em publicações científicas em química do estado sólido.

Fonte: Web of Science. Elaborado pelo autor.

Embora o Brasil apresente um baixo índice de colaboração internacional, o número de publicações conjuntas com pesquisadores de outros países vem subindo ao longo dos últimos anos. A Figura 25 apresenta a evolução das publicações totais de pesquisadores brasileiros em relação a evolução das publicações com colaboração internacional. Como pode ser observado, as primeiras publicações em parceria ocorreram somente em 2004 e, a partir desse ano, subiram consideravelmente e em 2016 representaram aproximadamente 31% do total de publicações de pesquisadores brasileiros.

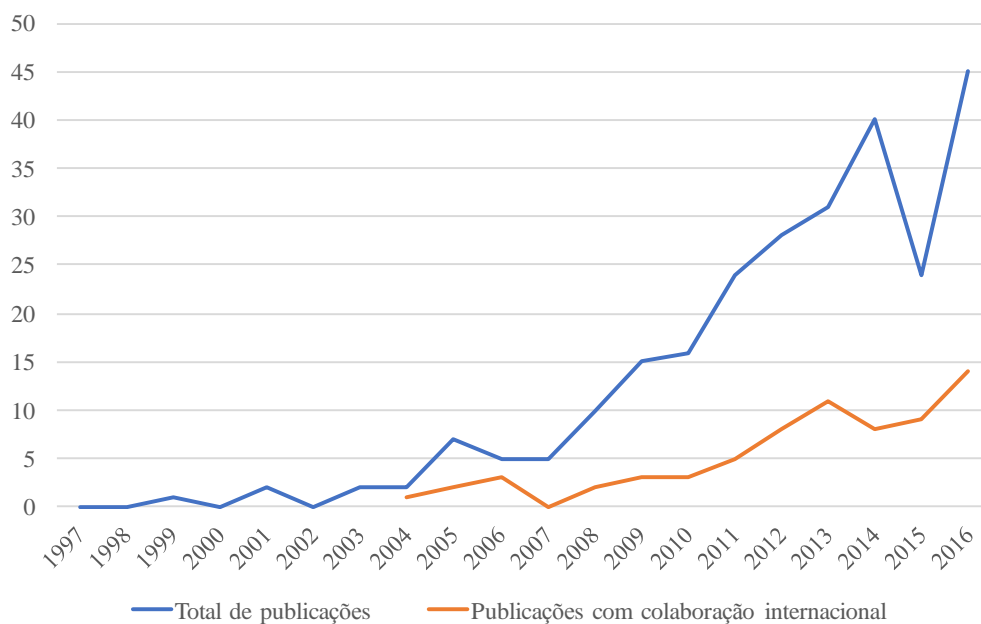


Figura 25 – Evolução do total de publicações de autores brasileiros e publicações com colaboração internacional.

Fonte: Web of Science. Elaborado pelo autor.

5.2.1.2

Principais organizações em publicações científicas na área de química do estado sólido no Brasil

Nesta subseção são apresentados os indicadores referentes as quinze principais organizações em número de publicações científicas na área de química do estado sólido na indústria farmacêutica no Brasil para o período de 1997 a 2016. Como pode ser observado na Figura 26, a Universidade de São Paulo (USP) aparece no topo da lista com 76 publicações, sendo responsável por aproximadamente 26% de todas as publicações de autores brasileiros. Em seguida estão a Universidade Estadual de Campinas (Unicamp) com 35 publicações, a Universidade Estadual Paulista (UNESP) com 32 e a Universidade Federal de Santa Catarina (UFSC) com 30.

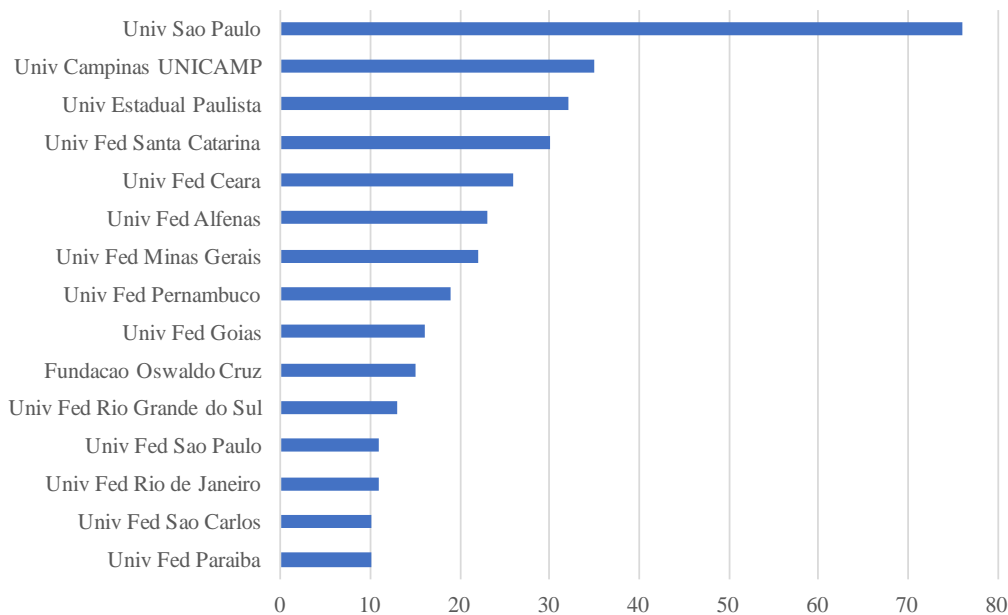


Figura 26 – Principais organizações em publicações científicas em química do estado sólido no Brasil.

Fonte: Web of Science. Elaborado pelo autor.

As organizações listadas somente uma não se caracteriza como universidade, sendo esta a Fundação Oswaldo Cruz (Fiocruz) localizada no estado do Rio de Janeiro. Enquanto instituição de pesquisa e desenvolvimento, a Fiocruz possui importante atuação no cenário farmacêutico brasileiro, sendo responsável pelo desenvolvimento de diversos projetos estratégicos na área farmacêutica. Em especial, destaca-se a atuação da unidade técnico-científica intitulada Instituto de Tecnologia em Fármacos (Farmanguinhos).

A Figura 27 apresenta a evolução das principais organizações brasileiras em número de publicações em química do estado sólido ao longo do período analisado. Como pode ser observado, A Universidade Federal do Rio de Janeiro (UFRJ), a Universidade Federal da Paraíba (UFPB) e a Universidade Estadual de Campinas (Unicamp) foram responsáveis pelas primeiras publicações no período, mas não mantiveram atividade constante nos anos seguintes. A Universidade de São Paulo (USP), por sua vez, manteve desde 2005 atuação em publicações na área tendo apresentado um forte crescimento a partir de 2012. Além disso, pode-se observar que a Universidade Federal de São Paulo (UNIFESP) é um novo entrante na área, tendo iniciado suas publicações somente em 2014.

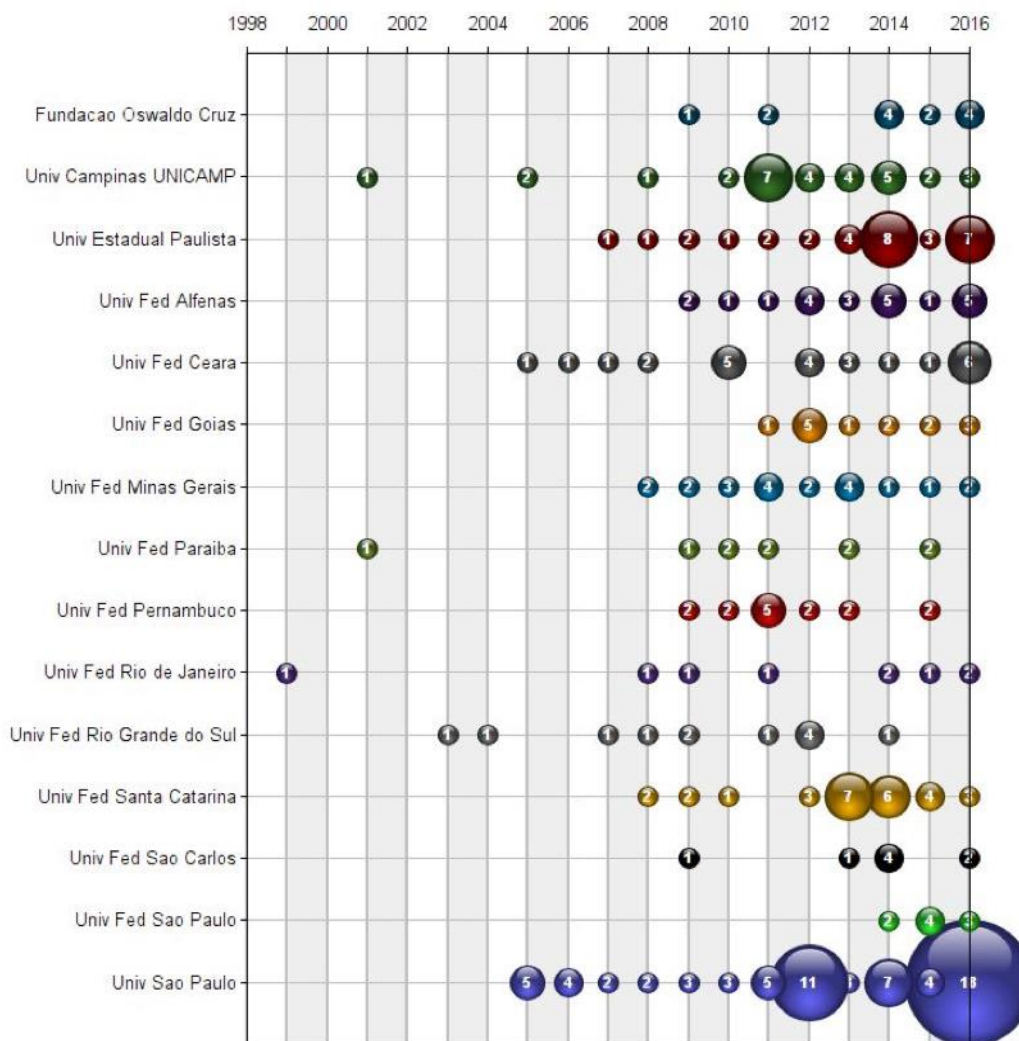


Figura 27 – Evolução das principais organizações brasileiras em número de publicações em química do estado sólido (1997-2016).

Fonte: Web of Science. Elaborado pelo autor.

A Figura 28 apresenta a rede de colaboração entre as principais organizações identificadas neste estudo. Como pode ser observado, existe uma grande interação entre as universidades brasileiras de diferentes regiões no que se refere a pesquisa em química do estado sólido, sendo possível identificar algumas redes mais fortes de interação. Como exemplo, pode-se destacar a colaboração entre a Universidade de São Paulo (USP), a Universidade Federal de Goiás (UFG) e a Universidade Federal de Alfenas (UNIFAL), a qual envolve universidades de três estados distintos.

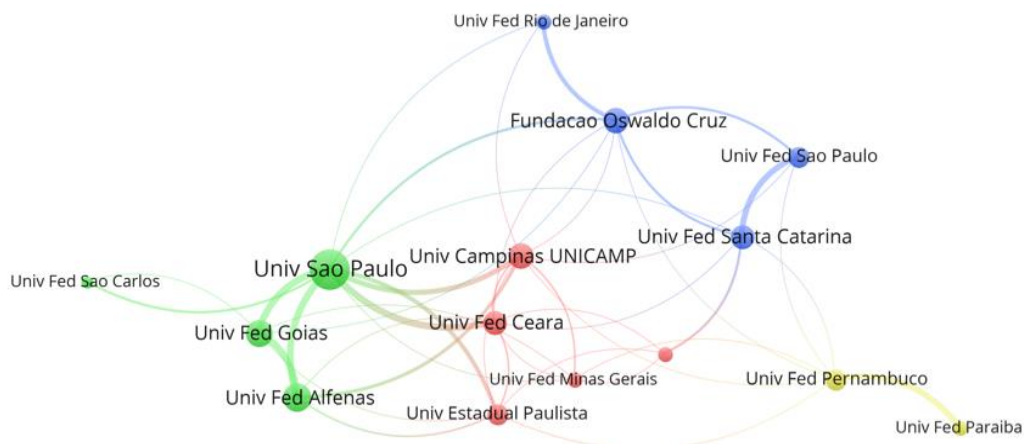


Figura 28 – Rede de colaboração entre as principais organizações brasileiras em publicações em química do estado sólido.

Outra rede de colaboração que pode ser identificada na Figura 28 é formada pela interação entre a Universidade Federal de Santa Catarina (UFSC), a Universidade Federal de São Paulo (UNIFESP) e a Fundação Oswaldo Cruz (Fiocruz). Uma análise mais detalhada das publicações desse grupo de organizações indica que suas pesquisas conjuntas envolvem duas linhas principais, sendo estas o efeito do polimorfismo e do tamanho de partícula nas propriedades do IFA Efavirez e o estudo de cocristais.

5.2.2

Indicadores tecnológicos em química do estado sólido no Brasil

Esta subseção apresenta a análise dos indicadores tecnológicos do patenteamento em química do estado sólido no Brasil. Dentre os 4.665 documentos de patente recuperados a partir da estratégia de busca empregada nesta pesquisa para o período de 1997 a 2015, somente cinco foram depositados por residentes brasileiros (Quadro 6). A Universidade de São Paulo (USP) foi responsável por três depósitos de patente, enquanto a Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio) e a Universidade Federal de Minas Gerais (UFMG) foram responsáveis por um depósito cada.

Quadro 6 – Principais depositantes de patente em química do estado sólido no Brasil.

Nº da patente	PI 0206723-4
Título	FASES LÍQUIDO-CRISTALINAS (LAMELAR E CÚBICA) COMO SISTEMA DE LIBERAÇÃO PARA USO TÓPICO DAS PRÓ-DROGAS: ÁCIDO 5-AMINOLEVULÍNICO (5-ALA) E SEUS ÉSTERES DERIVADOS METIL, BUTIL, HEXIL, OCTIL E DECIL, USADOS NA TERAPIA FOTODINÂMICA (TFD) E/OU FOTODIAGNÓSTICO (FD)
Depositante	Universidade de São Paulo - USP (BR/SP)
Ano de depósito	2002
Inventores	Antonio Claudio Tedesco / Rozane de Fátima Turchiello / Flávia Cristina Bonilha Vena / Maria Vitória Lopes Badra Bentley / Regilene Steluti

Nº da patente	PI 0904938-0
Título	MÉTODO DE OBTENÇÃO DE UM CO-CRISTAL DE LAMIVUDINA, SEU SAL MALEATO, FORMULAÇÕES FARMACÊUTICAS E SEU USO
Depositante	Universidade de São Paulo - USP (BR/SP) / Fundação de Amparo a Pesquisa do Estado de São Paulo - FAPESP (BR/SP)
Ano de depósito	2009
Inventores	Javier Alcides Ellena / Felipe Terra Martins / Nikolas Papparis

Nº da patente	PI 0907290-0
Título	MÉTODO DE OBTENÇÃO DE NOVA FORMA CRISTALINA DE LAMIVUDINA, SEU SAL CLORIDRATO, FORMULAÇÕES FARMACÊUTICAS E SEUS USOS
Depositante	Universidade de São Paulo - USP (BR/SP) / Fundação de Amparo a Pesquisa do Estado de São Paulo - FAPESP (BR/SP)
Ano de depósito	2009
Inventores	Javier Alcides Ellena / Felipe Terra Martins / Nikolas Papparis

Nº da patente	BR 10 2013 033006 0
Título	MÉTODO DE PREPARAÇÃO DA FORMA CRISTALINA ANIDRA DE HIDRAZONA DERIVADA DA ISONIAZIDA, POLIMORFO CRISTALINO DA FORMA ANIDRA PRODUZIDO, USO DO MESMO PARA O TRATAMENTO DAS DOENÇAS DE ALZHEIMER, PARKINSON E OUTRAS DESORDENS NEURODEGENERATIVAS E COMPOSIÇÃO FARMACÊUTICA
Depositante	Faculdades Católicas Mantenedora da Pontifícia Universidade Católica do Rio de Janeiro - PUC RIO (BR/RJ)
Ano de depósito	2013
Inventores	Nicolás Adrián Rey / Leonardo Viana De Freitas

Nº da patente	WO 2006000069
Título	PHARMACEUTICAL COMPOSITIONS CONSISTING OF CYCLODEXTRIN AND AN ANTIMONY DERIVATIVE, THEIR PREPARATION AS WELL AS THEIR USE
Depositante	Universidade Federal de Minas Gerais – UFMG (BR/MG)
Ano de depósito	2004
Inventores	Ruben Dario Sinisterra Millan / Rosemary Ochoa Bejarano / Cynthia Peres Demichele / Frederic Jean Georges Frezard / Alan Lane de Melo / Lucas Antonio Miranda Ferreira

Ao analisarmos os depósitos de residentes brasileiros, pode-se os dois depósitos da Universidade de São Paulo (USP) envolvendo o estado sólido do IFA lamivudina, um agente antirretroviral utilizado para o tratamento de infecção por HIV como fármaco isolado ou como adjuvante no tratamento. Os depósitos da USP envolvem um método de obtenção de uma nova forma cristalina de lamivudina e um método de obtenção de um co-cristal da lamivudina. Tendo em vista que a lamivudina é utilizada hoje como parte do coquetel básico oferecido pelo Sistema Único de Saúde (SUS). Depositadas em 2009, as duas patentes ainda aguardam análise por parte do Instituto Nacional da Propriedade Industrial (INPI).

Apesar da grande relevância da atuação da USP nos desenvolvimentos relacionados à lamivudina, nenhum dos depósitos possuem extensão para outros países, tendo sido depositados somente no escritório brasileiro. Tal fato restringe o potencial de apropriação dessas invenções já que a falta de proteção em outros países importantes como Estados Unidos diminui o interesse de exploração. Por outro lado, o depósito da Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio) envolvendo método de preparação da forma cristalina anidra de hidrazona possui extensão do depósito para o Escritório Europeu e Estados Unidos, evidenciando uma maior força da patente depositada.

De modo geral, os poucos depósitos de patente de residentes brasileiros em química do estado sólido retratam a dificuldade do país em converter seu desempenho científico em tecnologias para o mercado. Embora ocupe a 9ª posição em número de publicações científicas, o Brasil não apresenta atuação relevante em depósitos de patente sendo necessário o desenvolvimento de ações mais assertivas para que se possa mudar esse cenário.

5.3

Identificação de competências: estudo de caso acetato de dexametasona

Esta subseção apresenta a análise de lideranças e competências específicas em química do estado sólido, no mundo e no Brasil, a partir dos resultados obtidos pelo método NanoBusiness Researcher Quality Index (NB-RQ). São apresentadas e analisadas abaixo os Top10 dos escores que avaliam o pesquisador sob cinco perspectivas diferentes ($5 V's = V_{pub}^*, V_{val}^*, V_{acad}^*, V_{ind}^*$ e V_{proj}^*), bem como pelo

Index NB-RQ proposto nessa tese. Os Top10 foram obtidos a partir dos 20.625 e 1.014 pesquisadores identificados no mundo e no Brasil respectivamente, com competência em química do estado sólido com base na metodologia descrita no Capítulo 4 dessa tese.

Sabendo que cada w_i representa o peso ou grau de importância de cada score para compor o NB-RQ, os w_i utilizados para a identificação de lideranças e competências específicas no mundo e no Brasil receberam os seguintes valores:

$$NB - RQ(\text{mundo})$$

$$= 0,25V_{pub}^* + 0,10V_{val}^* + 0,05V_{acad}^* + 0,35V_{ind}^* + 0,25V_{proj}^*$$

$$NB - RQ(\text{Brasil}) = 0,05V_{pub}^* + 0,05V_{val}^* + 0,05V_{acad}^* + 0,75V_{ind}^* + 0,10V_{proj}^*$$

Como pode ser observado, foram aplicados pesos diferenciados na identificação de lideranças e competências específicas globais (mundo) e local (Brasil). Esses pesos foram estabelecidos após a análise de sensibilidade do processo de otimização NB-RQ que teve como objetivo encontrar os pesquisadores ótimos entre os pesquisadores na dimensão global e local.

Devido a limitações na identificação de pesquisadores homônimos, particularmente chineses, uma restrição ao modelo proposto nessa tese foi estabelecida. Desse modo, foi excluído manualmente todos os pesquisadores homônimos de origem chinesa da lista de pesquisadores globais (mundo) e local (Brasil).

5.3.1

Identificação de lideranças e competências científicas no Mundo e no Brasil

5.3.1.1

Identificação de lideranças e competências científicas em volume de publicações no mundo

Dentre as cinco diferentes perspectivas para avaliação dos pesquisadores, a Figura 29 apresenta o Top10 de lideranças e competências específicas em

química do estado sólido no mundo com maior volume de publicação (V_{pub}^*) entre os 20.625 pesquisadores identificados.



Figura 29 – Pesquisadores Top10 pela perspectiva (V_{pub}^*) Volume de Publicações – Mundo.

Os quatro primeiros pesquisadores possuem vínculo com as principais organizações que apresentaram, na subseção 5.1.1.2, os maiores indicadores de número de publicações científicas na área de química do estado sólido (University of Copenhagen, University of Cambridge e Purdue University). O volume de publicações do professor Tomas Rades é um indicador de sua liderança e contribuição para o desempenho científico da University of Copenhagen na área de química do estado sólido no mundo. Seu interesse de pesquisa está associado a formulação e administração de medicamentos e caracterização física do estado cristalino sólido e líquido da matéria. As pesquisas de Rades nessas áreas visa melhorar a terapia de drogas através da formulação adequada de medicamentos, buscando aumentar a compreensão das propriedades físico-químicas dos medicamentos, bem como combina as ciências e tecnologias físicas, químicas e biológicas para formulação e otimização de drogas para uso humano e veterinário. Desde 1999 o pesquisador Tomas Rades é consultor da Divisão de

Desenvolvimento e Formulação Pré-Clínica da Roche na Suíça e atua também desde 2007 como Membro do Conselho Científico da empresa australiana de biotecnologia, Phosphagenics. Essa interação universidade-empresa, portanto, pode explicar o seu maior index NB-RQ quando comparada aos outros professores listados entre o Top10 – Volume de Publicações.

5.3.1.2

Identificação de lideranças e competências científicas em validade das publicações no mundo

Sabemos que o escore V_{val}^* , ou seja, Validade das Publicações, é computado a partir da média do fator de impacto dos periódicos em que o pesquisador publicou seus artigos. Desse modo, a Figura 30 apresenta o Top10 de lideranças e competências específicas em química do estado sólido no mundo.

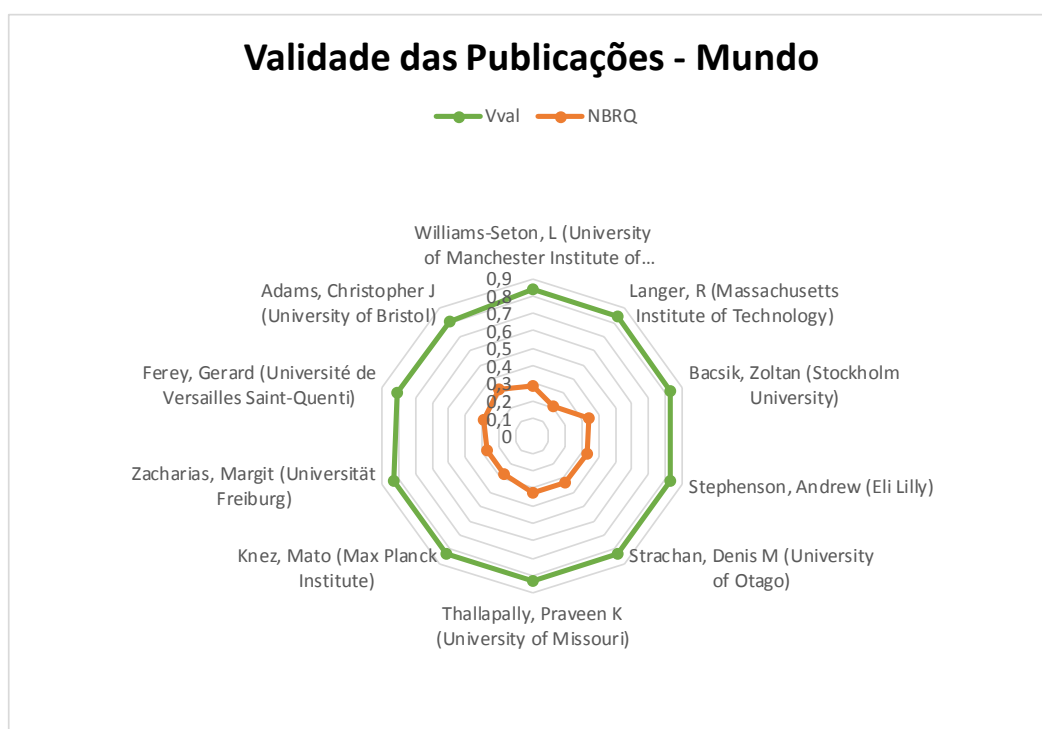


Figura 30 – Pesquisadores Top10 pela perspectiva (V_{val}^*) Validade das Publicações – Mundo.

Observa-se que não existe entre os Top 10 um pesquisador líder na perspectiva Validade das Publicações, sendo seus respectivos escores $V_{val}^* = 0,8$ em média. Por outro lado, apesar da professora da Maragit Zacharias da

Universität Freiburg pertencer a esse grupo estritamente acadêmico, sua atual função como Vice-Presidente para Inovação e Transferência de Tecnologia de uma universidade fundada em 1457 pode indicar as iniciativas promovidas na Alemanha para aumentar sua taxa de conversão dos avanços científicos em desenvolvimento tecnológico frente ao forte desempenho de países como os EUA, China e Índia, conforme apresentado na subseção 5.1.2.1 – Figura 19.

Entre esses líderes na promoção de fatores de impacto em química do estado sólido identificados destaca-se a presença do engenheiro químico Robert S. Langer. Atualmente Langer é o professor do Instituto David H. Koch do MIT. O professor Langer escreveu mais de 1.400 artigos, além de possui mais de 1.300 patentes que estão licenciadas ou sublicenciadas para mais de 350 empresas farmacêuticas, químicas, de biotecnologia e de dispositivos médicos. Ele é o engenheiro mais citado da história (h-index 246 de acordo com o Google Scholar). Apesar do excelente desempenho do Dr. Robert Langer no meio acadêmico e na indústria, Langer possui o menor index-NBRQ, especificamente na área de química do estado sólido, quando comparados aos outros pesquisadores integrantes do Top10 – Validade das Publicações. Esse resultado se deve ao fato que o professor Langer não possui estritamente pesquisas em química do estado sólido e tão pouco patentes associadas especificamente a acetato de dexamatasona.

5.3.1.3

Identificação de lideranças e competências científicas em valor acadêmico no mundo

Ao lembramos que o valor acadêmico do pesquisador está intimamente ligado em quanto a sua produção científica é citada pelos seus demais pares da comunidade, não é difícil entender o impacto dos pesquisadores da Pfizer no curva do escore V_{acad}^* - Valor Acadêmico, conforme apresentado na Figura 31.

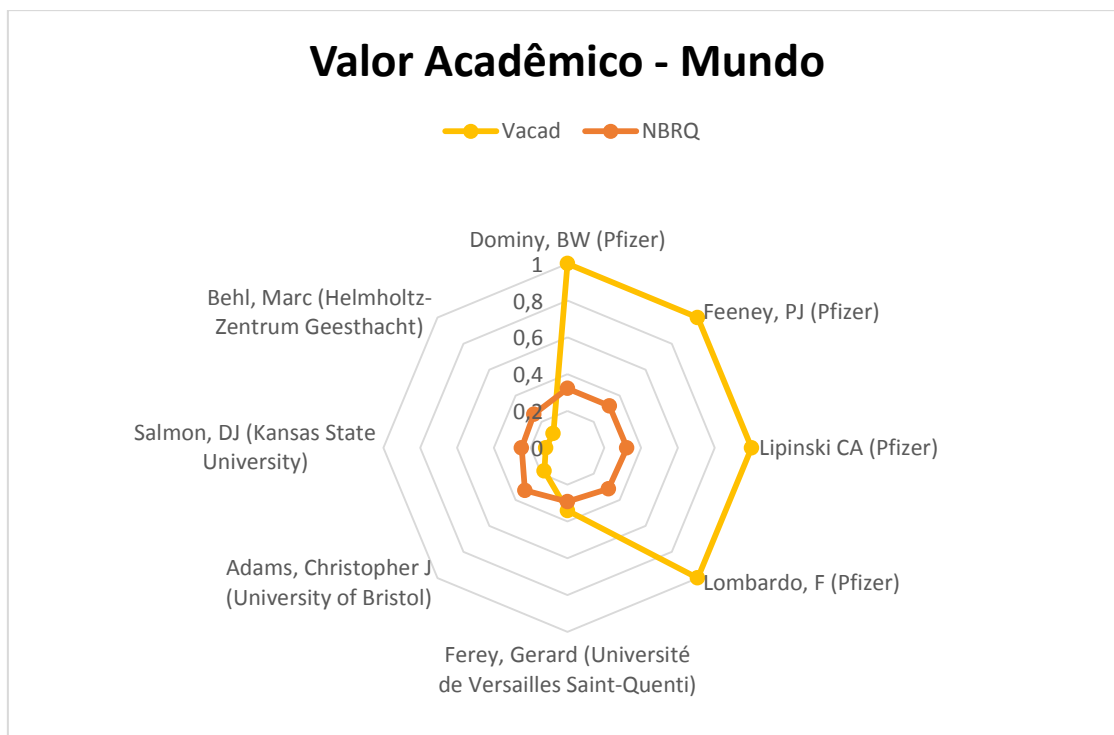


Figura 31 – Pesquisadores Top10 pela perspectiva (V_{acad}^*) Valor Acadêmico – Mundo.

Os três pesquisadores da Pfizer (Christopher A. Lipinski, Franco Lombardo, Beryl W. Dominy, Paul J. Feeney) possuem juntos 3.866 citações acumuladas no período entre 1997 a 2016 com apenas a publicação de um artigo, intitulado “*Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings*”. Desse modo, os três pesquisadores são líderes no número de citações na área de química do estado sólido pela comunidade científica e tecnológica mundial. Por outro lado, não foi possível identificar entre esses líderes ou qualquer pesquisador do Top10 algum que tenha qualquer nível de interação com a indústria farmoquímica e farmacêutica.

5.3.1.4

Identificação de lideranças e competências científicas em valor industrial no mundo

Dentre os 5 escores que avaliam o pesquisador sob perspectivas diferentes ($5 V's = V_{pub}^*, V_{val}^*, V_{acad}^*, V_{ind}^*$ e V_{proj}^*), o valor industrial (V_{ind}^*) é o único escore obtido a partir da extração de informações da base de dados de patentes, tendo

vido utilizados os dados de depósitos de patentes analisados ao longo da pesquisa e extraídos da base *Derwent Innovations Index*. Desse modo, o escore V_{ind}^* é uma variável relevante – necessário, mas não exclusiva e tão pouco suficiente – na identificação de lideranças e competências científicas com capacidade de transformação de conhecimento em aplicações tecnológicas. A Figura 32 apresenta o Top10 entre os 20.625 pesquisadores identificados.

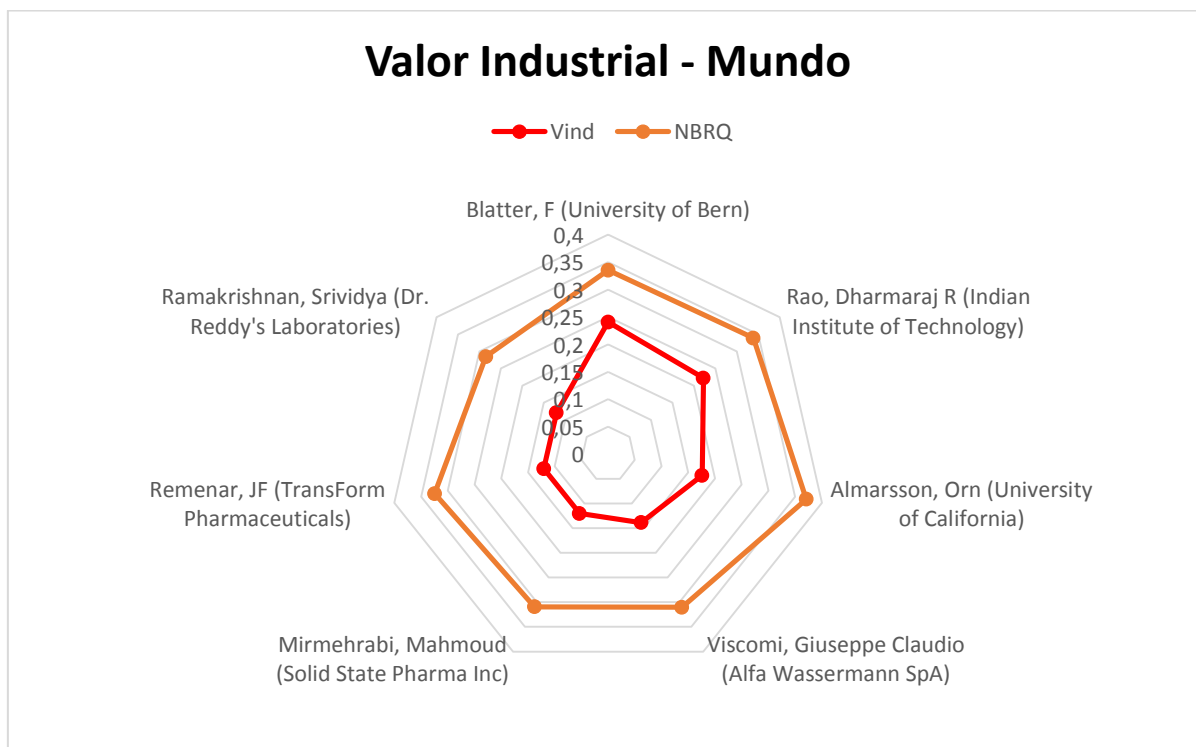


Figura 32 – Pesquisadores Top10 pela perspectiva (V_{ind}^*) Valor Industrial – Mundo.

Verifica-se que o escore do valor industrial (V_{ind}^*) de todos os pesquisadores são inferiores aos seus receptivos index NB-RQ. Essa característica indica que apenas avaliar o pesquisador pelo seu número de patentes não é suficiente para a identificação de líderes e competências científicas com a capacidade de geração de produtos, empresas e emprego de alto valor agregado e baseado na ciência contribuído, assim, para o aumento do nível de competitividade de um país.

Ainda assim, podemos identificar como líder entre o Top10 o pesquisador Fritz Blatter. Com doutorado na University of Bern, Fritz trabalha por mais de 15 anos em química de estado sólido de drogas, polimorfismo, cristalização, ciência de superfície e resolução de problemas relacionados a esse assunto. Ele realizou

estudos sobre mais de 100 insumos farmacêuticos ativos, vários deles agora são *blockbusters*. Possui experiência de pesquisa e atividade gerencial em empresas como IBM, Novartis e Berkeley Lab. Atualmente é gerente de projetos da Solvias.

5.3.1.5

Identificação de lideranças e competências científicas no mundo com valor para o caso acetato de dexametasona

O V_{proj}^* - Valor do Projeto é estimado a partir do grau de afinidade da produção científica do pesquisador com os tópicos específicos de interesse definidos dentro do tema de estudo de caso do acetato e dexametasona. Foram definidos ainda como tópicos de interesse desse caso os tópicos Polimorfismo e Sais. Sendo assim, as publicações foram correlacionais com os tópicos supracitados, permitindo identificar lideranças e competências científicas conforme Figura 33.



Figura 33 – Pesquisadores Top10 pela perspectiva (V_{proj}^*) Valor do Projeto – Mundo.

O fato de ser pesquisador do Indian Institute of Technology e dirigente de uma equipe de 145 cientistas de uma empresa indiana, a Cipla Limited, atuante na área de desenvolvimento de novos processos para fabricação de ingredientes

farmacêuticos ativos complexos, formas polimórficas alternativas de APIs com mais de 200 patentes e pedidos de patentes, torna o pesquisador Dharmaraj Rao um entre os 10 líderes identificados com competências nos tópicos Polimorfismo e Sais. Hoje a Cipla Limited possui mais de 2.000 produtos em 65 categorias terapêuticas; com mais de 40 formas de dosagem, cobrindo um amplo espectro de doenças que vão desde doenças transmissíveis, não transmissíveis, comuns e emergentes até mesmo doenças raras. A Cipla Limited é uma das maiores empresas farmacêuticas fabricantes de genéricos do mundo com forte presença em mais de 170 países.

5.3.1.6

Identificação de lideranças e competências científicas no mundo – NB-RQ: caso acetato de dexametasona

Mesmo com forte influência dos pesquisadores identificados em Volume de Publicação (V_{pub}^*), com o index NB-RQ foi possível identificar 2 tipos de liderança e competência científica entre os Top10 da Figura 34. O primeiro perfil é caracterizado pelo pesquisador com vínculo com a Universidade e que possui capacidade de resolução de problemas demandados pela indústria farmoquímica e farmacêutica, bem como a busca por depósito de patentes com interesse no seu licenciamento. Nesse caso, podemos classificar os professores Thomas Rades, Jukka Rantanen, William Jones e Lynne Taylor.

O segundo perfil é caracterizado pelo professor-empendedor que concebe e viabiliza empresas ou organizações (com ou sem fins lucrativos) *spin-off* das Universidades e Instituições de Pesquisa. Entre o Top10 da Figura 34 temos os seguintes pesquisadores com esse perfil: Dario Braga, Orn Almarsson, Dhamaraj Rao, Nair Rodriguez-Hornedo e Guy van den Mooter e Fritz Blatter.

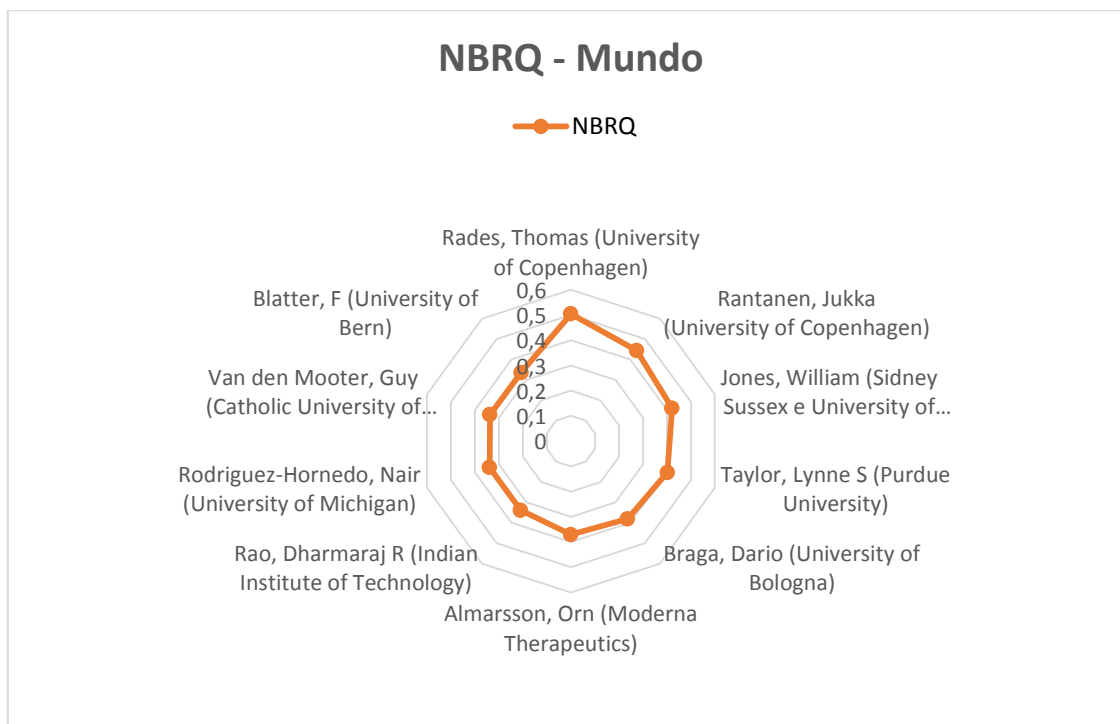


Figura 34 – Pesquisadores Top10 pela perspectiva NB-RQ – Mundo.

Os dois perfis oferecem uma relevante contribuição para seus respectivos países e parceiros. Considerando que já foi descrito o perfil de Thomas Rades, podemos destacar o perfil do líder dos professores-empresendedores, Dario Braga. O Dr. Dario Braga é professor do Departamento de Química da Universidade de Bolonha na Itália, considerada a mais antiga universidade do mundo ocidental. É, também, Coordenador do Programa de Doutorado em Nanociências para Medicina e Meio Ambiente. Seu grupo de pesquisa atua nas seguintes linhas:

- a) Polimorfismo de cristal molecular, investigação de condições de cristalização e nucleação de novas formas polimórficas, co-cristais e solvatadas de cristais moleculares orgânicos e inorgânicos. Esta linha de pesquisa inclui alguns esforços cooperativos e contratos com empresas farmacêuticas;
- b) Concepção e preparação e caracterização de sais hídricos inorgânicos-organometálicos híbridos ligados a hidrogênio para aplicação em tecnologia de células de combustível sólido e investigação dos aspectos teóricos da ligação de hidrogênio entre e com sistemas iônicos;
- c) Design, síntese e caracterização de blocos organometálicos para a preparação de superestruturas à base de sistemas heterometálicos e para a engenharia de estruturas de rede;

- d) Reações isentas de solventes: exploração das reações de gás sólido envolvendo um substrato sólido organometálico que absorve e libera reversivelmente moléculas pequenas;
- e) Reações mecanicamente ativadas entre cristais ou dentro de cristais para formar novos materiais cristalinos ou novos produtos moleculares sem solvente condições;
- f) Preparação de co-cristais iônicos e moleculares
- i) Determinação estrutural a partir de dados de difração do pó

Em 2005, o grupo de pesquisa do professor Braga concebeu uma empresa de *spin-off* (PolyCrystalLine) no campo da química de estado sólido e do polimorfismo de cristais. A empresa é totalmente independente e emprega vários pesquisadores do laboratório do professor Braga.

O Quadro 7 apresenta o vínculo de cada pesquisador identificado pelo método NB-RQ com participação em atividades de interação Universidade-Empresa ou iniciativas empreendedoras no mercado de química do estado sólido.

Quadro 7 – Instituições de pesquisa e empresa com vínculo com os pesquisadores Top10 NB-RQ – Mundo.

Pesquisador	Instituição de Pesquisa	Empresa
Rades, Thomas	University of Copenhagen	Roche e Phosphagenics
Rantanen, Jukka	University of Copenhagen	Consultor de mais de 20 empresa na Europa, EUA e Asia
Jones, William	University of Cambridge	AbbVie
Taylor, Lynne S	Purdue University	SSCI
Braga, Dario	University of Bologna	PolyCrystalLine
Almarsson, Orn	Moderna Therapeutics	Transform Pharmaceuticals (aquirida pela Johnson & Johnson)
Rao, Dharmaraj R	Indian Institute of Technology	Cipla Ltd
Rodriguez-Hornedo, Nair	University of Michigan	Transform Pharmaceuticals (aquiririda pela Johnson & Johnson)
Van den Mooter, Guy	Catholic University of Leuven	FORMAC Pharmaceuticals N.V
Blatter, F	University of Bern	Solvias AG
Rades, Thomas	University of Copenhagen	Roche e Phosphagenics

5.3.1.7

Identificação de lideranças e competências científicas no Brasil – NB-RQ: caso acetato de dexametasona

Como pode ser observado na Figura 35, o Brasil possui líderes com competências científicas em química do estado sólido que poderão colaborar com o projeto acetato de dexametasona. Essa competência está concentrada na Universidade de São Paulo, organização com maior indicador de publicação científica em química do estado sólido no Brasil conforme Figura 25 da subseção 5.2.1.2. Com base na Figura 26 da subseção 5.2.1.2 é possível afirmar que os pesquisadores do Top10 NB-RQ são responsáveis pela forte evolução do número de publicações da USP em química do estado sólido a partir de 2012.

O perfil do Top10 NB-RQ de pesquisadores brasileiros é caracterizado pela sua capacidade de resolução de problemas demandados pela indústria farmacêutica e química, bem como pela busca por depósito de patentes com interesse no seu licenciamento. Por outro lado, apenas um pesquisador entre os Top10 é caracterizado como professor-empresário por conceber e viabilizar empresas ou organizações (com ou sem fins lucrativos) *spin-off* das Universidades e Instituições de Pesquisa. Esse perfil é exclusivo da pesquisadora Selma Gutierrez Antonio da UNESP.

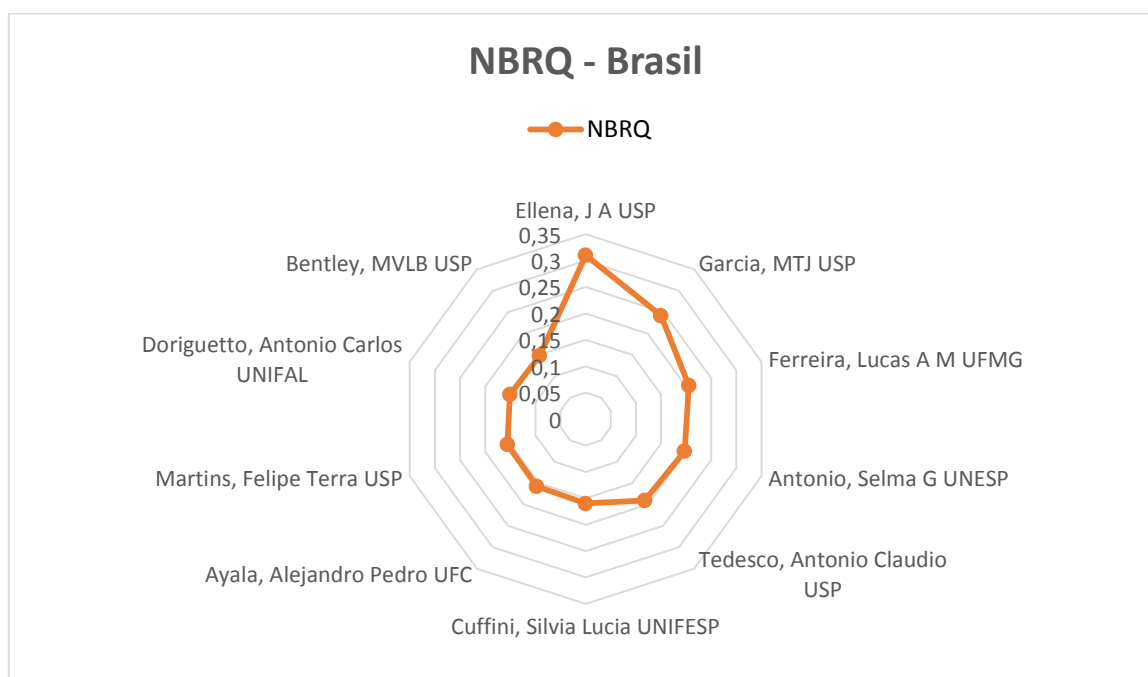


Figura 35 – Pesquisadores Top10 pela perspectiva NB-RQ – Brasil.

A Dra. Selma, possui graduação em química pela Universidade Estadual Paulista Júlio de Mesquita Filho (2003), mestrado em Ciência e Tecnologia de Materiais pela Universidade Estadual Paulista Júlio de Mesquita Filho (2006) e Doutorado na Universidade Estadual Paulista Júlio de Mesquita Filho com projeto de Caracterização polimorfos em fármacos por Difração de raios X e refinamento de estrutura pelo Método de Rietveld. Atualmente é Pós-doc na Universidade Estadual Paulista Júlio de Mesquita Filho com o projeto Controle de Medicamentos por Difração de Raios X. Tem experiência na área de Química, com ênfase em Físico Química e Inorgânica, atuando principalmente nos seguintes temas: Método de Rietveld, indexação e determinação de cristalina com dados de difração de raios X por policristais, análise de polimorfismo de fármacos e análise quantitativa de fase. Sua atuação na Nox Consultoria e Treinamento tem como foco o treinamento do método de Rietveld para profissionais da indústria.

O pesquisador líder do Top10 NB-RQ Brasil é o professor Javier A. Ellena da USP, que tem como linha de pesquisa estudos de difração de raios-x em monocristal de compostos de interesse biológico e farmacológico.

O Grupo de Cristalografia do Instituto de Física de São Carlos (IFSC) é coordenado pelo Dr. Javier que mantém suas atividades de pesquisa e desenvolvimento, ensino e extensão, nas áreas de Cristalografia Estrutural, Cristalografia em Ciência dos Materiais, Polimorfismo em Cristais, Cristalografia de Proteínas, Modelagem Molecular, Biologia Molecular, Planejamento de Fármacos, Química Medicinal e Computacional. O Grupo caracteriza-se pela forte interdisciplinaridade; ótima produtividade científica; grande envolvimento na formação de mestres, doutores e pós-doutores; excelentes níveis de captação de recursos junto às agências de fomento nacionais e internacionais; além do desenvolvimento de projetos de inovação tecnológica em parceria com empresas, especialmente, do ramo farmacêutico.

6

Conclusões

6.1

Química do Estado Sólido

Sendo atendido todos os objetivos estabelecidos para elaboração dessa tese, podemos concluir que a difusão do conhecimento científico em química do estado sólido na indústria farmoquímica e farmacêutica global ocorreu nos últimos 20 anos devido a pressões regulatórias, sendo o emblemático caso do IFA ritonavir em 1998 e o lançamento do guia do *International Committee on Harmonization* (ICH) Q6A em 1999 marcos históricos relevantes.

A compreensão das propriedades do estado sólido de um IFA mostra-se crítica no desenvolvimento de formulações em função de seus impactos na biodisponibilidade e solubilidade dos fármacos. Portanto, a necessidade do controle e monitoramento contínuo do estado sólido na indústria farmacêutica, mostra-se crítica ao considerarmos que aproximadamente 70% dos medicamentos disponíveis são produzidos com formulações sólidas, para uso oral, como comprimidos, cápsulas, drágeas, entre outras. Sendo que dentre de mais de 350 compostos estudados, aproximadamente 90% apresentaram variações de sua forma no estado sólido.

Devido as possíveis transições de formas sólidas de um mesmo IFA ao longo de toda a cadeia, desde a produção do IFA até sua formulação, armazenamento, distribuição e uso, novos desafios surgiram para também para as ciências da farmacovigilância e da metrologia legal. O progressivo ganho de importância da temática da química do estado sólido na indústria farmacêutica, sobretudo no que tange ao controle e monitoramento do polimorfismo, foi diretamente influenciado não somente pela identificação de problemas de produção e na eficácia de medicamentos, mas também pelo avanço das disputas patentárias envolvendo formas polimórficas de IFAs.

Além das pressões regulatórias, a dinâmica competitiva da indústria farmoquímica e farmacêutica com necessidades constatadas de inovações em

produtos e processos ofereceram uma janela de oportunidade para pesquisadores com competências em química do estado sólido protegerem suas inovações. Essas inovações baseadas na cristalografia e engenharia de cristais passaram a contribuir com as modificações de uma ou mais propriedades dos compostos como a estrutura físico-química, método de síntese do IFA, forma farmacêutica, farmacodinâmica e farmacocinética na busca de novos medicamentos.

As oportunidades oferecidas pela química do estado sólido pela sua aplicabilidade no desenvolvimento de novos medicamentos e, sobretudo, no crescente mercado de medicamentos genéricos, passou a ser explorada por países emergentes como a China e Índia. Hoje, estima-se que cerca de dois terços da produção de IFAs destinada à venda no mundo sejam provenientes da Ásia. Por outro lado, a Índia é hoje o maior fornecedor de medicamentos genéricos do mundo, sendo responsável por aproximadamente 20% das exportações globais em volume.

Além da capacidade produtiva da China e da Índia, hoje esses países emergentes estão entre os principais países que depositam o maior número de patentes relacionadas a química do estado sólido. A China destaca-se pelo forte crescimento no número de depósitos de patentes em química do estado sólido nos últimos 20 anos, com um crescimento anual médio de 33,73%. Além disso, deve-se ressaltar que a China apresentou um número de depósitos de patentes (1028) superior ao número de publicações científicas no período do estudo (736) associado a química do estado sólido. Ao contrário da China, a Índia se destaca por possuir estratégia de proteção de suas patentes com extensão para mercados importante como os EUA e países da Europa.

6.2

Da ciência ao mercado

Os indicadores apresentados nesta tese demonstram que a adoção da inovação aberta e a exploração novos modelos de negócio como *joint ventures* entre empresas farmacêuticas e outros atores externos tem se apresentado como alternativas interessantes para aumentar a produtividade da P&D no setor farmacêutico mundial.

A prática de *outsourcing* de diversas etapas da cadeia produtiva e inovativa farmacêutica, sobretudo as com maiores exigências regulatórias e de maior especialização como o caso das competências associadas em química do estado sólido, favorecem a realização de atividades de P&D em parceria com empresas *start-ups* e universidades.

Ainda que o Brasil tenha alcançado a 9ª posição entre os países com maior número de publicações científicas no mundo, a adoção de inovação aberta e práticas de *outsourcing* foram baixos quando comparado com o mundo. Por outro lado, apesar do baixo índices de aplicação desse modelo no Brasil (sintetizado no NB-RQ index apresentado nessa tese), o grande crescimento do mercado de medicamentos genéricos, tem favorecido a criação de competências específicas no Brasil que, no longo prazo, permitirão o avanço em direção a atividades associadas com a geração de produtos inovadores farmacêuticos.

Com cerca de 1.000 pesquisadores com competência em química do estado sólido liderados por pesquisadores que possuem uma importante produção científica em colaboração com organizações internacionais, o Brasil possui atualmente as bases para formulação de uma estratégia de fortalecimento da competitividade da indústria farmoquímica e farmacêutica baseado do desenvolvimento científico e tecnológico da área de química do estado sólido. Essa estratégia deve ser viabilizada com a regulamentação, publicada no último dia 08 de fevereiro de 2018, do Marco Legal da Ciência, Tecnologia e Inovação (Lei nº 13.243/2016) que tem como objetivo desburocratizar as atividades de pesquisa inovação no país, reduzindo a distância entre a ciência e o mercado.

Essa estratégia deve aproveitar o nicho identificado nessa tese, priorizando desse modo as atividades de pesquisa ligadas ao estudo de propriedades do estado sólido de novos fármacos ou processos para obtenção de novos tipos de sólidos para fármacos já conhecidos. Para isso, deve-se garantir uma maior capacitação dos pesquisadores brasileiros em subáreas como solubilidade, dissolução, cristalização e co-cristais. Para isso, os pesquisadores brasileiros devem não só aumentar o índice de colaboração com universidades internacionais como, também, com empresas de qualquer porte. As universidades localizadas em países como Reino Unido, França, Itália, Alemanha e Dinamarca identificadas nessa tese, são indicadas para o estabelecimento de parcerias de P&D frutíferas.

De modo geral, embora possua líderes e competências científicas na área em questão responsáveis por elevar o Brasil para a 9ª posição em número de publicações científicas, o Brasil não apresenta atuação relevante em depósitos de patente sendo necessário o desenvolvimento de ações mais assertivas para que se possa mudar esse cenário. Os poucos depósitos de patente de residentes brasileiros em química do estado sólido retratam a dificuldade do país em converter seu desempenho científico em tecnologias para o mercado.

6.3

Ciência de Dados (Método NB-RQ)

Segundo o conceito de *Research Intelligence* e de *Research Analytics* (CANTÚ-ORTIZ, 2017, p. 2), as atividades de *Research intelligence* comparam as métricas de realização para produzir rankings de vários tipos, dossiês de benchmarking ou declarações de tendências científicas para oferecer feedback aos analistas responsáveis pela avaliação de resultados, avaliação de *throughput* ou alteração de estratégia. A *Research Analytics* é entendida como a confluência de análise de dados aplicada a produtos científicos encontrados em bases de dados bibliométricas e usando metodologias de Cienciometria e *Research Intelligence* de forma prospectiva e prescritiva.

Desse modo, podemos concluir que com base nos resultados apresentados nessa tese, que o método NB-RQ (NanoBusiness Researcher Quality Index) está alinhado aos conceitos de *Research Intelligence* e *Research Analytics* proposta por CANTÚ-ORTIZ. Com os resultados obtidos foi possível mitigar a dificuldade em identificar competências e lideranças específicas quando comparada aos métodos tradicionais, dado que o NB-RQ captura não só a essência da multidisciplinaridade como, também, contempla dados de depósito de patente associados a informações de publicações científicas.

Foi possível demonstrar a robustez do método, bem como a sua capacidade de identificar pesquisadores em áreas específicas, oferecendo soluções para apoio à tomada de decisão e identificação de pesquisadores relevantes para o desenvolvimento do setor farmoquímico e farmacêutico.

Mesmo que os resultados obtidos tenham atendido os objetivos propostos nessa tese, devemos considerar a implementação de melhorias no método NB-RQ

no futuro próximo. Para mitigar as limitações percebidas no critério utilizado para definição de pesos, deve-se buscar uma solução mais sistemática de se estimar os pesos ou implementar no método uma outra forma de calcular o NB-RQ, sem a utilização de pesos, por exemplo, via a utilização de componentes principais.

Seria também interessante incorporar outras dimensões no método NB-RQ como, por exemplo:

- Vint_univ-ind: valor do pesquisador na interação universidade indústria;
- V_mentoria: número de alunos/pesquisador orientado pelo pesquisador e
- V_grupo: número de co-autores que figuram nas pesquisas realizadas pelo pesquisador.

Mesmo com as limitações apresentadas na tese, seria útil usar outras bases de artigos e patentes, além das utilizadas nessa tese. Devido à importância dos resultados científicos e tecnológicos de origem asiática, devemos encontrar ou desenvolver soluções para mitigar o problema dos homônimos.

Por fim, seria interessante utilizar números Fuzzy para estimar o grau de pertinência de uma Palavra-chave e um Tópico. Desse modo, seria possível incluir o volume de incerteza expresso pelos diferentes pesquisadores nessa inferência (ao invés somente de usar a média ponderada dos mesmos), providenciando na saída final da metodologia intervalos de predição da afinidade de um Tópico e um Autor.

6.4

Trabalhos e projetos realizados e em andamento

Os trabalhos e projetos realizados e em andamento associados ao tema dessa tese são apresentados no Apêndice C.

Referências bibliográficas

ABRIMOV, Y. (Ed.). Computational pharmaceutical solid state chemistry. Hoboken: John Wiley & Sons, 2016.

AJE - AMERICAN JOURNAL EXPERTS. **Relatório AJE de Publicações Acadêmicas:** Brasil. Durham, 2016. Disponível em: <<https://www.aje.com/en/arc/dist/docs/Brazil-scholarly-publishing-report-2016-translated.pdf>>. Acesso em: jul 2017.

ANVISA – AGÊNCIA NACIONAL DE VIGILÂNCIA SANITÁRIA. Denominações Comuns Brasileiras. Disponível em: <http://www.anvisa.gov.br/hotsite/farmacopeiabrasileira/conteudo/cont_dcb_glosario.htm> Acesso em 18 set. 2017.

ARCHAMBAULT, E. et al. Comparing bibliometric statistics obtained from the Web of Science and Scopus. **Journal of the American Society for Information Science and Technology**, v. 60, n. 7, p. 1320-1326, 2009.

ARONSON, J. K. Something new every day: defining innovation and innovativeness in drug therapy. **The Journal of Ambulatory Care Management**, v. 31, n. 1, p. 65-68, 2008.

BAUER et al. Ritonavir: an extraordinary example of conformational polymorphism. **Pharmaceutical Research**, v. 18, p. 859-866, 2001.

BERNSTEIN, J. Polymorphism and patents. In: BERNSTEIN, J. **Polymorphism in molecular crystals**. Nova York: Oxford University Press, 2002, p. 297-307.

BHARDWAJ, R. **Control and Prediction of Solid-State of Pharmaceuticals: Experimental and Computational Approaches**. Suíça: Springer Nature, 2016.

BORNMANN, L.; LEYDESDORFF, L. Scientometrics in a changing research landscape. **EMBO Reports**, v.15, n. 12, p. 1228-1231, 2014.

BOUYSSOU, D.; MARCHANT, T. Consistent bibliometric rankings of authors and of journals. **Journal of Infometrics**, v. 4, n. 3, p. 365-378, 2010.

BYRN, S.; ZOGRA, G.; CHEN, X. **Solid-State Properties of Pharmaceutical Materials**. Hoboken: John Wiley & Sons, 2017.

CANTÚ-ORTIZ, F. (Ed.). **Research Analytics: Boosting University Productivity and Competitiveness through Scientometrics**. Boca Raton: Taylor & Francis Group, 2018.

DIMASI, J.; GRABOWSKI, H.; HANSEN, R. Innovation in the pharmaceutical industry: New estimates of R&D costs. **Journal of Health Economics**, v. 47, p. 20-33, 2016.

EVALUATE. Evaluate Pharma World Preview 2015, Outlook to 2020. Disponível em: <<http://info.evaluategroup.com/rs/607-YGS-364/images/wp15.pdf>> Acesso em: 15 set. 2017.

GAVEL, Y.; ISELID, L. Web of Science and Scopus: A journal title overlap study. **Online Information Review**, v. 32, n. 1, p. 8-21, 2008.

GREGOLIN, J. A. et al. Análise da produção científica a partir de indicadores bibliométricos. In: FUNDAÇÃO DE AMPARO À PESQUISA DO ESTADO DE SÃO PAULO. **Indicadores de ciência, tecnologia e Inovação em São Paulo 2004**. São Paulo: FAPESP, 2005.

GSK plc. Discovery Partnerships with Academia. Disponível em:<<http://www.dpac.gsk.com/>> Acesso em: 11 set. 2017.

HANCOCK, B.; PARKS, M. What is the True Solubility Advantage for Amorphous Pharmaceuticals?. **Pharmaceutical Research**, v. 17, n. 4, p. 397-404, 2000.

HIRSCH, J. E. An index to quantify an individual's scientific research output. **Proceedings of the National Academy of Sciences**, v. 102, p. 16569–16572, 2005.

HUANG, L. F.; TONG, W. Q. Impact of solid state properties on developability assessment of drug candidates. **Advanced Drug Delivery Reviews**, v. 56, n. 3, p. 321-334, 2004.

IBGE – INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. Pesquisa de Inovação 2014. Rio de Janeiro: IBGE, 2016.

IBM – INTERNATIONAL BUSINESS MACHINES CORPORATION. IBM Consumer Products Industry Blog. Disponível em:<<https://www.ibm.com/blogs/insights-on-business/consumer-products/2-5-quintillion-bytes-of-data-created-every-day-how-does-cpg-retail-manage-it/>> Acesso em: 16 nov. 2017.

IFPMA – INTERNATIONAL FEDERATION OF PHARMACEUTICAL MANUFACTURERS & ASSOCIATIONS. **The Pharmaceutical Industry and Global Health** - Facts and Figures 2017. Genebra: IFPMA, 2017.

KLAVANS, R.; BOYACK, K. Toward an objective, reliable and accurate method for measuring research leadership. **Scientometrics**, v. 82, p. 539-553, 2010.

LAIRD, T. Polymorphism and impurities. **Organic Process Research and Development**, v. 8, n. 3, p. 301-302, 2004.

LEOPOLD, E.; MAY, M.; PASS, G. Data mining and text mining for science and technology research. In: MOED, H.; GLÄZEL, W.; SCHMOCH, U. **Handbook of quantitative science and technology research: the use of publication and patent statistics in studies of S&T systems**. Nova York: Kluwer Academic Publishers, 2004.

LIMA, D.; SILVEIRA, C. O patenteamento de polimorfos na indústria farmacêutica e o acesso a medicamentos. **Physis**, v. 21, n. 4, p. 1515-1536, 2011.

LLINÀS, A.; GOODMAN, J. M. Polymorph control: past, present and future. **Drug Discovery Today**, v. 13, p. 198-210, 2008.

MAHESHWARI, C. et al. Tailoring aqueous solubility of a highly soluble compound via cocrystallization: Effect of cofomer ionization, pH_{max} and solute-solvent interactions. **CrystEngComm**, v. 14, n. 14, p. 4801-4811, 2012.

MALERBA, F.; ORSENIGO, L. Innovation and market structure in the dynamics of the pharmaceutical industry and biotechnology: towards a history-friendly model. **Industrial and Corporate Change**, v. 11, n. 4, p. 667-703, 2002.

MARZANO, F. Políticas de Inovação no Brasil e nos Estados Unidos: a busca da competitividade – oportunidades para a ação diplomática. Brasília: Fundação Alexandre Gusmão, 2011.

MILANEZ, D. **Elaboração de Indicadores de Ciência e Tecnologia para o Monitoramento de Avanços Tecnológicos em Nanocelulose**. São Carlos, 2015. 221p. Centro de Ciências Exatas e de Tecnologia, Universidade Federal de São Carlos.

MILES, I.; SARITAS, O.; SOKOLOV, A. **Foresight for Science, Technology and Innovation**. Suíça: Springer Nature, 2016.

MOED, H.; GLÄZEL, W.; SCHMOCH, U. **Handbook of quantitative science and technology research: the use of publication and patent statistics in studies of S&T systems**. Nova York: Kluwer Academic Publishers, 2004.

MONGEON, P.; PAUL-HUS, A. The journal coverage of Web of Science and Scopus: a comparative analysis. **Scientometrics**, v. 106, n. 1, p. 213-228, 2016.

MORAES, F. T. Brasil aumenta produção científica, mas impacto dos trabalhos diminui. **Folha de São Paulo**, São Paulo, 16 de out de 2017. Disponível em: <<http://www1.folha.uol.com.br/ciencia/2017/10/1927163-brasil-aumenta-producao-cientifica-mas-impacto-dos-trabalhos-diminui.shtml>> Acesso em: 10 nov. 2017.

MURDANDE, et al. Solubility advantage of amorphous pharmaceuticals: II. Application of quantitative thermodynamic relationships for prediction of solubility enhancement in structurally diverse insoluble pharmaceuticals. **Pharmaceutical Research**, v. 27, n. 12, p. 2704-2714, 2010.

PARANHOS, J. **Interação entre empresas e instituições de ciência e tecnologia no sistema farmacêutico de inovação brasileiro: estrutura, conteúdo e dinâmica.** Rio de Janeiro, 2010. 343p. Instituto de Economia, Universidade Federal do Rio de Janeiro.

PAUL, et al. How to Improve R&D productivity: the pharmaceutical industry's grand challenge. **Nature Reviews Drug Discovery**, v. 9, p. 203-214, 2010.

PRADO, L. D.; ROCHA, H. V. Estado Sólido na Indústria Farmacêutica: Uma Breve Revisão. **Revista Virtual de Química**, v. 7, n. 6, p. 2080-2112, 2015.

PUDIPEDDI, M.; SARAJUDDIN, A. Trends in Solubility of Polymorphs. **Journal of Pharmaceutical Sciences**, v. 94, n. 5, p. 929-939, 2005.

QuintilesIMS INSTITUTE. **Outlook for Global Use of Medicines through 2021.** Parsippany, 2016. Disponível em: <<https://www.iqvia.com/-/media/iqvia/pdfs/institute-reports/global-outlook-for-medicines-through-2021.pdf>> Acesso em: 5 set 2017.

REPÚBLICA FEDERATIVA DO BRASIL. A Revisão da Lei de Patentes: inovação em prol da competitividade. Brasília: Câmara dos Deputados, Edições Câmara, 2013. Disponível em:< <http://www2.camara.leg.br/a-camara/estruturaadm/altosestudios/temas/seminarios/lancamento-patentes-9-10-13/a-revisao-da-lei-de-patentes>>. Acesso em: 23 nov. 2017.

SCHUHMACHER, A.; GASSMANN, O.; HINDER, M. Changing R&D models in research-based pharmaceutical companies. **Journal of Translational Medicine**, v. 14, n. 105, p. 1-11, 2016.

SILVA, et al. Validation of the method of quantitative phase analysis by X-ray diffraction in API: case of Tibolone. **Journal of Physics: Conference Series**, v. 733, n. 1, 2016.

SINDUSFARMA. Vendas Mercado Farmacêutico, 2017. Disponível em:< [://sindusfarma.org.br/cadastro/index.php/site/ap_indicadores](http://sindusfarma.org.br/cadastro/index.php/site/ap_indicadores)> Acesso em: 10 nov. 2017.

STAHLY, G. P. Diversity in single- and multiple-component crystals. The search for and prevalence of polymorphs and cocrystals. **Crystal Growth Design**, v. 7, n. 6, p. 1007-1026, 2007.

VAN ECK, N.; WALTMAN, L. Text mining and visualization using VOSviewer. Disponível em:<<https://arxiv.org/abs/1109.2058>>. Acesso em: 18 nov. 2017.

WIPO – WORLD INTELLECTUAL PROPERTY ORGANIZATION. Intellectual Property Statistics. Disponível em:<<http://www.wipo.int/ipstats/en/>>. Acesso em 18 nov. 2017.

WOEGINGER, G. J. An axiomatic analysis of Egghe's g-index. **Journal of Informetrics**, v. 2, n. 4, p. 364-368, 2008.

ZHANG, et al. Comparing keywords plus of WOS and author keywords: A case study of patient adherence research. **Journal of the Association for Information Science and Technology**, v. 67, n. 4, p. 967-972, 2016.

ANEXO A



ELSEVIER

Contents lists available at ScienceDirect

Journal of Pharmaceutical Sciences

journal homepage: www.jpharmsci.org

Pharmaceutics, Drug Delivery and Pharmaceutical Technology

New Polymorph Form of Dexamethasone Acetate



Ronaldo Pedro da Silva^{1,2}, Mateus Felipe Schuchter Ambrósio^{1,3},
 Luciana Almeida Piovesan³, Maria Clara Ramalho Freitas⁴,
 Daniel Lima Marques de Aguiar^{5,6}, Bruno Araújo Cautiero Horta⁷,
 Eugenio Kahn Epprecht², Rosane Aguiar da Silva San Gil⁷, Lorenzo do Canto Visentin^{3,*}

¹ Universidade Federal do Rio de Janeiro, Campus de Xerém, Estrada de Xerém, N° 27, Xerém, Duque de Caxias, Rio de Janeiro 25245-390, Brazil

² Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Industrial, Rua Marquês de São Vicente, 225, Rio de Janeiro, RJ, 22451-900, Brazil

³ NanoBusiness Informação e Inovação Ltda, Polo Tecnológico de Xerém, Prédio 32, Av. Nossa Senhora das Graças, 50, Xerém, Duque de Caxias, Rio de Janeiro 25250-020, Brazil

⁴ Universidade Federal Fluminense, Instituto de Física, Laboratório de difração de raios X(LDRX-UFF), Av. Gal. Milton Tavares de Souza, Campus Praia Vermelha, Niterói, RJ, 24210-346, Brazil

⁵ Universidade Federal do Rio de Janeiro, Escola de Belas Artes, Av. Pedro Calmon 500, Rio de Janeiro, RJ, 21941-901, Brazil

⁶ Centro Universitário Augusto Motta, Av. Paris 84, Rio de Janeiro, Rio de Janeiro 21041-020, Brazil

⁷ Universidade Federal do Rio de Janeiro, Instituto de Química, Av. Athos da Silveira Ramos, 149, Rio de Janeiro, RJ, 21941-611, Brazil

ARTICLE INFO

Article history:
 Received 22 June 2017
 Received in revised form 18 September 2017
 Accepted 13 October 2017
 Available online 12 October 2017

Keywords:
 polymorph
 crystal structure
 molecular modeling
 X-ray polycrystal diffraction and
 CPDMPAS NMR

ABSTRACT

A new monohydrated polymorph of dexamethasone acetate was crystallized and its crystal structure characterized. The different analytical techniques used for describing its structural and vibrational properties were: single crystal and polycrystal X-ray diffraction, solid state nuclear magnetic resonance, infrared spectroscopy. A Hirshfeld surface analysis was carried out through self-arrangement cemented by H-bonds observed in this new polymorph. This new polymorph form appeared because of self-arrangement via classical hydrogen bonds around the water molecule.

© 2018 American Pharmacists Association®. Published by Elsevier Inc. All rights reserved.

Introduction

Dexamethasone belongs to the compound class of steroids. The compounds of this class show anti-inflammatory activity and are currently the drugs of choice for the treatment of several diseases, despite their adverse side effects. In dermatology, the use of topical steroid preparations represents a major treatment option. For instance, dexamethasone shows a good pharmacological profile against dermatosis, which is a disease that affects humans worldwide. This compound is a synthetic adrenal corticosteroid with potent anti-inflammatory properties. It acts by binding to specific nuclear steroid receptors and also interferes with NF-κB activation

and apoptotic pathways. According to Weeks et al.,¹ its properties show a fine relation with conformation structure, especially with the A-ring conformation.

The polymorphism effect in the solid state occurs when the same molecule exists as multiple crystalline phases in the crystal lattice.^{2,3} For the pharmaceutical industry, this aspect is relevant and must be monitored since active pharmaceutical ingredients can suffer from this phenomenon.^{4,5} When the polymorph forms contain either stoichiometric or nonstoichiometric amounts of a solvent, these forms are also commonly known as hydrates⁶ and solvatomorphs.⁷

In the last years, the crystallographic aspects and their consequences around polymorphism effects have been the focus of several books.^{8–11} For the pharmaceutical industry, the main concern are the effects of the appearance of unwanted polymorphs within the active pharmaceutical ingredients, formulations, and tablets, a problem that can appear throughout the production chain and therefore must be very well monitored.¹² Since polymorph

This article contains supplementary material available from the authors by request or via the Internet at <https://doi.org/10.1016/j.xphs.2017.10.001>.

* Correspondence to: Lorenzo do Canto Visentin (Telephone: +55-21-2145-3218; Fax: +55-21-2679-9518).

E-mail address: lorenzo.visentin@nanobusiness.com.br (L.C. Visentin).

<https://doi.org/10.1016/j.xphs.2017.10.001>

0022-3549/© 2018 American Pharmacists Association®. Published by Elsevier Inc. All rights reserved.

forms of same molecule can show different physicochemical properties when compared with biologically active form of interest, the monitoring is critical.

Many works identifying solvates and polymorphs in steroid hormones date from the 1960s and 1970s.^{13,14} In 1975, the monohydrate crystal structure of the 9 α -Fluoro-16 α -methyl-II [1,17,21-trihydroxy-1,4-pregnadiene-3,20-dione-21-Acetate] was determined by Terzis and Theophanides.¹⁵ According to this study, this molecule is a potent inducer of liver microsomal enzymes as well as a powerful glucocorticoid. In 2001, Mallet et al.¹⁶ studied the desolvation process and the solid-state transitions of dexamethasone acetate solvated by DMSO and water. The academic interest in polymorphism has manifested itself in the increased number of publications, annual reviews articles, and patents appearing during 2004¹⁷ and 2005.¹⁸

To distinguish different polymorphs, one important aspect is the study of intermolecular interactions, which defines the solid packing behavior. Taking this into account, Hirshfeld surface (HS) analysis^{19–23} offers an interesting approach to investigate self-arrangements through hydrogen bonds and prediction of polymorphs. This situation represents a serious problem for the pharmaceutical industry and regulatory agencies in production and distribution of safe and efficient solid drugs.

Emerging as a short-range technique, solid-state nuclear magnetic resonance (NMR) provides information about a solid system and its electronic structure. Taking this into account, NMR spectroscopy has been used to achieve a detailed characterization of powder samples of drugs, for example, even polymorphs with minimal differences can be analyzed by NMR.^{24,25} In some cases, solid-state interactions (e.g., dipolar couplings) make the spectrum interpretation a challenge. Fortunately, these difficulties can often be circumvented by computational chemistry calculations. Indeed, a number of theoretical methods aim at calculating the magnetic response by using density functional theory (DFT) within the cluster approximation.^{26,27} The gauge including projector augmented waves²⁸ (GIPAW) approach is a DFT-based method used for determination of physical parameters measured by NMR and takes into account the natural solid translational invariance. GIPAW-computed chemical shifts have been widely used for investigating organic solids, especially in the NMR crystallography approach. Cross-polarization magic angle spinning carbon-13 (¹³C CPMAS) and GIPAW calculations were not previously reported in literature for both dexamethasone acetate (DEX-I and DEX-II) crystal structures.

In this work, a new monohydrated form of the dexamethasone acetate molecule (DEX-II) was obtained, which was crystallized and characterized by different analytical techniques. Structural properties were investigated by single crystal X-ray diffraction (SCXRD) and polycrystal X-ray diffraction (PXRD), solid-state nuclear magnetic resonance (solid-state NMR), and infrared spectroscopy (IR).

Material and Methods

Solvents

The solvents used in the crystallization tests were ethanol P.A. 99.5%, acetone P.A. 99.5%, and methanol P.A. 99.5% (Sigma-Aldrich, São Paulo, Brazil).

Crystallization Experiment

For DEX-I

In a beaker containing 30 mL of methanol was solubilized with 100 mg of dexamethasone acetate at room temperature. The mixture was left standing to slowly evaporate. After 1 week, the colorless single crystals suitable for X-ray analysis were obtained.

For DEX-II

In a beaker containing 15 mL of ethanol and 15 mL of acetone was solubilized with 100 mg of dexamethasone acetate at room temperature. The mixture was left standing to slowly evaporate. After 1 week, the colorless single crystals suitable for X-ray analysis were obtained.

Fourier Transform Infrared Spectroscopy

The infrared spectrum was collected in the Perkin Elmer Spectrum GX Fourier transform infrared (FTIR) spectrometer. Measurements were performed in the range of 4000 to 400 cm⁻¹, with resolution of 4 cm⁻¹, step 1 cm⁻¹ totalizing 16 scans.

Solid-State NMR Experiments

¹³C CPMAS-NMR experiments were performed on a Bruker Avance III NMR spectrometer (B₀ = 9.4T), operating at Larmor frequency of 100.3 MHz for ¹³C. The measurements were carried out by using a 4 mm CPMAS probehead and zirconia rotors stopped with Kel-F caps. To excite ¹³C, a cross polarization (CP) pulse sequence was used after several optimizations. Acquisitions were performed using CP.ramp.100 pulse sequence with 5 μ s proton 90 degree pulse. The contact and recycle times of 5 μ s, 2000 μ s, and 4 s were found, respectively. The Fourier transformation was carried out after 1024 accumulations with exponential multiplication of 10 Hz. Both samples (DEX-I and DEX-II) were spun at 11 kHz. The α -glycine carboxyl ($\delta_{c=O}$ = 176.02 ppm) resonance was used as external secondary signal reference for chemical shift calibration.

Magnetic Response Calculations

All electronic calculations were carried out using codes available in Quantum-Espresso package²⁹ which implements periodic boundary conditions (PBCs) within the DFT framework. The Kohn-Sham one-electron wavefunctions were expanded self-consistently using a plane-wave basis set adjusted to yield less than 1 mRy/atom convergence in total energy. Satisfactory results were found with a kinetic energy cutoff of 55 Ry for the wavefunction and 220 Ry for the charge density cutoff. The *k*-point sampling for the Brillouin zone integration was carried out by sums 3 \times 2 \times 1 within the Monkhorst-Pack method³⁰ *k*-mesh grid. The exchange and correlation contribution was calculated using the Perdew–Burke–Ernzerhof (PBE) approach³¹ for the generalized gradient approximation (GGA). The interaction between valence and core electrons was treated using the PAW approach, available in PSLibrary.³² The equilibrium geometries were obtained by an initial proton optimization. DEX-I crystal structure was available in literature,¹⁵ and DEX-II had not been reported until the present work.

Ab initio magnetic response calculations were performed within the GIPAW approach, which is implemented in the Quantum Espresso package.³² The GIPAW calculations yielded the absolute chemical shielding tensors (σ) which were converted into its isotropic values (σ_{iso}) by the equation $\sigma_{iso} = \text{Tr}[\sigma/3]$. The isotropic chemical shifts (δ_{iso}) were calculated by the equation $\delta_{GIPAW} = \delta_{ref} - \delta_{iso}$, as suggested by Harris, by using glycine as reference.³³

Polycrystal X-Ray Diffraction Experiment

PXRD patterns were obtained from a PANalytical X'Pert PRO diffractometer at 295(2) K with Cu K α radiation (λ = 1.5418 Å), in the Bragg-Brentano reflection geometry, coupled graphite monochromator and PIXcel 1D detector with length active of 3347° (2 θ). Samples were collected in the 2 θ range between 2 and 50° with steps of 0.026° and 100 s of step size, and using a scan speed (°/s) of 0.067335.

Table 1
Crystal Data and Structure Refinement for DEX-I and DEX-II

Name	DEX-I ¹⁵	DEX-II
Empirical formula	C ₂₄ H ₃₁ FO ₆ ·H ₂ O	C ₂₄ H ₃₁ FO ₆ ·H ₂ O
Formula weight	452.50	452.50
Temperature (K)	295	293(2)
Wavelength (Å)	1.54178	1.54178
Crystal system	Orthorhombic	Orthorhombic
Space group	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
Unit cell dimensions (Å)		
<i>a</i>	15.292 (3)	6.6009 (1)
<i>b</i>	19.143 (3)	15.7827 (3)
<i>c</i>	7.808 (2)	22.8607 (5)
Volume (Å ³)	2285.67	2381.63 (8)
Z	4	4
Density (calculated) (mg/m ³)	1.308	1.262
F(000)	968	968
Crystal size (mm)	0.10 × 0.12 × 0.25	0.386 × 0.288 × 0.064
Reflections collected	2141	74,223
Independent reflections	1949 (<i>R</i> _(int) = 0.142)	4838 (<i>R</i> _(int) = 0.1167)
Completeness to theta	98.0	99.6
Absorption correction	Semi-empirical	Semi-empirical
Max. and min. transmission	0.940 and 0.890	0.950 and 0.746
Refinement method	Full matrix least-squares on F ²	Full matrix least-squares on F ²
Data/restraints/parameters	1949/0/239	4838/0/297
Fit indices [<i>I</i> > 2σ(<i>I</i>)]	<i>R</i> ₁ = 0.0280; <i>wR</i> ₂ = 0.0290	<i>R</i> ₁ = 0.0564; <i>wR</i> ₂ = 0.1360
Residuals (all data)	<i>R</i> ₁ = 0.0320; <i>wR</i> ₂ = 0.0290	<i>R</i> ₁ = 0.0855; <i>wR</i> ₂ = 0.1540
Least-squares difference peak and hole (e.Å ⁻³)	0.210 and -0.230	0.504 and -0.158
CCDC number	1157946	1514791

Crystal X-Ray Diffraction Experiment

X-ray data for the compound (DEX-II) were collected from a D8 Venture diffractometer,³⁴ at 295(2) K, using Incoatec *I*_μS focus X-ray source and CuK α monochromatic radiation. The crystal was mounted in the *Kappa* Goniometer, linked with a θ/θ 100 detector for reflections collection. The average data and cell parameters refinement were obtained using the *SHELX*³⁴ programs. Data reduction and integration were processed using *SAINT v8.34A*³⁵ and absorption correction performed by *SADABS*^{35,36} programs. The structure was solved in *WinGX*³⁷ package. Direct Methods via *SHELXS*³⁸ and refined via *SHELXL*³⁸ by a full matrix least-squares treatment with anisotropic displacement parameters for all nonhydrogen atoms. H atoms of the unsaturated carbon were positioned geometrically (C–H = 0.93 Å for C_{sp}² atoms) and treated as riding on their respective C atoms, with *U*_{iso}(H) values set at 1.2*U*_{eq} C_{sp}². The hydrogen atoms of methyl and methylene groups were also positioned geometrically (C–H = 0.96 and 0.97 Å for C_{sp}³ atoms) and treated as riding on their respective C atoms, with *U*_{iso}(H) values set at -1.5*U*_{eq} C_{sp}³ and -1.2*U*_{eq} C_{sp}³. The hydrogen atom in the hydroxyl groups was positioned geometrically (O–H = 0.82 Å for O_{sp}³ atoms) and treated as riding on their respective O atoms, with *U*_{iso}(H) values set at -1.5*U*_{eq} O_{sp}³. The water hydrogen atoms were located in a Fourier map and free refined to position, the distance between O(7)–H(1W) atoms is 1.039(3) Å and for O(7)–H(2W) is 1.028(4) Å. To verify the number of water molecules present in the unit cell, the squeezed procedure was performed using *Platon* program.³⁹ The result shows a total potential solvent accessible void volume (SOLV-Map value) of 65 Å³ and electron count voids/cell equal to 10 electrons per unit cell, which is equivalent to only 1 water molecule present in the unit cell.³⁹ The crystal data are listed in Table 1 for DEX-I¹⁵ and DEX-II.

Results and Discussion

Spectroscopy Analysis

The infrared spectroscopic⁴⁰ data for DEX-I and DEX-II show slight differences in the profile of some spectrum absorption bands.

Considering both forms, the absorptions appear in the range between 3700 and 3150 cm⁻¹ for hydroxyl groups and 1780 and 1620 cm⁻¹ for carbonyl groups. The main absorption bands for the 2 polymorphs are listed in Table 2, and Figure 1 shows a comparative spectrum between DEX-I and DEX-II. The small differences observed in spectrum are in the range of 3150–3700 cm⁻¹, which could be due to different hydrogen-bonding with water molecule.

¹³C NMR Analysis

¹³C CPMAS-NMR spectra of dexamethasone acetate polymorphs (Fig. 2) depicts 24 and 22 NMR resonances for DEX-I and DEX-II, respectively. Taking only the ¹³C CPMAS-NMR analysis into account, for both crystal structures, the number of NMR signals suggests only 1 molecule in asymmetric unit, in good agreement with the present crystallographic data. In addition, the comparison between ¹³C magnetic responses clearly highlights the intrinsic differences of each crystal structure. The signal shapes and chemical shifts of DEX-I and DEX-II are almost the same at δ 0–33 and δ 37–200 ppm range. Out of 16 resonances, only 3 have $\Delta\delta > 1$ ppm (δ 93.4/91.9, 155.5/157.2 and δ 170.8/172.3 ppm). Indeed, the major differences between both ¹³C spectra were found within the δ 33–37 ppm (Fig. S1) range, where DEX-I exhibits 3 resonances (δ 33.0, δ 34.9 and δ 35.7 ppm), whereas DEX-II shows just 1 (δ 34.6 ppm).

In fact, the molecular structure of dexamethasone acetate has 24 magnetically different carbons, whereas DEX-II ¹³C CPMAS-NMR has only 22 signals. Taking into account the dexamethasone acetate ¹³C{¹H} data, published by Hickey et al.⁴¹ (Table 3), the

Table 2

The Many Bands Symmetrical and Anti-symmetrical Stretching and Deformations of the Groups (O–H), (C=O) and (C=C) for DEX-I and DEX-II

Compound	OH	C=O	C=C
DEX-I	3583, 3457, 3344, 3235	1740, 1720, 1665	1618, 1600
DEX-II	3621, 3543, 3490, 3388	1750, 1724, 1663	1614, 1605

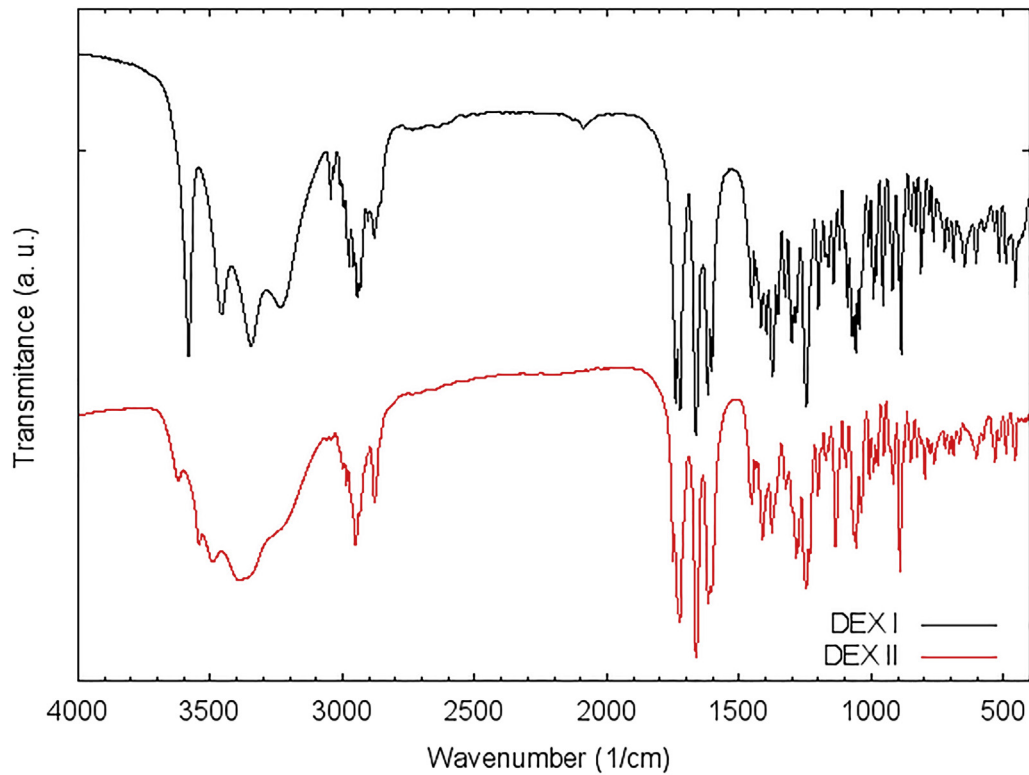


Figure 1. Comparison of the 2D FTIR spectra of DEX-I (black) and DEX-II (red) polymorphs.

als at $\delta 100.5/102.3$ ppm (DEX-I) and $\delta 100.3/102.1$ ppm (DEX-II) spectra could be assigned as C24, due to the ^{19}F - ^{13}C coupling. A direct comparison of the ^{13}C signals for the solid state and the previously published liquid state (dioxane- d_8)⁴¹ does not allow an unambiguous assignment of both DEX-I and DEX-II ^{13}C CPMAS-R spectra. All experimental chemical shifts are summarized in Table 3.

Following a previous paper⁴² the equilibrium geometries were found by a simple proton relaxation. In addition, DEX-II and DEX-I has 260 atoms ($4 \times \text{C}_{24}\text{H}_{31}\text{FO}_6 \cdot \text{H}_2\text{O}$) within its unit cell and the relaxation of the entire cell, considering the degrees of freedom of all heavy atoms, would be too computationally intensive.

^{13}C GIPAW shielding (σ) of each polymorph was plotted versus its experimental chemical shift counterpart (Fig. 3) to achieve the

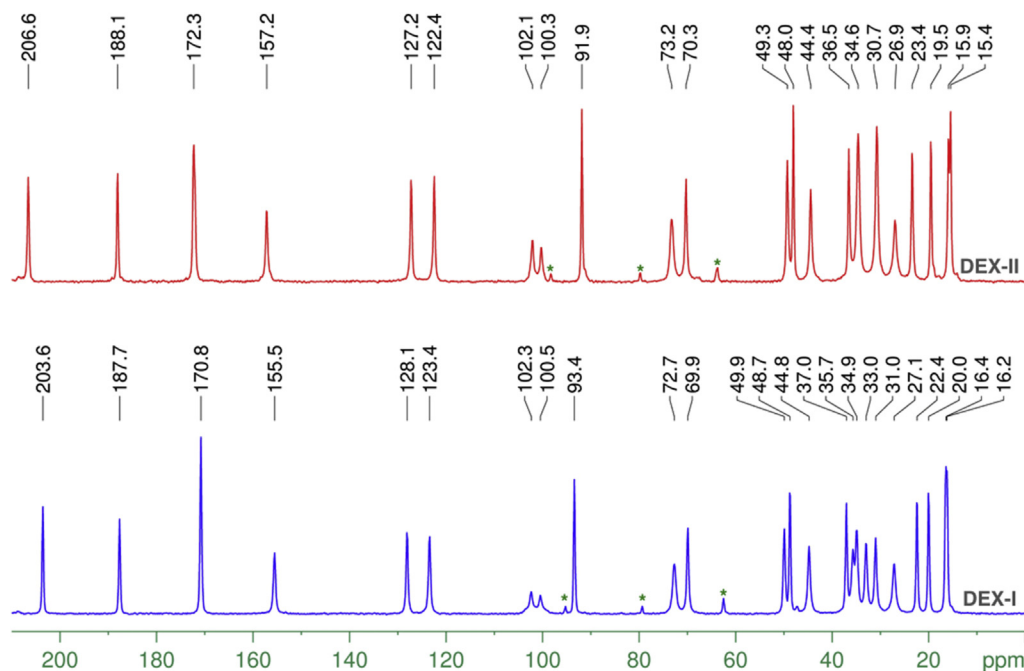


Figure 2. ^{13}C CPMAS spectra for dexamethasone acetate ($\text{C}_{24}\text{H}_{31}\text{FO}_6$) polymorphs (DEX-I and DEX-II). The spinning side bands are indicated with asterisks.

Table 3
Experimental ($^{13}\text{C}\{^1\text{H}\}$ and ^{13}C CPMAS) and Theoretical (GIPAW) ^{13}C Magnetic Response of and Dexamethasone Acetate (DEX) and Dexamethasone Acetate Polymorphs, DEX-I, and DEX-II

Hickey ⁴¹ $^{13}\text{C}\{^1\text{H}\}^a$ - (ppm)		This Work ^{13}C CPMAS/GIPAW - (ppm)						
δ -DEX		δ -DEX-I				δ -DEX-II		
#	Experimental	#	Theoretical	Experimental	$\Delta\delta$	Theoretical	Experimental	$\Delta\delta$
C15	14.8	C10	13.4	16.2	2.8	14.1	15.4	1.3
C10	16.7	C17	16.6	16.4	-0.2	14.3	15.9	1.6
C17	20.1	C15	16.7	20	3.3	17.5	19.5	2
C11	23.4	C11	20.6	22.4	1.8	24.4	23.4	-1
C22	28	C22	25.2	27.1	1.9	29.9	26.9	-3
C21	31.1	C7	31.7	31	-0.7	30.5	30.7	0.2
C7	32.6	C23	36.0	33	-3	35.4	34.6	-
C23	34.8 (18 Hz)^b	C8	36.1	34.9	-1.2	38.4	-	-
C4	36.6	C4	38.2	35.7	-2.5	33.7	36.5	3.7
C8	36.9	C21	38.3	37	-1.3	32.8	36.5	3.7
C6	44.5	C6	50.8	44.8	-6	46.2	44.4	-1.8
C2	48.5 (19.4 Hz) ^b	C2	52.2	48.7	-3.5	53.4	48	-5.4
C5	48.9	C5	53.5	49.9	-3.6	50.2	49.3	-0.9
C14	n.d.	C14	69.3	69.9	0.6	72.2	70.3	-1.9
C3	72.3 (37.5 Hz) ^b	C3	75.3	72.7	2.6	79.1	73.2	-5.9
C9	91.7	C9	99.2	93.4	-5.8	98.1	91.9	-6.2
C24	101.1 (176.2 Hz) ^b	C24	110.0	100.5	-8.6	111.2	100.3	8.8
C	125.4	C19	122.9	123.4	0.5	123.4	122.4	1
C	130.2	C12	128.3	128.1	-0.2	128.7	127.2	-1.5
C	152	C1	163.2	155.5	-7.7	167.2	157.2	-10
C	165.9	C16	167.3	-	-	169.7	-	-
C	170.6	C20	176.2	170.8	-5.4	185.0	172.3	-12.7
C	185.6	C18	189.1	187.7	-1.4	182.1	188.1	6
C	204.6 (2.8 Hz) ^b	C13	214.2	203.6	-10.6	216.8	206.6	-10.2

^a as highlight the assignment conflicts.

^b analysis carried out in dioxane-d₈ (δ 66.66 ppm).

^c values given in parentheses are ^{13}C - ^{19}F coupling constants.

linear equation in which the y-intercept is itself the reference shielding (σ_{ref}). The reference shielding obtained for each polymorph was used to calculate the GIPAW chemical shifts and were directly compared with experimental magnetic response.

It can be clearly seen from our *ab initio* calculations that the GIPAW approach exhibits an excellent experimental-theoretical correlation with $r^2 > 0.99$ and a slope close to unit, |1.0282| for DEX-I and |1.0336| for DEX-II. In addition, this good agreement is a strong indication of a reliable crystal data. Considering that both asymmetric unities (DEX-I/DEX-II) contain only 2 molecules (1 dexamethasone acetate + 1 water), the straightforward comparison between $^{13}\text{C}\{^1\text{H}\}$, ^{13}C CPMAS-NMR, and GIPAW-calculated chemical shifts emerges as an affordable strategy to achieve the ^{13}C CPMAS resonances assignment. GIPAW approach

yields 24 different theoretical shieldings (σ) which were converted into chemical shifts by using the y-intercept (σ_{ref}) of each calculation. Table 3 lists the average chemical shift and theoretical shielding.

Hickey et al.⁴¹ reported 23 (+1 that was hidden by solvent peak; dioxane-d₈ at δ 66.66 ppm) resonances in their solution-state NMR study for DEX. Taking only the sequence of chemical shifts into account, DEX-I and DEX-II ^{13}C CPMAS clearly exhibit almost the same values, except for δ 165.9 ppm that cannot be observed in both solid-state NMR spectra.

Although our *ab initio* results shown a good theoretical-experimental correlation, 9 out of 24 carbons (C7, C8, C10, C15, C16, C17, C20, C21, C23) exhibit assignments that conflict compared with the Hickey et al.⁴¹ proposal. *Per se* these discrepancies may not

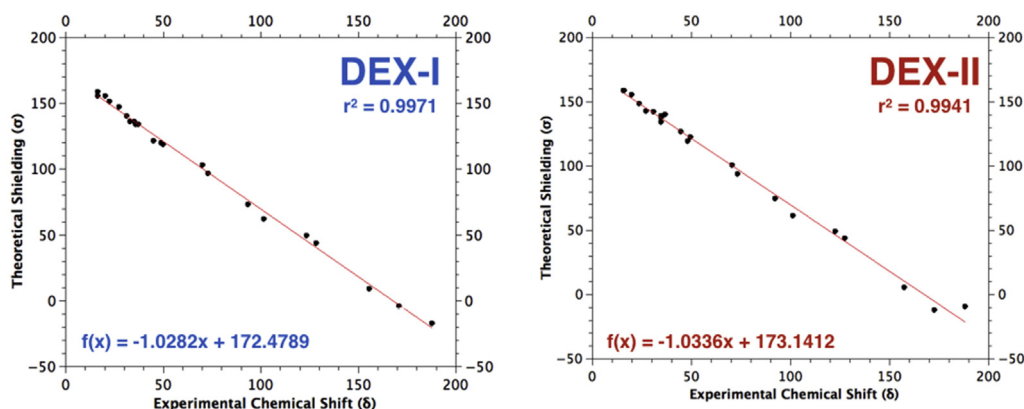


Figure 3. Theoretical ^{13}C shielding versus experimental chemical shift of dexamethasone acetate monohydrate polymorphs, DEX-I (left) and DEX-II (right).

mean an error, but some differences between solid and solution-state NMR. Some interesting example about the influence of packing over chemical shifts relies on C16/C20 resonances and the region at δ 30–40 ppm range. Both ^{13}C CPMAS (DEX-I and DEX-II) depict only 1 resonance at δ 165–175 ppm range, for example, the signals C16/C20 have only 1 chemical shift value for both carbons, which explains the value of $\Delta\delta$. In the same way, DEX-II ^{13}C CPMAS spectra at δ 30–40 ppm range show a single signal, whereas theoretical calculation predicts 3 resonances. In fact, the description of ^{13}C magnetic response of fluorinated compounds with periodic DFT calculations using GGA-PBE functionals is not optimal,⁴² which explains the $\Delta\delta > 8.5$ ppm for C24.

Crystal Structure

The crystal structure characterization of DEX-II reveals different structural properties, in relation of known structure DEX-I.

The structural formula and scheme of the DEX-II with the numbering scheme, is shown in Figure 4. Selected bond lengths and bond angles are listed in Table 4.

This class of compound shows a basic chemical structure of 4 rings of carbon atoms forming the steroid molecular skeleton⁴⁴ in an arrangement corresponding to that of perhydrocyclopentanophenanthrene nucleus.

The angles and bond lengths for DEX-II agrees with the range of values and bond lengths reported for similar structures^{15,16,45-48} and unsaturated and saturated carbon bonds. The calculated bond lengths reported in the literature for $\text{C}_{sp}^2 - \text{C}_{sp}^2$, $\text{C}_{sp}^2 - \text{C}_{sp}^3$ and $\text{C}_{sp}^3 - \text{C}_{sp}^3$ in the context of the basic steroid molecular skeleton show the following average values respectively (Å): 1.317–1.358, 1.493–1.522, and 1.517–1.548. The International Tables for Crystallography⁴⁹ list these parameters for typical bond lengths, which are 1.326 Å for $\text{C}_{sp}^2 - \text{C}_{sp}^2$ ($\sigma = 0.012$, $q_l = 1.318$ Å $q_u = 1.334$ Å), 1.507 Å for $\text{C}_{sp}^2 - \text{C}_{sp}^3$ ($\sigma = 0.015$, $q_l = 1.499$ Å and $q_u = 1.517$ Å), and 1.530 Å for $\text{C}_{sp}^3 - \text{C}_{sp}^3$ ($\sigma = 0.015$, $q_l = 1.521$ Å and $q_u = 1.539$ Å).

In DEX-I, the molecular structure shows bond average lengths in the order of (in Å) 1.325–1.327, 1.501–1.502, and 1.528–1.536 for

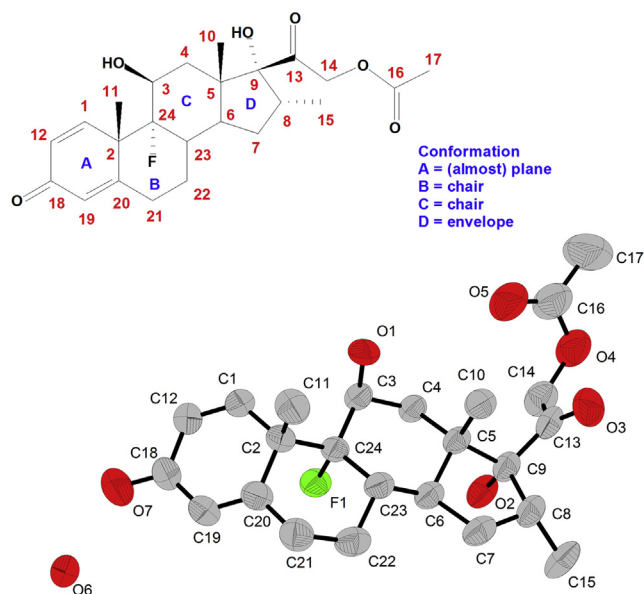


Figure 4. Molecular structure with numbering scheme and conformation label for DEX-II and view of the ORTEP⁴³ plot with displacement thermal ellipsoids plot at 50% probability level. The hydrogen atoms were omitted for better visualization.

Table 4
Bond Lengths [Å] and Angles [°] for DEX-I and DEX-II

DEX-I ¹⁵		DEX-II	
Bond Lengths [Å]			
O(7)–C(18)	1.2450(2)	O(7)–C(18)	1.2315(8)
O(1)–C(3)	1.4211(3)	O(1)–C(3)	1.4172(8)
O(17)–C(17)	1.4368(2)	O(17)–C(17)	1.4258(1)
O(3)–C(13)	1.2072(1)	O(3)–C(13)	1.2073(9)
O(4)–C(14)	1.4338(2)	O(4)–C(14)	1.4393(2)
O(6)–H(4)	0.834(2)	O(6)–H(4)	1.039(3)
O(6)–H(5)	0.975(1)	O(6)–H(5)	1.028(3)
F(1)–C(24)	1.4302(2)	F(1)–C(24)	1.4187(6)
Angles [°]			
O(5)–C(16)–O(4)	123.1(1)	O(5)–C(16)–O(4)	121.5(6)
O(5)–C(16)–C(17)	125.7(1)	O(5)–C(16)–C(17)	126.5(6)
O(4)–C(16)–C(17)	111.2(1)	O(4)–C(16)–C(17)	111.9(6)
C(16)–O(4)–C(14)	116.4 (3)	C(16)–O(4)–C(14)	115.6(4)
H(4)–O(6)–H(5)	106.4 (2)	H(4)–O(6)–H(5)	121.1(3)
O(7)–C(18)–C(19)	121.6(1)	O(7)–C(18)–C(19)	122.3(4)
O(7)–C(18)–C(12)	121.1(1)	O(7)–C(18)–C(12)	120.1(5)
O(2)–C(9)–C(13)	109.0(2)	O(2)–C(9)–C(13)	108.2(3)
O(2)–C(9)–C(8)	107.9(2)	O(2)–C(9)–C(8)	113.2(3)
O(2)–C(9)–C(5)	111.3(3)	O(2)–C(9)–C(5)	106.8(3)
H(4)–O(6)–H(5)	106.4 (2)	H(4)–O(6)–H(5)	121.1(3)
O(7)–C(18)–C(19)	121.6(1)	O(7)–C(18)–C(19)	122.3(4)
O(3)–C(13)–C(14)	120.2(4)	O(3)–C(13)–C(14)	121.2(5)
O(3)–C(13)–C(9)	122.8(3)	O(3)–C(13)–C(9)	123.2(4)
O(4)–C(14)–C(13)	111.6(2)	O(4)–C(14)–C(13)	112.3(4)

$\text{C}_{sp}^2 - \text{C}_{sp}^2$, $\text{C}_{sp}^2 - \text{C}_{sp}^3$ and $\text{C}_{sp}^3 - \text{C}_{sp}^3$, respectively.¹⁵ For DEX-II, these bond lengths are (in Å) 1.321–1.339, 1.499–1.501, and 1.523–1.534, respectively.

In DEX-II, the other most important bond lengths are C18–O7_{carbonyl} [1.232(5) Å], C16–O5_{carbonyl} [1.203(7) Å], C3–O1_{hydroxyl} [1.417(5) Å], and C9–O2_{hydroxyl} [1.425(5) Å]. Their values are slightly shorter than in DEX-I.¹⁵ This is due to the attraction and repulsion forces from self-arrangement generated and cemented by the classical hydrogen bonds around the water molecule. In DEX-I, 3 H-bonds around water molecule are identified, whereas in DEX-II 4. This causes a change in the torsion angles in DEX-II relative to DEX-I. The most important torsion angles affected in DEX-II were C5–C4–C3–O1, O2–C9–C13–O3 and O2–C9–C13–C14 and C14–O4–C16–O5, which are (in °) 75.7(4), –156.9(4), 23.4(5),

Table 5
Geometric Parameters for Hydrogen Bonding in the (DEX-I and DEX-II), (Å/°) From Intermolecular Interactions

D–H...A	D–H	H...A	D...A	∠D–H...A
For DEX-I ¹⁵				
Classical H-Bonds from water				
O(1)–H(2)···O(6)	0.88	1.96	2.8417(7)	173
O(6)–H(5)···O(5)	0.97	1.98	2.9463(8)	168
O(6)–H(4)···O(2) [#]	0.83	1.99	2.8190(7)	171
Classical H-Bonds out water				
O(2) [#] –H(3) [#] ···O(7)	0.88	1.91	2.7825(7)	172
Non-Classical H-Bond out water				
C(7)–H(11)···(O7) [#]	1.05	2.59	3.6258(9)	167
For DEX-II				
Classical H-Bonds from water				
O(1) ⁱ –H(2) ⁱ ···O(6)	0.82	2.05	2.833(4)	160
O(6)–H(5) ⁱ ···O(7)	1.04	1.71	2.726(5)	166
O(6)–H(4) ⁱ ···O(5) ⁱ	1.03	1.89	2.767(5)	142
O(2) ⁱⁱ –H(3) ⁱⁱ ···O(6)	0.82	2.01	2.795(4)	161
Non-Classical H-Bond out water				
C(14) ⁱⁱ –H(19) ⁱⁱ ···(O1)	0.97	2.58	3.443(5)	148

Symmetry code (i) = $\frac{1}{2} - x, 1 - y, \frac{1}{2} + z$; (ii) = $-\frac{1}{2} + x, 1/5 - y, 2 - z$; (#) = $\frac{1}{2} + x, \frac{1}{2} - y, -z$.

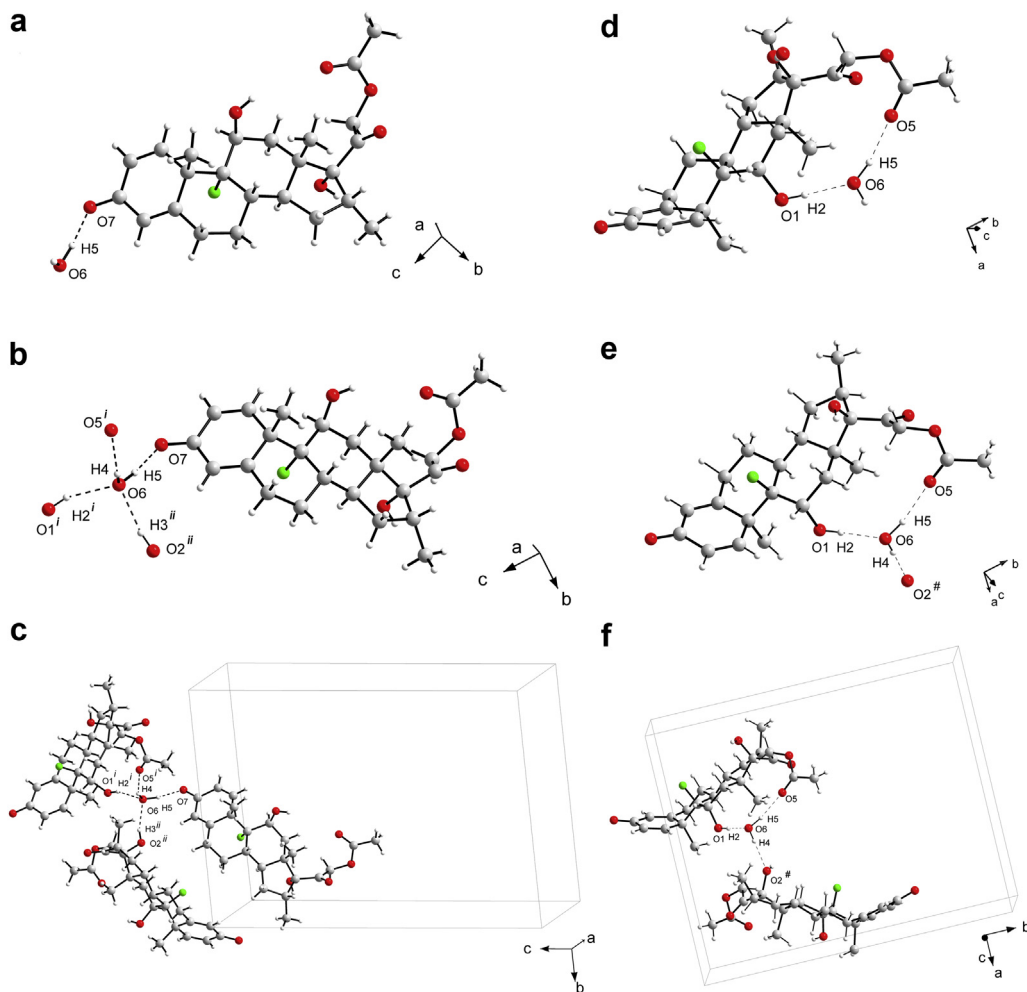


Fig 5. View of the crystal structures and their different packing modes and self-arrangement by classical H-bonds. In this way we have: (a) ORTEP⁴³ plot for asymmetric unit of DEX-I; (b) ORTEP⁴³ plot for tetragonal arrangement around water molecule in DEX-II; (c) view of tecton plot for DEX-II; (d) ORTEP⁴³ plot for asymmetric unit of DEX-II; (e) ORTEP⁴³ plot for trigonal arrangement around water molecule in DEX-I; (f) view of tecton plot for DEX-I; (symmetry code [i] = $\frac{1}{2} - x, 1 - y, \frac{1}{2} + z$; [ii] = $-1/2 + x, 1/5 - y, 2 - z$; [#] = $\frac{1}{2} + x, \frac{1}{2} - y, z$).

and $-1.6(8)$. The respective values for DEX-I are $77.7(2)$, $-148.3(2)$, $30.4(2)$, and $-1.7(2)$.

Hydrogen Bonds and Self-Arrangement

For DEX-II, the self-arrangement analysis shows that each molecule of water is surrounded by 3 dexamethasone acetate molecules. This assembly shows 4 interactions through H-bonds between 3-dexamethasone acetate molecule and 1 water solvate. It has been observed that the self-arrangement of DEX-I and DEX-II molecules formed tecton units.^{50,51}

The main geometric parameters from classical and nonclassical H-bonds are described in Table 5. The Figure 5 shows the asymmetric units of DEX-II evidencing 1 H-bond and for DEX-I 2 H-bonds intermolecular interactions. Besides showing tetragonal and trigonal self-arrangement around the water molecule by classical H-bonds for DEX-II and DEX-I, respectively. In addition, it also shows the respective tecton units for DEX-II and DEX-I.

In both DEX-II and DEX-I, the tecton units in the solid packaging form a 3D-framework. The graph-set analysis implemented in the Mercury⁵² program shows 2 motifs formed by DEX-II self-arrangement, which are $R_2^2(13)$ involving only classical H-bonds and $R_2^2(9)$ involving both classical and nonclassical H-bonds. All H-bonds were calculated using the Platon program and agree with

related literature.^{15,39,53} The 4 H-bonds show translation units along the *a*-axis, *b*-axis, and *c*-axis. Infinite 3D-framework were formed with directional [010], [001], and [100] vectors base. Of the 4 interactions surrounding the water solvate, 2 are bifurcated H-bonds and 2 are linear H-bond, which are $O(1)^i-H(2)^{ii}\cdots O(6)$, $O(6)-H(5)\cdots O(7)$, $O(6)-H(4)\cdots O(5)^i$ and $O(2)^{ii}-H(3)^{iii}\cdots O(6)$,

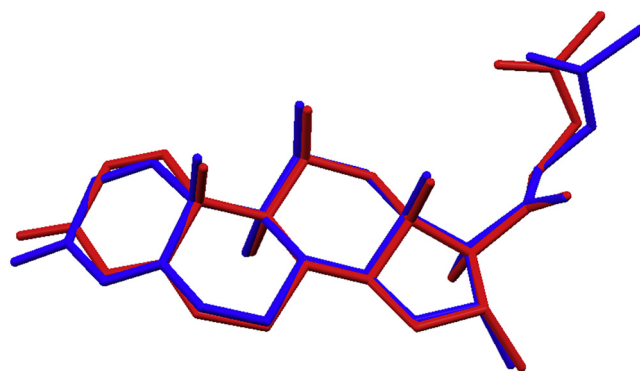


Figure 6. View of the superposition of DEX-I (blue) and DEX-II (red) after least-squares fitting.

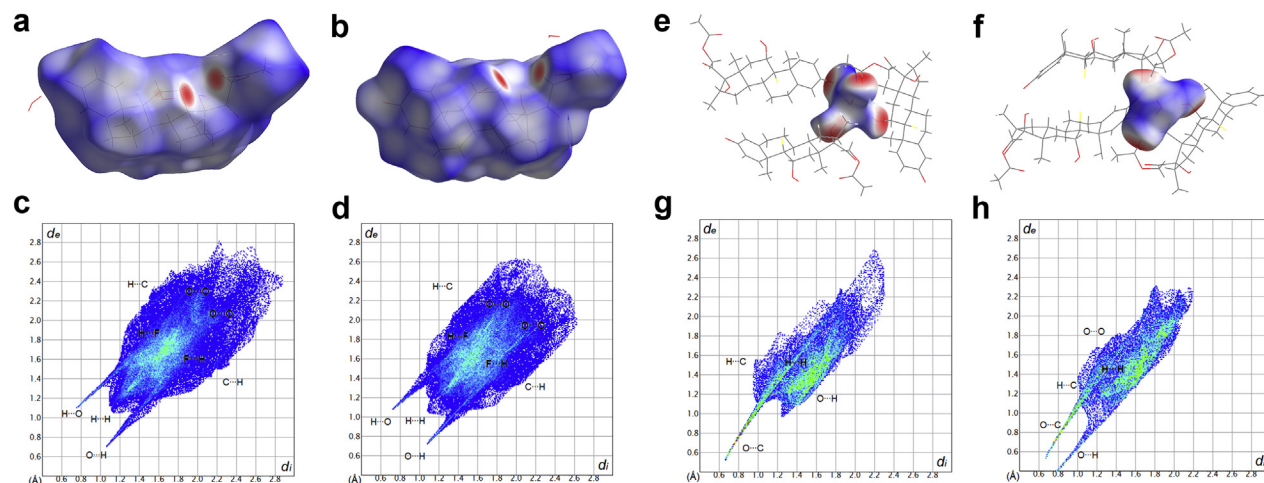


Figure 7. View HS surface and FPP in both crystal structure of DEX-II and DEX-I, in: (a) HS for respective assignment for the dexamethasone acetate monohydrate in DEX II; (b) HS for respective assignment for the dexamethasone acetate monohydrate in DEX I; (c) FPP for respective assignment for the dexamethasone acetate monohydrate in DEX II; (d) FPP for respective assignment for the dexamethasone acetate monohydrate in DEX I; (e) HS for respective assignment for the water in DEX II; (f) HS for respective assignment for the water in DEX I; (g) FPP for respective assignment for the water in DEX II; and (h) FPP for respective assignment for the water in DEX I.

respectively. The H-bond analysis around water molecule in DEX-I and DEX-II shows different interaction modes. DEX-I packing has 1 and less compared with DEX-II crystal structure. The water molecule forms 3 interactions in DEX-I, whereas in DEX-II 4.

Conformational Analysis

The 6-membered rings (rings A, B, and C) in the DEX-II molecule exhibited almost plane conformations for ring A, and chair conformations for both rings B and C, whereas the 5-membered ring (ring D) showed an envelope conformation. The 2 dexamethasone acetate monohydrate conformers were superimposed to

compare their conformational differences. The fitting was performed by selecting all atoms. The analysis was carried out using the Mercury⁵² program. No significant differences were observed in both conformations from DEX-I and DEX-II, except for a slight difference in the acetate group and carbonyl terminal group (Fig. 6). These groups are involved in the intermolecular interactions.

Hirshfeld Surface Analysis of Dexamethasone Acetate Monohydrate

HS¹⁹⁻²³ and Fingerprint plot (FPP)⁵⁴ analysis of the supramolecular environment are of fundamental importance in their unequivocal description of molecular packing, and of great

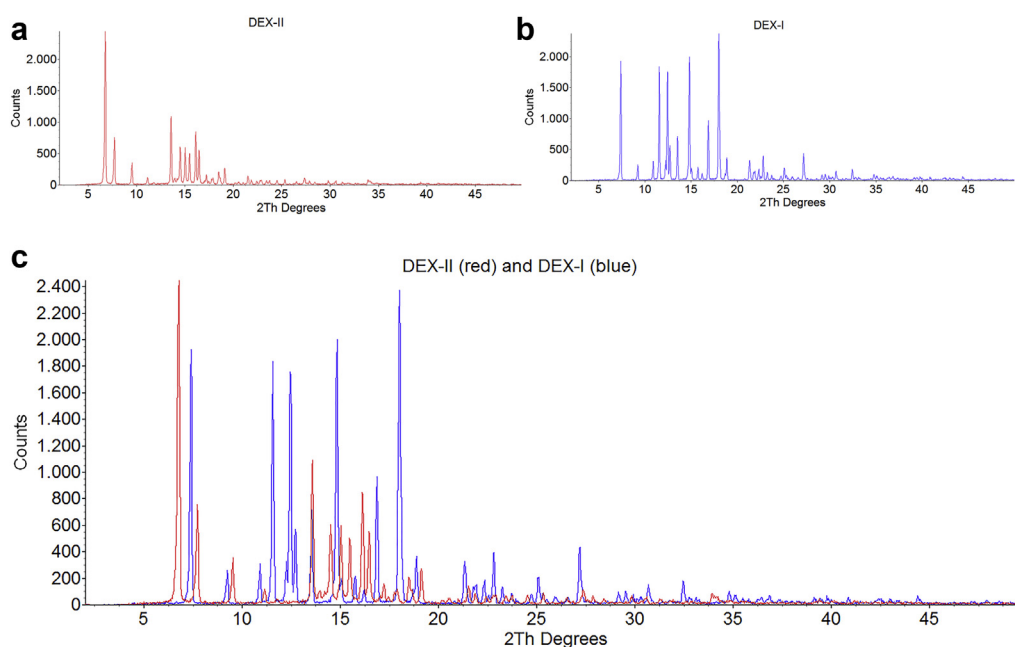


Figure 8. View of the profile standard polycrystal diffraction and overlap between DEX-I (blue) and DEX-II (red) forms: (a) pure XRD pattern DEX-II; (b) pure XRD pattern DEX-I; and (c) overlap XRD patterns.

Table 6
Indexed Polycrystal Diffraction With Position [$^{\circ}$ 2Th.], Interplanar Distance [Å], and Relative Intensity [%] From Standard Polycrystal X-Ray to DEX-I and DEX-II

DEX-I			DEX-II		
Pos. [$^{\circ}$ 2Th.]	d-spacing [Å]	Rel. Int. [%] ^a	Pos. [$^{\circ}$ 2Th.]	d-spacing [Å]	Rel. Int. [%] ^a
7.3679(6)	11.98862	78.38	6.7381(6)	13.10758	100
11.5382(6)	7.66317	76.23	7.689(1)	11.48816	28.49
12.4336(7)	7.11322	80.47	9.497(2)	9.30472	13.22
12.696(1)	6.96661	25.11	13.5556(8)	6.52689	46.91
13.517(1)	6.5453	30.24	14.489(2)	6.10825	24.15
14.7951(6)	5.98276	91.54	15.006(1)	5.8992	23.94
16.8411(8)	5.26024	44.24	15.461(1)	5.72652	21.92
17.983(1)	4.92878	100	16.104(1)	5.49919	36.89
22.780(1)	3.90058	19.31	16.445(1)	5.38615	23.31
27.168(1)	3.27969	20.47	19.107(2)	4.64116	11.6

$$^a I(2\theta) = I_{hkl} [\eta L(2\theta - 2\theta_0) + (1 - \eta) G(2\theta - 2\theta_0)].$$

importance in polymorphs studies. Crystalexplorer²² was employed for the construction of HS and FPP as well.

Figure 7 shows the respective HS and FPP for the 2 polymorphs. The HS was constructed in an isosurface at d_{norm} where $W(r) = 0.5$. The HS analyses for DEX-II/DEX-I, (Figs. 7a and 7b), show a d_{norm} over color scale observed of the $-0.658/-0.623$ (red) to $1.857/1.460$ Å (σ), and volume of $566.12/573.39$ Å³, respectively. The same approach was applied for HS analyses around the water solvate unit for DEX-II/DEX-I, which is shown in Figures 7e and 7f, with respect to d_{norm} over color scale observed of $-1.253/-1.272$ (red) to $4/1.432$ Å (blue), and volume of $76.95/64.46$ Å³. The corresponding FPP within respective assignments are presented in Figures 7c, 7d, 7g, and 7h for DEX-II/DEX-I.

As already discussed, the investigation of polymorphic structure can be done by the construction of HS and the comparison between respective FPP of the selected molecules. In this research, it was observed that the results from HS analysis and the FPP from the molecular structure alone were not conclusive as can be seen in Figures 7a–7d. The inconclusiveness of the HS and FPP data is a consequence of the high degree of similarity between the conformations of DEX-II and DEX-I.

For this reason, the identification of DEX-II and DEX-I forms by HS and FPP is only possible when the analysis is extended for supramolecular arrangement, through its respective tecton units. This analysis is shown in Figures 7e–7h and indicates the possibility to clearly identify the differences in HS and FPP.

The analysis of HS combined with FPP suggests that the strongest interactions correspond to classical H-bonds, involving O–H...O. Considering the HS analysis, it was possible to conclude that these interactions occur between dexamethasone acetate and water molecules.

As can be seen in Figures 7g and 7h, DEX-II, and DEX-I show similar O...C interactions in FPP. However, when we analyze the O...H interactions in both supramolecular arrangements, it is possible to see a clear difference between them. In DEX-II, an additional O...H bond was observed, which is not present in DEX-I,

exhibiting a different shape in the FPP. These differences in the supramolecular arrangements are conclusive and allow the distinguishing of DEX-II and DEX-I.

Qualitative Analysis From XRD Data

The data acquisition for the standard polycrystal diffraction for DEX-II and DEX-I were performed with high resolution equipment, for verifying whether the samples were really pure, Figure 8. The standard diffraction patterns of DEX-II and DEX-I are quite distinct, allowing a well-defined qualitative characterization. The main diffraction peaks do not overlap in both polycrystal diffraction patterns. For indexing peaks, a pseudo-Voigt peak profile function was used, and the main peaks for DEX-I and DEX-II are $2\theta^{\circ} = 7.37$; 11.54; 12.43; 14.8; 17.98 and 6.74; 7.69; 13.56; 14.49; 16.10, respectively. The others peaks are shown in Tables 6 and 7 is the show of the difference between both primitive cell.

Conclusion

Dexamethasone acetate monohydrate can form polymorphs in different crystallization solutions. This highlights the importance of monitoring, knowing, and controlling the polymorphism in this substance. We have obtained and isolated a new single crystal of dexamethasone acetate as a solvated form. The crystallization experiment in ethanol and acetone mixture, resulted in DEX-II, a new polymorph that is, reported and characterized in the present work.

Different analytical techniques were applied in the study of the structural properties of these polymorphs: SCXRD and PXRD X-ray diffraction, solid-state NMR, and IR.

The study has shown that the new polymorph appeared because of different self-arrangement around a water solvate. The water molecule in DEX-II shows a tetrahedral arrangement and the one in DEX-I show a trigonal arrangement. Only 1 additional H-bond is sufficient to induce a crystallization of the new polymorph with a distinct and unique crystal lattice. The unambiguous conclusion was derived from HS analysis, especially from FPP of water molecules in crystal lattice of the DEX-I and DEX-II polymorphs.

The qualitative study demonstrated different polycrystal diffraction patterns for DEX-II and DEX-I, which have been clearly identified.

Acknowledgments

The authors are grateful to their sponsors, *CNPq*, *FAPERJ*, and to *Laboratório de Difração de Raios X* from Universidade Federal Fluminense, (LDRX-UFF), Brazil. A special acknowledgment to Daniel Giacometti for management support in this project.

Table 7
The Origin *hkl* Plains for Cell Lattice in DEX-I and DEX-II

h	k	l	m	d	th2	Cell lattice
DEX-I						
2	0	0	2	7.63998	11.57328	$a = 15.27996$
0	2	0	2	9.55109	9.25186	$b = 9.10218$
0	0	2	2	3.89854	22.79162	$c = 7.797080$
DEX-II						
2	0	0	2	3.30035	26.99439	$a = 6.600710$
0	2	0	2	7.91934	11.16370	$b = 15.83868$
0	0	2	2	11.41858	7.73620	$c = 22.83716$

References







- Weeks CM, Duax WL, Wolff ME. Comparison of the molecular structures of six corticosteroids. *J Am Chem Soc.* 1973;95:2865–2868.
- Raw AS, Furness MS, Gill DS, Adams RC, Holcombe Jr FO, Yu LX. Regulatory considerations of pharmaceutical solid polymorphism in abbreviated new drug applications (ANDAs). *Adv Drug Deliv Rev.* 2004;56:397–414.
- Vippagunta SR, Brittain HG, David JW, Grant DJW. Crystalline solids. *Adv Drug Deliv Rev.* 2001;48:3–26.
- Brittain HG. Polymorphism and solvatomorphism 2010. *J Pharm Sci.* 2012;101:464–484.
- Linden A. Virtual issue on polymorphism. *Acta Crystallogr C.* 2011;67(Pt 12):e15.
- U.S. Food and Drug Administration. Scientific considerations of polymorphism in pharmaceutical solids: abbreviated new drug applications. Silver Spring, MD: FDA; 2002. Available at: https://www.fda.gov/ohrms/dockets/ac/02/briefing/3900B1_04_Polymorphism.htm. Accessed June 1, 2017.
- Yang D, Gong N, Zhang L, Lu Y. Isostructurality among 5 solvatomorphs of betulin: X-Ray structure and characterization. *J Pharm Sci.* 2016;105(6):1867–1873.
- Byrn SR, Pfeiffer RR, Stowell JG. *Solid State Chemistry of Drugs*. 2nd ed. West Lafayette: SSCI Inc; 1999.
- Brittain HG. Methods for the characterization of polymorphs and solvates. In: Brittain HG, ed. *Polymorphism in Pharmaceutical Solids*. New York: Marcel Dekker, Inc; 1999:227–278.
- Bernstein J. *Polymorphism in Molecular Crystals*. Oxford: Clarendon Press; 2002.
- Hilfiker R. *Polymorphism in the Pharmaceutical Industry*. Weinheim, Germany: Wiley-VCH; 2006.
- Antos OMM, Reis MED, Jacon JT, Lino MES, Simões JS, Doriguetto AC. Polymorphism: an evaluation of the potential risk to the quality of drug products from the Farmácia Popular Rede Própria. *Braz J Pharm Sci.* 2014;50(1):1–24.
- Shiles JA. Optical crystallographic properties of some glucocorticoids. *J Pharm Sci.* 1961;50:464–468.
- Kuhnert-Brandstätter M, Gasser P. Solvates and polymorphic modifications of steroid hormones. *Microchem J.* 1971;16(4):590–601.
- Ferzic A, Theophanides T. The crystal and molecular structure of 9 α -fluoro-16 α -methyl-11 β ,17,21-tri-hydroxy-1,4-pregnadiene-3,20-dione 21-acetate monohydrate. *Acta Cryst.* 1975;B31:796–801.
- Vallet F, Petit S, Petit MN, et al. Solvent exchange between dimethylsulfoxide and water in the dexamethasone acetate structure. *J Phys IV.* 2001;11:253–259.
- Brittain HG. Polymorphism and solvatomorphism. 2004. In: Brittain HG, ed. *Profiles of Drug Substances, Excipients, and Related Methodology*. Vol. 32. Amsterdam: Elsevier Academic Press; 2005:263–283.
- Brittain HG. Polymorphism and solvatomorphism 2005. *J Pharm Sci.* 2007;96:705–728.
- McKinnon JJ, Mitchell AS, Spackman MA. Hirshfeld surfaces: a new tool for visualising and exploring molecular. *Chem Eur J.* 1998;4(11):2136–2141.
- ayatilaka D, Grimwood DJ. Tonto: a fortran based object-oriented system for quantum chemistry and crystallography. In: Sloop PMA, Abramson D, Zogdanov AV, Gorbachev YE, Dongarra JJ, Zomaya AY, eds. *Computational Science—ICCS 2003. ICCS 2003. Lecture Notes in Computer Science*. Vol. 2660. Berlin, Heidelberg: Springer; 2003.
- Spackman MA, McKinnon JJ, Jayatilaka D. Electrostatic potentials mapped on Hirshfeld surfaces provide direct insight into intermolecular interactions in crystals. *CrystEngComm.* 2008;10(4):377–388.
- Spackman MA, Jayatilaka D. Hirshfeld surface analysis. *CrystEngComm.* 2009;11(1):19–32.
- McKinnon JJ, Fabbiani FPA, Spackman MA. Comparison of polymorphic molecular crystal structures through Hirshfeld surface analysis. *Cryst Growth Des.* 2007;7(4):755–769.
- Aguiar DLM, San Gil RAS, Borre LB, Marques MRC, Gemal AL. Evaluation of polymorphs in cephalixin medicines by ¹³C solid state NMR. *Int J Pharm Pharm Sci.* 2011;3(3):293–298.
- Souza EF, Chagas CA, Ramalho TC, et al. Study on the formation of crystalline vanadium nitride (VN) in low temperature through a fully solid-state synthesis route. *J Phys Chem C.* 2013;117(48):25659–25668.
- Kutzelnigg W. Theory of magnetic susceptibilities and NMR chemical shifts in terms of localized quantities. *Isr J Chem.* 1980;19:193–200.
- Pickard CJ, Mauri F. All-electron magnetic response with pseudopotentials: NMR chemical shifts. *Phys Rev B.* 2001;63:245101.
- Uldry AC, Griffin JM, Yates JR, et al. Quantifying weak hydrogen bonding in uracil and 4-cyano-4'-ethynylbiphenyl: a combined computational and experimental investigation of NMR chemical shifts in the solid state. *J Am Chem Soc.* 2008;130(3):945–954.
- Giannozzi P, Baroni S, Bonini N, et al. QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. *J Phys Condens Matter.* 2009;21(39):395502.
- Monkhorst HJ, Pack JD. Special points for Brillouin-zone integrations. *Phys Rev B.* 1976;13:5188–5192.
- Perdew JP, Burke K, Ernzerhof M. Generalized gradient approximation made simple. *Phys Rev Lett.* 1996;77(18):3865–3868.
- QuantumEspresso, pseudopotential library—PS library. Available at: <http://www.quantum-espresso.org/pseudopotentials/pslibrary/>. Accessed June 1, 2017.
- Harris RK, Hodgkinson P, Pickard CJ, Yates JR, Zorin V. Chemical shift computations on a crystallographic basis: some reflections and comments. *Magn Reson Chem.* 2007;45:S174–S186.
- Bruker S. *APEX Suite of Crystallographic Software, APEX 3 Version 2015-5.2*. Madison, WI: Bruker AXS Inc; 2015.
- Bruker S. *Version 8.34A and SADABS, Version 2014/5*. Madison, WI: Bruker AXS Inc; 2014.
- Sheldrick GM. *SADABS, Program for Empirical Absorption Correction of Area Detector Data*. Germany: University of Göttingen; 1996.
- Farrugia LJ. WinGX suite for small-molecule single-crystal crystallography. *J Appl Cryst.* 1999;32:837–838.
- Sheldrick GM. A short history of SHELX. *Acta Crystallogr A.* 2008;64:112–122.
- Spek AL. Structure validation in chemical crystallography. *Acta Cryst.* 2009;D65:148–155.
- Silverstein RM, Bassler GC, Morrill TC. *Spectrometric Identification of Organic Compounds*. 5th ed. Somerset, NJ: John Wiley & Sons, Inc; 1991.
- Hickey JP, Butler IS, Pouskouleli G. Carbon-13 NMR spectra of some representative hormonal steroids. *J Magn Reson.* 1980;38:501–506.
- Sene S, Berthomieu D, Donnadiou B, et al. A combined experimental-computational study of benzoxaborole crystal structures. *CrystEngComm.* 2014;16:4999–5011.
- Farrugia LJ. ORTEP-3 for windows - a version of ORTEP-III with a graphical user interface (GUI). *J Appl Cryst.* 1997;30:565–566.
- Makin HLJ, Gower DB, Kirk B. *Steroid Analysis*. Dordrecht, the Netherlands: Springer Science Business Media; 1995. Available at: <http://www.springer.com/gp/book/9789401730785#aboutBook>. Accessed June 1, 2017.
- Rohrer DC, Duax WL. 9 α fluoro-16 α -methyl-11 β ,17 α ,21-trihydroxy-1,4-pregnadiene-3,20-dione (dexamethasone) (C22H29FO5). *Cryst Struct Commun.* 1977;6:123–126.
- Raynor JW, Minor W, Chruszcz M. Dexamethasone at 119 K. *Acta Cryst.* 2007;E63:o2791–o2793.
- Maccaroni E, Giovenzana GB, Palmisano G, Botta D, Volante P, Masciocchi N. Structures from powders: diflorasone diacetate. *Steroids.* 2009;74(1):102–111.
- Valente EJ, Zubkowski JD, Lee KS, et al. 9 α -Fluoro-11 β ,17 α ,21-trihydroxy-1,4-pregnadiene-3,20-dione-21-acetate (9-fluoroprednisolone-21-acetate). *J Chem Cryst.* 1997;27(4):219–222.
- Allen FH, Watson DG, Brammer L, Orpen AG, Taylor R. Typical interatomic distances: organic compounds. In: Prince E, ed. *International Tables for Crystallography*. Vol. C, Chapter 9.5. New York, NY: Wiley; 2006:790–811.
- Simard S, Su D, Wuest JD. Use of hydrogen bonds to control molecular aggregation. Self-assembly of three-dimensional networks with large chambers. *J Am Chem Soc.* 1991;113(12):4696–4698.
- Fyfe MCT, Stoddart JF. Synthetic supramolecular chemistry. *Acc Chem Res.* 1997;30(10):393–401.
- Macrae CF, Edgington PR, McCabe P, et al. Mercury: visualization and analysis of crystal structures. *J Appl Cryst.* 2006;39:453–457.
- Jeffrey GA, Saenger W. *Hydrogen Bonding in Biological Structures*. Berlin: Springer-Verlag; 1991:20.
- Spackman MA, McKinnon JJ. Fingerprinting intermolecular interactions in molecular crystals. *CrystEngComm.* 2002;4:378–392.

ANEXO B

Access Structures Sign In

Your query was: Compound name: dexamethasone acetate and the search returned 3 records. New Search

Select all Download Selected View Selected

<input checked="" type="checkbox"/>	UZAJUC		Deposition Number(s): 1514791 Space Group: P 2 ₁ ; 2 ₁ ; (19) Cell: a 6.60090(10)Å b 15.7827(3)Å c 22.8607(5)Å, α 90° β 90° γ 90° Compound Name: 9-fluoro-11,17-dihydroxy-16-methyl-3,20-dioxopregna-1,4-dien-21-yl acetate monohydrate Synonyms: dexamethasone acetate monohydrate	
<input type="checkbox"/>	XURDAP		Deposition Number(s): 163804 Space Group: P 2 ₁ ; 2 ₁ ; (19) Cell: a 7.696(1)Å b 16.810(2)Å c 20.465(2)Å, α 90° β 90° γ 90° Compound Name: 9α-Fluoro-16α-methyl-11β,17,21-trihydroxy-1,4-pregnadiene-3,20-dione acetate dimethylsulfoxide solvate Synonyms: Dexamethasone acetate dimethylsulfoxide solvate	
<input type="checkbox"/>	XURDAP01		Deposition Number(s): 163803 Space Group: P 2 ₁ ; 2 ₁ ; (19) Cell: a 7.647(1)Å b 16.712(1)Å c 20.155(2)Å, α 90° β 90° γ 90° Compound Name: 9α-Fluoro-16α-methyl-11β,17,21-trihydroxy-1,4-pregnadiene-3,20-dione acetate dimethylsulfoxide solvate Synonyms: Dexamethasone acetate dimethylsulfoxide solvate	

[CCDC Home](#)
[Deposit Structures](#)
[Access Structures](#)
[About This Service](#)

CSD Entry: UZAJUC Sign In

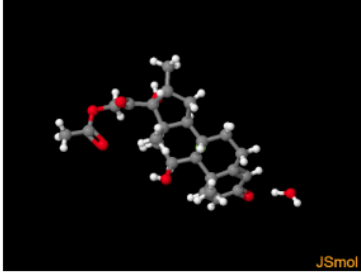
Your query was: Compound name: dexamethasone acetate and the search returned 3 records. Back to Search List New Search

Results	
Database Identifier	Deposition Number
<input checked="" type="checkbox"/> UZAJUC	1514791

[Download](#)

UZAJUC : 9-fluoro-11,17-dihydroxy-16-methyl-3,20-dioxopregna-1,4-dien-21-yl acetate monohydrate
 Space Group: P 2₁; 2₁; (19), Cell: a 6.60090(10)Å b 15.7827(3)Å c 22.8607(5)Å, α 90° β 90° γ 90°

3D viewer

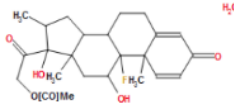


JSmol

[H](#)
[Disorder](#)
[↕](#)
[Menu](#)
[Open](#)

Style **Labels** **Packing** **Measure**
 (Ball and Stick) (No Labels) (None) (None)

Chemical diagram





[View group symbols key](#)

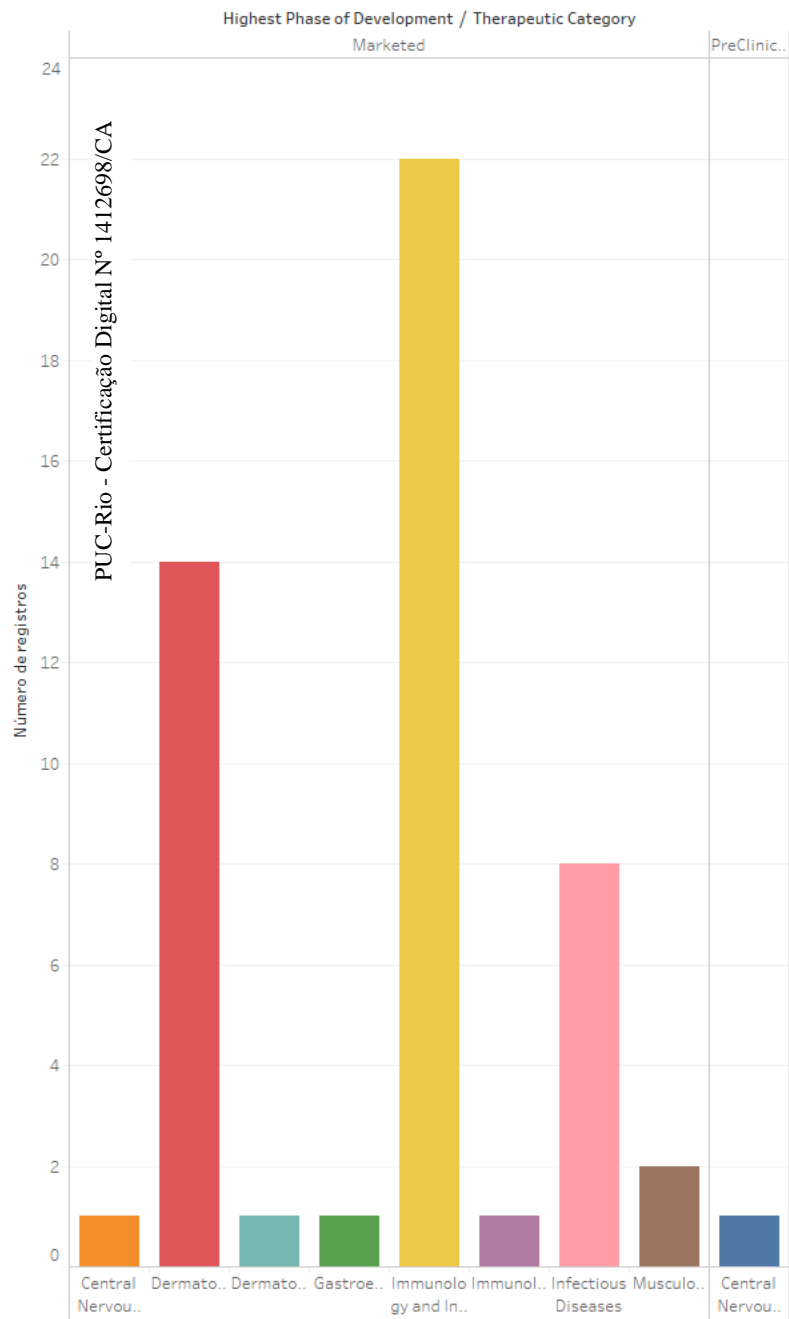
Additional details

Deposition Number	1514791
Data Citation	Lorenzo do Canto Visentin, Mateus Felipe Schuchter Ambrósio, Luciana Almeida Piovesan, Bruno Araujo Cautiero Horta, Maria Clara Ramalho Freitas, Ronaldo Pedro da Silva CCDC 1514791: Experimental Crystal Structure Determination, 2016, DOI: 10.5517/ccdc.csd.cc1mv87z
Additional Database Identifiers	UZAJUC01
Synonyms	dexamethasone acetate monohydrate
Deposited on	03/11/2016

Associated publications

-  Lorenzo do Canto Visentin, Mateus Felipe Schuchter Ambrósio, Luciana Almeida Piovesan, Bruno Araujo Cautiero Horta, Maria Clara Ramalho Freitas, Ronaldo Pedro da Silva, *CSD Communication*, 2016
-  Ronaldo Pedro da Silva, Mateus Felipe Schuchter Ambrósio, Luciana Almeida Piovesan, Maria Clara Ramalho Freitas, Daniel Lima Marques de Aguiar, Bruno Araujo Cautiero Horta, Eugenio Epprecht, Rosane Aguiar da Silva San Gil, Lorenzo do Canto Visentin, *Journal of Pharmaceutical Sciences*, 2018, 107, 672, DOI: 10.1016/j.xphs.2017.10.001

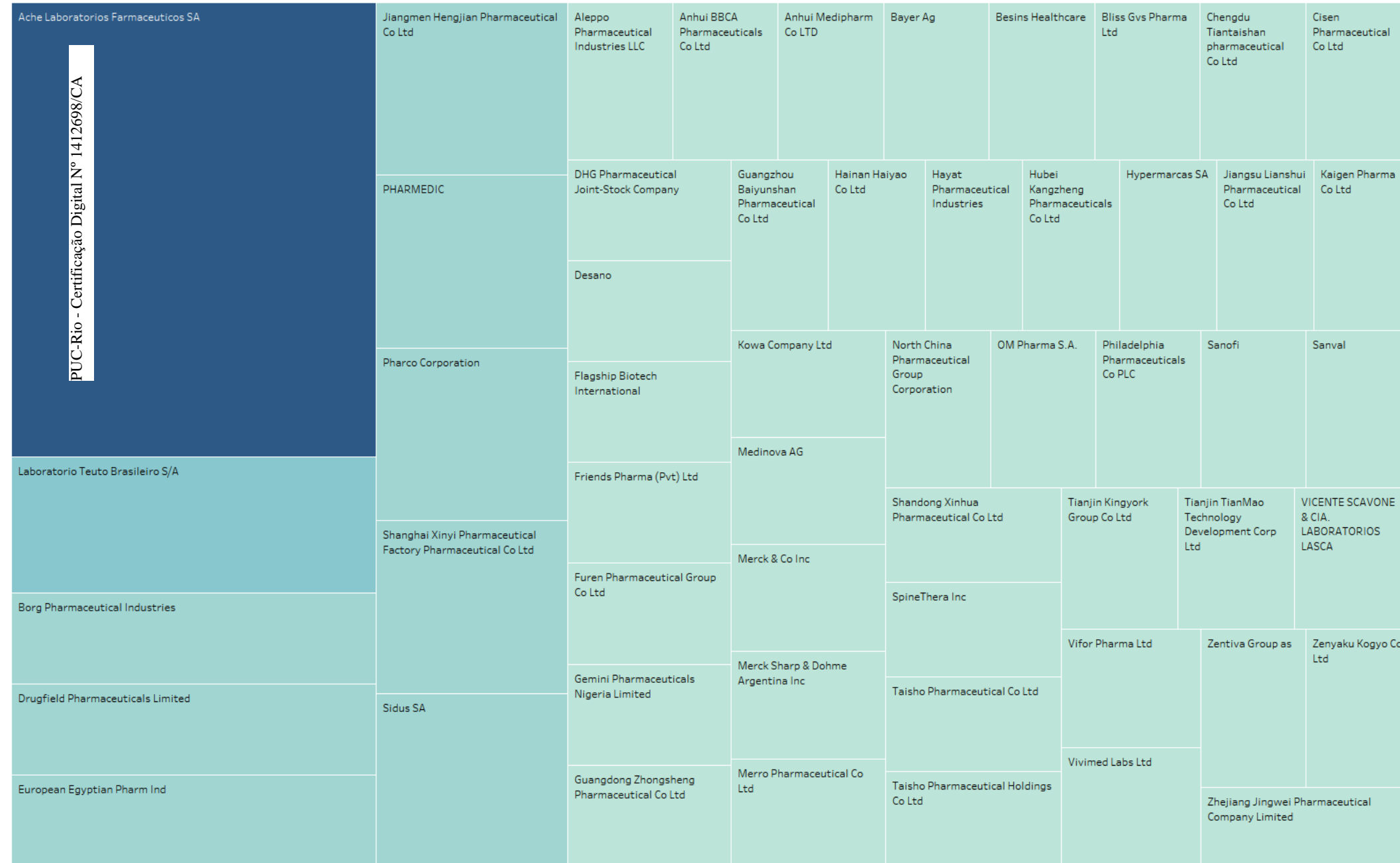
ANEXO C

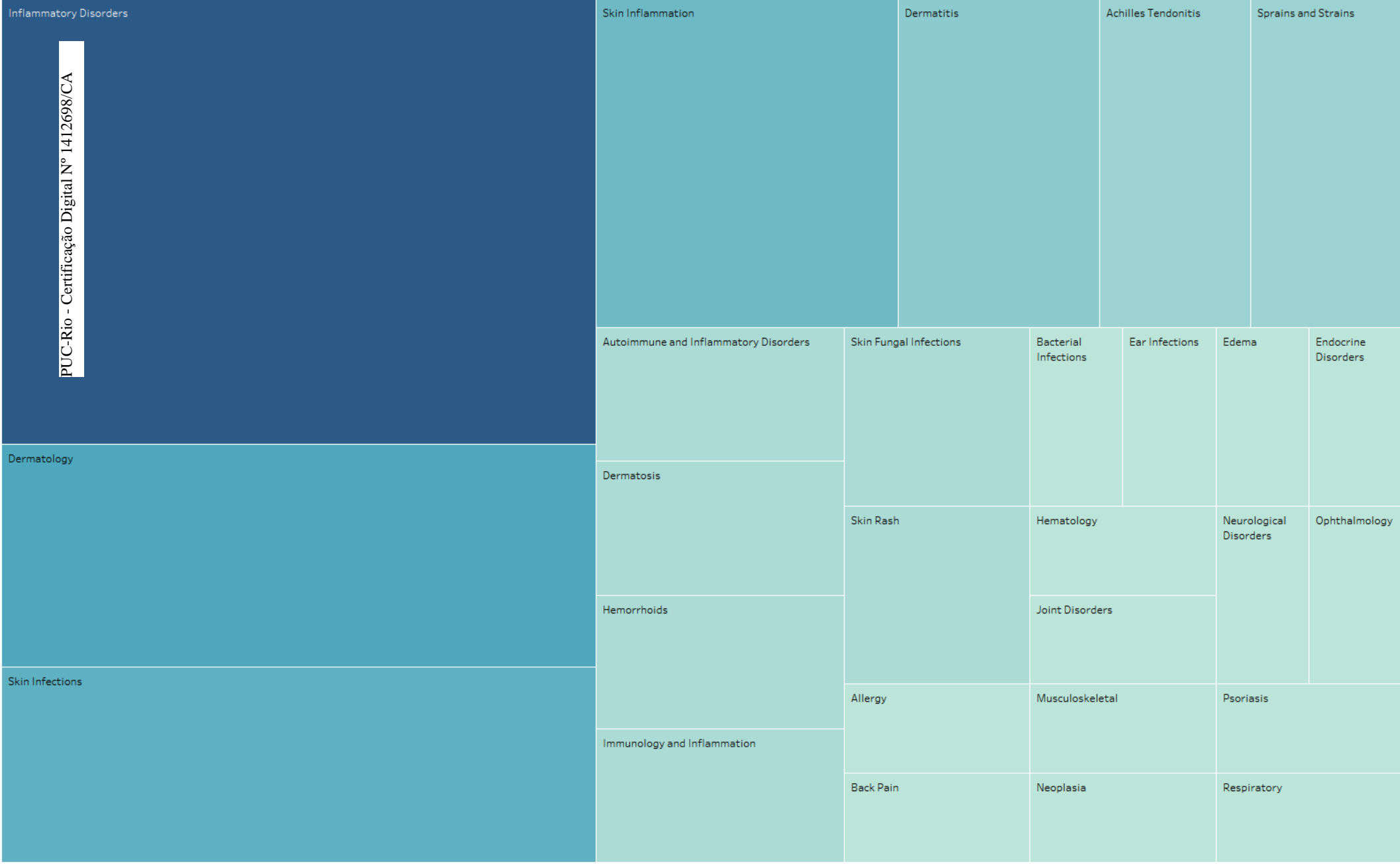




Número de produtos globais/Categoria Terapeutica

Therapeutic Category	Condition Treated	
Central Nervous System	Back Pain	1
	Neurological Disorders	1
Dermatolog	Dermatitis	4
	Dermatology	8
	Dermatosis	2
	Psoriasis	1
	Skin Rash	2
Edema	Edema	1
Endocrine, bolic and ..	Endocrine Disorders	1
Gastroente	ly Hemorrhoids	2
Hematolog	Hematology	1
Immunolog	Allergy	1
Inflammati	Autoimmune and Inflamm..	2
	Immunology and Inflamm..	2
	Inflammatory Disorders	16
	Skin Inflammation	6
Infectious I	ses Bacterial Infections	1
	Ear Infections	1
	Skin Fungal Infections	2
	Skin Infections	7
Musculoske...	Achilles Tendonitis	3
	Joint Disorders	1
	Musculoskeletal	1
	Sprains and Strains	3
Oncology	Neoplasia	1
Ophthalmology	Ophthalmology	1
Respiratory	Respiratory	1





Localização

PUC-Rio - Certificação Digital N° 1412698/CA



ANEXO D

PAPER • OPEN ACCESS

Validation of the method of quantitative phase analysis by X-ray diffraction in API: case of Tibolone

To cite this article: R P Silva *et al* 2016 *J. Phys.: Conf. Ser.* **733** 012030

View the [article online](#) for updates and enhancements.

Related content

- [Quantitative phase analysis of modified hardened cement paste](#)
O V Demyanenko, N O Kopanitsa, Yu S Sarkisov et al.
- [The magneto-X-ray Kerr effect in ferritic steel](#)
Jan Bonarski and Jan Karp
- [Instability of CaLi₂ at high pressure](#)
J. S. Tse, D. D. Klug, S. Desgreniers et al.

Validation of the method of quantitative phase analysis by X-ray diffraction in API: case of Tibolone

R P Silva^{1,2}, M F S Ambrósio^{1,5}, E K Epprecht², R R Avillez³, C A Achete⁴, A Kuznetsov⁴, L C Visentin⁵

¹ Universidade Federal do Rio de Janeiro, Campus de Xerém, 25245-390, Duque de Caxias, RJ – Brasil

² Pontifícia Universidade Católica do Rio de Janeiro, Centro Técnico-Científico, Departamento de Engenharia Industrial, 22451-900, Rio de Janeiro, RJ - Brasil

³ Pontifícia Universidade Católica do Rio de Janeiro, Centro Técnico-Científico, Departamento de Engenharia Química e de Materiais, 22451-900 - Rio de Janeiro, RJ – Brasil

⁴ INMETRO, Instituto Nacional de Metrologia, Qualidade e Tecnologia, 25250-020, Duque de Caxias, RJ - Brasil

⁵ NanoBusiness Informação e Inovação Ltda, 25250-020, Duque de Caxias, RJ – Brasil

E-mail: lorenzo.visentin@nanobusiness.com.br

Abstract. In this study, different structural and microstructural models applied to X-ray analysis of powder diffraction data of polymorphic mixtures of known concentrations of Tibolone were investigated. The X-ray data obtained in different diffraction instruments were analysed via Rietveld method using the same analytical models. The results of quantitative phase analysis show that regardless of the instrument used, the values of the calculated concentrations follow the same systematics with respect to the final errors. The strategy to select a specific analytical model that leads to lower measurement errors is here presented.

1. Introduction

The quantitative phase analysis by X-ray diffraction (XRD) of organic solids which compose active pharmaceutical ingredients (API) is susceptible to many factors (instrumental, knowledge of materials properties, and method of data treatment) that influence the final result.

The Rietveld method is accepted as the most accurate approach to extract from X-ray data the values of concentrations of phases in phase mixtures [1]. However, the choice of structural and microstructural model in this method of analysis becomes a critical factor from the point of view of its contribution to the errors of calculated concentrations. This situation represents a serious problem for the pharmaceutical industry and regulatory organizations in production and distribution of safe and efficient solid drugs.

This study investigates, in a systematic manner, different assumptions about structural and microstructural properties of phases of Tibolone during the data modelling. Tibolone is a synthetic steroid used to relieve hypo-estrogenic symptoms and protect against bone loss in post-menopausal

women [2]. The general and specific constraints for the reliable quantitative phase analyses were established on the base of multiple experimental data and variable concentrations of phases of Tibolone.

2. Materials and methods

2.1. Sample Preparation

Three standard mixtures of known concentrations, referred hereafter as M1, M2 e M3, were prepared using Tibolone pure Form I (commercial) and Form II (prepared according to Araujo, et.al.) [2]. The concentrations of the mixtures are shown in the Table 1. A calibrated analytical balance with the error of indication of -0,01 mg and the expanded measurement uncertainty (coverage factor $k = 2$) of 0.01 mg was used to establish the real phase proportions in the studied samples.

The overall mass of each prepared sample was approximately 1 g. The mixtures were placed in eppendorfs and stirred in vibratory mixer mill for 30 minutes. Table 1 presents the acceptable lower limits (M1, preferably M2) concentrations for a crystalline pure Form I [3], and the expected guess for a detection limit of Form II in Tibolone.

Table 1. Sample identification and phase composition.

Tibolone (API)	Nominal [F1-F2] (%)	Real [F1-F2] (%)
M1	90-10	90.09-9.91
M2	95-5	94.99-5.01
M3	99-1	98.73-1.27

F1= Tibolone Form I and F2 = Tibolone Form II

2.2. X-ray data collection

The instruments used for X-ray data collections, their basic configurations, and sample preparation aspects are listed in the Table 2.

Table 2. Instruments and sample preparation types used in X-ray experiments.

Instruments	Geometry	Tube	Detector	Powder Sample
X'Pert Pro Panalytical ¹	Reflection, Bragg-Brentano	CuK α	PiXcel	Mixed and pressed *Mixed, grinded and pressed
D8 Focus Bruker ²	Reflection, Bragg-Brentano	CoK α	Scintillator	Mixed and pressed
Agilent Technologies ³	Transmission, Debye Scherrer	CuK α	Charge-coupled device (CCD)	Mixed powder in capillary

¹ NanoBusiness Informação e Inovação Ltda.

² Inmetro - Laboratório de Difração e Espectroscopia, Divisão de Metrologia de Materiais (DIMAT).

³ Inmetro - Diretoria de Metrologia Aplicada às Ciências da Vida (DIMAV).

2.3. Structural and Microstructural models

Eight models of Form I and Form II of Tibolone were employed to calculate the concentrations of these phases in the samples. Different constraints for structural and/or microstructural properties of phases distinguish these models (see Table 3).

Model 1 assumes the fixed structural parameters for Form I and Form II of Tibolone taken from data base of Cambridge Crystallography Data Centre (No.: Form I 1126814 and Form II 1126815), and isotropic crystallite size peaks broadening (Voigt function) for diffraction peak shape fitting. Subsequent models include other refinement options, as shown in Table 3.

Table 3. Description of models.

Structural and Microstructural Models	Model Constraints							
	1	2	3	4	5	6	7	8
1. Isotropic crystallite size broadening	●	●	●	●	●	●	●*	●
2. Lattice Parameters	●	●	●	●	●	●	●	●
3. Texture (spherical harmonics)	●	●	●	●	●	●	●	●
4. Hydrogen atoms inserted in Form I	●	●	●	●	●	●	●	●
5. Isotropic atomic displacement parameter, B_{eq}	●	●	●	●	●	●	●	●
6. Atomic coordinates (constrained to maximum displacement of 0.0016)	●	●	●	●	●	●#	●	●
7. Anisotropic crystallite size broadening	●	●	●	●	●	●	●#	●
8. Anisotropic atomic displacement parameters, U_{ij}	●	●	●	●	●	●	●	●

● Applied or variable

● Not applied or not variable

* Applied only to Form II

Applied only to Form I

3. Results and Discussion

In Figure 1 it is shown the selected diffraction patterns of samples M1, M2 and M3 obtained on instrument 1. Figure 2 shows the experimental X-ray diffraction pattern of M1 sample and the one calculated from the Rietveld refinement of model 8.

The use of the latter model results in the excellent agreement between calculated and real concentrations of phases in the sample. Figure 3 summarises the relative errors of the calculated concentrations of Form II for all samples (Table 1) and experimental conditions (Table 2) used in this study.

As one can expect, the values of relative errors increase with decreasing concentration of Form II. This is a general trend, which can vary for a particular structural and microstructural model. One can also note that the results obtained on instruments 1 and 2 have markedly better agreement with the real values of concentrations than the ones of instrument 1* and 3.

The main distinction between these sets of conditions is the 2Θ resolution of diffraction peaks. This resolution is worse in case of instruments 3 and 1* due to a low angular resolution of CCD detector and additional crystallite size peak broadening induced by grinding, respectively.

The inset of Figure 3 compares the values of real relative errors (the relative differences between the real and calculated concentrations) and standard deviations for concentrations of Form II obtained from Rietveld refinement of a particular model. The calculated and real errors are comparable for models that result in a weighted profile residual (R_{wp} -factor) bellow 12 and a goodness of fit (GoF) bellow 2 (see Figure 4).

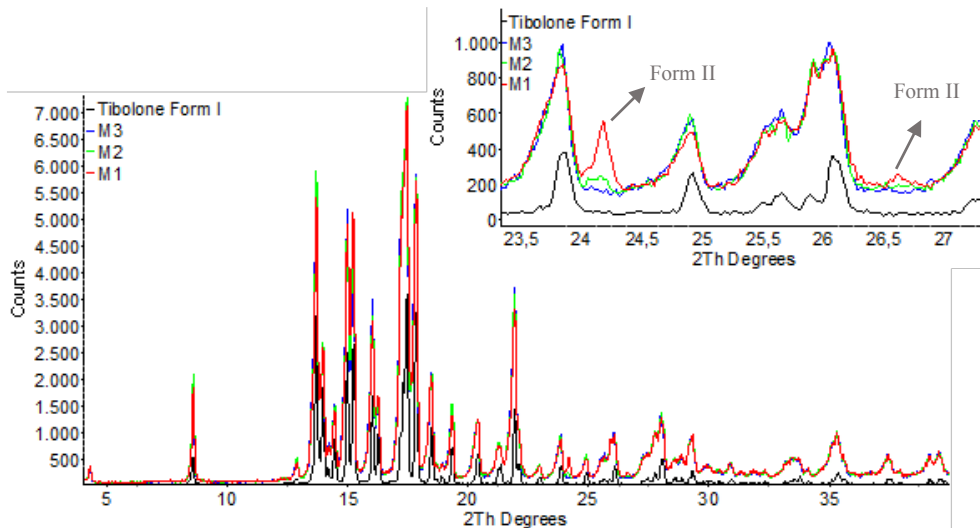


Figure 1. Diffraction patterns of M1, M2 and M3 samples. Inset emphasises the 2θ regions of two not overlapped peaks of Form I and Form II.

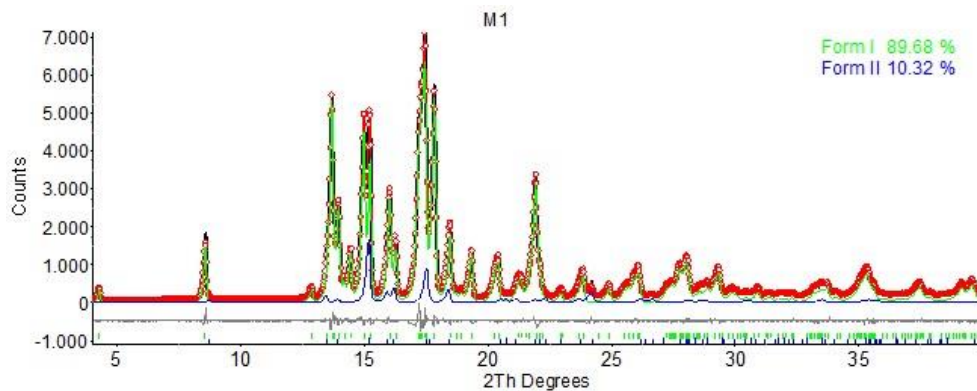


Figure 2. Experimental (black curve), calculated (red curve) and difference (grey) from Rietveld refinement (model 8) XRD patterns of the sample M1. Form I (green curve) and Form II (blue curve)

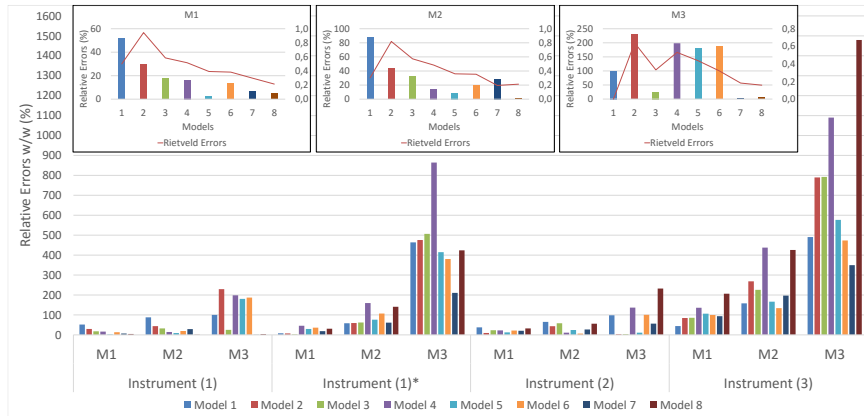


Figure 3. Relative errors in concentration. The three extended figures is from measures in the instrument 1.

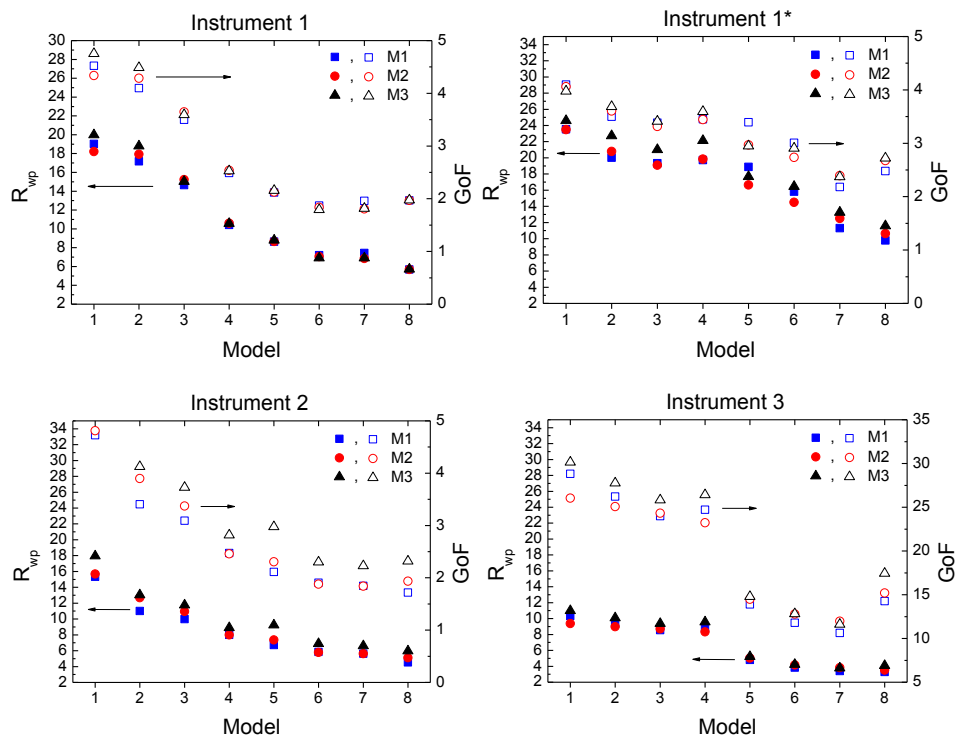


Figure 4. R_{wp} and GoF indexes ($R_{wp} = \sqrt{\sum w_m(Y_{o,m} - Y_{c,m})^2 / \sum w_m Y_{o,m}^2}$; $GoF = chi2 = R_{wp} / R_{exp}$).

4. Conclusion

The obtained results showed that a good resolution of X-ray diffraction data is a necessary condition for reliable quantitative analysis of polymorphism in Tibolone. The Rietveld method of analysis of diffraction patterns gives reliable estimations of concentrations of Form II with respective errors for concentrations above 5 % w/w. The modelling of microstructural (texture) and structural (H atoms, ADPs) features was required to achieve this result. Additional models sophistications, such as consideration of anisotropy of crystallite size broadening and anisotropy of ADPs may be necessary to obtain reliable results in case of lower concentrations of Form II.

Acknowledgments

The e-Diffraction Lab from NanoBusiness Company is grateful to the agencies CNPq, FAPERJ and to INMETRO for the collaboration in this research. An especially acknowledgment to Daniel Giacometti by management support in this project.

5. References

- [1] Young R A 1993 *Introduction to the Rietveld method. - The Rietveld Method* (Oxford: Oxford University Press - IUCr Book Series) pp 1-39.
- [2] Araujo G L B, Faria D L A, Zaim M H, Carvalho F M S, Andrade F R D and Matos J R 2010 *J. Therm. Anal. Calorim.* **102** 233
- [3] Akzo Nobel N.V. G A J M T Sas *et. al.* Pharmaceutical composition which contains a pharmaceutically suitable carrier and the compound having the structure (7 α ,17 α)-17- hydroxy-7-methyl-19- nor-17- pregn-5(10)- en-20- yn-3- one. EP0389035A1, 12 mar. 1990, 26 sep. 1990.

APÊNDICE A

CENTRO TÉCNICO CIENTÍFICO
DEPARTAMENTO DE ENGENHARIA INDUSTRIAL
MEMORANDO TÉCNICO

MEMORANDO TÉCNICO No. 01/18

**NEW APPROACHES FOR VALIDATION AND EVALUATION OF
MEASUREMENT UNCERTAINTY BY POLYCRYSTAL X-RAY
DIFFRACTION (PXRD)**

**Ronaldo Pedro da Silva
Lorenzo do Canto Visentin
Luciana Almeida Piovesan
Mateus Felipe Schuchter Ambrósio
Eugenio Kahn Epprecht
Paulo Roberto Guimarães Couto
Jailton Carreteiro Damasceno**

26 de Janeiro de 2018

PUC

Pontifícia Universidade Católica do Rio de Janeiro
Rua Marquês de São Vicente, 225 - CEP 22451-900
Rio de Janeiro - Brasil

New approaches for validation and evaluation of measurement uncertainty by Polycrystal X-Ray Diffraction (PXRD)

*Ronaldo Pedro da Silva,¹ Lorenzo do Canto Visentin,¹ Luciana Almeida Piovesan,¹
Mateus Felipe Schuchter Ambrósio,¹ Eugenio Kahn Epprecht,² Paulo Roberto
Guimarães Couto,³ Jailton Carreteiro Damasceno.³*

¹ Nanobusiness Informação e Inovação Ltda, Incubadora de Projetos, Instituto Nacional de Metrologia, Qualidade e Tecnologia – Inmetro, Duque de Caxias, Brazil

² Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Industrial, Rua Marquês de São Vicente, 225, Rio de Janeiro, RJ, 22451-900, Brazil

³ Instituto Nacional de Metrologia, Qualidade e Tecnologia – Inmetro, Duque de Caxias, Brazil.

Abstract: Following the requirements established in ISO / IEC 17025: 2005, a measurement result is reliable when traceability is evidenced and also has proven metrological reliability. The initial goal of ISO GUM 2008 is the international harmonization of methodology evaluation measurement uncertainty. In this way it is possible that comparisons between measurement results are performed, an activity necessary for the existence of mutual recognition between metrological systems. Only the uncertainty of a measurement does not ensure that a result holds the quality appropriate to a specific need. Thus, prior to the evaluation of the measurement uncertainty of a result, its validation must be performed. This work presents new methodological approaches for the metrological validation and measurement uncertainty assessment for X-ray diffraction by polycrystals (DRXP).

Keywords: polymorphism, validation, uncertainty, metrology, diffraction

1. Introduction and context

According to the requirements established in ISO / IEC 17025: 2005 [1], for a reliable measurement result, it must present evidence regarding its traceability and also

have proven its metrological reliability. These evidences are established as important properties for the evaluation of the quality of a measurement. However many times the measurement results in the various areas are presented without the concern of these important and necessary requirements [2]. ISO GUM 2008 [3] has as its initial objective the international harmonization of measurement uncertainty assessment, so that it is possible to compare the measurement results. Comparison of measurement results is fundamental to the existence of mutual recognition among metrological systems, an important action that bases the needs of an increasingly globalized market [2].

It should be noted that only the measurement uncertainty assessment does not ensure its quality, even if the result is obtained by a standardized method. Thus, the need for validation of a result that is a step prior to the evaluation of its measurement uncertainty arises. The validation of a measurement result has the purpose of confirming by examination and supplying objective evidence that the specific requirements for a given intended use are met [1].

The literature at moment show a minimal information about validation and evaluation of measurement uncertainty in Polycrystal X-Ray Diffraction (PXRD) in related to the control of polymorphism in active pharmaceutical ingredients (API) and solid drugs. Following the guidelines of Food and Drug Administration Amendments Act (FDAAA), 2007 [4], the pharmaceutical industries should monitor the effect of polymorphism on drugs. The most drugs are produced in solid form. In this way the chemical active substances may present polymorphism. A phenomenon of possible undesired consequences for the drug, such as decreased efficacy, absence of therapeutic activity or even rendering it toxic [5]. The most widely accepted definitions for polymorphism are based on the ability of the solid material with same chemical formula exist in two or more crystalline phases with different arrangements or conformations in the crystalline structure [6], which can lead to different chemical and physical properties. Many researchers consider important of polymorphism control in API and drugs in function of the fine relation between polymorphism appear and to possible differences in dissolution rate, which may directly affect their bioavailability and hence efficacy [7]. In addition, the process of solid drug development can generate the polymorphic transformation within the own process. This effect can be cause also processability problems, influencing processes of grinding, micronization, lyophilization, compression and granulation.

Therefore, the polymorphism affects the quality, safety and efficacy of the drug, making it necessary to control quality and stability from drug acquisition to final product storage, posing a challenge in the production of drugs of adequate and ratified quality [8].

For this monitoring the application of the single crystal X-ray diffraction (SCXRD) bring accurate information about lattice parameters and its crystalline structure, being considered the definitive technique for evidence of polymorphism [4]. The literature [4], [9] and [10] dedicate a chapter about to the characterization of crystalline and partially crystalline solids by PXRD.

However, the difficulty in producing single crystal suitable for diffraction procedure in single crystal can makes restricted the application of this technique. Thus, PXRD is considered other method for identification of the existence of polymorphism with excellent accurate and is, admittedly, one of the worldwide references for the identification of polymorphism in solid API. Other techniques are considered to be complementary in the characterization of the solid state of a substance, including here optical microscopy and scanning electron microscopy, differential scanning calorimetry and thermogravimetric analysis, Fourier transform infrared, Raman and Solid Nuclear Magnetic Resonance [9].

Diffraction does not observe the atomic and molecular world in the same way as microscopy, so the data generated from a diffraction experiment are always mathematical data of a reciprocal space to real space. Thus, the importance of having reliable data undoubtedly passes through the context of uncertainty measurement so that the raw data of this measurement, after statistical treatment, can be trusted within acceptable values of accuracy. In this context, the reproducibility of diffraction measurements must undergo The law established for diffraction is the Bragg law, which establishes the relationship between the diffraction angles and the interplanar distance, thus originating the phenomenon of diffraction and the possibility of diffraction data collecting. The Bragg law is described by equation 1.

$$n\lambda = 2d\sin\theta \quad 1$$

Where: Integer;

$n =$

$\lambda =$ Wavelength of incident X-rays;

$d =$ Interplanar distance;

$\theta =$ Diffraction angle.

Since 1920 the Bragg equation is applied to determine lattice and peak diffraction positions in θ and 2θ as a function of the diffraction incident angle on the sample. To identify with precision and accuracy the peak position diffraction is very important in many cases of diffraction studies such as phase identification, indexing of diffraction patterns, determination of lattice parameters, etc. The measurement uncertainty is present also on the method of "instrumental monitoring of diffraction" [11].

In relation to the statistical control of the diffraction data in order to increase confidence in measurements, Taupin [12] used the "BIAS" method and least squares method to optimize and improve accuracy of peak positions data, intensities and width from data complex numbers obtained by XRD. Afterwards, Huang [13] presented good results when using derivative methods, combining the first derivative of a cubic polynomial with the second derivative of quadratic/cubic polynomial, for precise determination of the diffraction peak position.

The International Union of Crystallography (IUCr) working group in 1995 reported the general recommendations for measurement uncertainty of types A and B in the field of crystallography [14]. In order to evaluate the statistical quality of the intensity data and 2θ angles as a function of the step, Rebmann et al [15] used systematic approaches by Poisson statistics for the intensities and the assumption of a Gaussian distribution for the observed 2θ angle around its average value.

Since the 2000s, important works were published regarding the estimation of errors in the measurement of unit cell parameters [16], the precise location of the diffraction peak and the expected uncertainty [17].

Given the importance of PXRD in controlling drug and drug polymorphism, this work presents new methodological approaches for metrological validation and measurement uncertainty assessment of crystalline substance analysis by PXRD.

2. Methodologies

2.1.1 DRX Calibration Curve

From the values indicated (2θ) by the DRX for each of the respective reference values ($2\theta_{\text{ref}}$) contained in the DRX certificate, its calibration curve was elaborated, which corrects the systematic error of the DRX for any 2θ . Thus, the result of the diffraction angle of the analysis of a sample from the DRX indication is calculated by the expression 2:

$$2\theta_{\text{sample}} \left[^0 \right] = a + b \cdot (2\theta_{\text{i analysis}}) \quad 2$$

Onde:

- $2\theta_{\text{sample}} =$ Sample diffraction angle;
- $a =$ Linear coefficient of the DRX calibration curve;
- $b =$ Angular coefficient of the DRX calibration curve;
- $2\theta_{\text{i analysis}} =$ Diffraction angle indicated by the DRX for the analysis.

2.1.2 Validation of the measurement method

2.1.2.1 Standard deviation of measurements

For validation, four-angle ($2\theta_{\text{ref}}$) DRX measurements of the Certified Reference Material were performed. For each of the ($2\theta_{\text{ref}}$) angles six replicate sequences were performed and each contained ten replicates of the measured angle values (2θ) by the XRD for each reference values ($2\theta_{\text{ref}}$). For each value (2θ) of the sequences were calculated the errors (e_i), their means and respective variances. For the validation of the error variances of each sequence, the control charter of the standard deviation was first elaborated and later the Cochran parametric test was applied.

Control chart of the standard deviations

For the preparation of the control chart of the standard deviation of sequences the upper limit used was the tolerance of $0,05^\circ$.

Cochran parametric test

For numbers of measurement sequences ($K_{\text{sequence}} > 2$) the comparison between the variances of the measurement sets is performed by comparing the Cochran statistic calculated (C_{cal}) with the table (C_{table}) for an established probability. The Cochran statistic calculated is by the expression 3.

$$C_{\text{cal}} = \frac{S_{\text{max}}^2}{\sum S_i^2} \quad 3$$

Where:

$S_{\text{max}}^2 =$ Maximum variance of all sequences K

$S_i^2 =$ Variance of each sequence K

The condition for considering that S_{max}^2 is equal to all variances (S_i^2) for a statistical probability is that $C_{\text{cal}} < C_{\text{table}}$. The values of " C_{table} " the values of "X" are shown in the attached table. If this is true then the pooled deviation for the variances is given by equation 4.

$$s_p = \sqrt{\frac{\sum v_i S_i^2}{\sum v_i}} \quad 4$$

Where:

$v_i =$ Number of degrees of freedom of each sequence K.

2.1.2.2 Mean of sequence errors

Control chart of the means of the sequence errors

The upper limit of 0.05 was used, as the tolerance, for the preparation of the control chart of the means of the errors of each of the sequences

2.2 Evaluation of measurement uncertainty

A incerteza avalia de forma pontual a qualidade de um resultado de medição. The evaluation of measurement uncertainty is in accordance with the ISO GUM 2008 methodology, which can be summarized in the following main steps: *i)* Definition of the measurand; *ii)* Diagram causes effect; *iii)* Evaluation of sources of uncertainty of input quantities; *iv)* Calculation of the sensitivity coefficients; *v)* Calculation of uncertainty components; *vi)* Calculation of combined uncertainty; *vii)* Calculation of the number of effective degrees of freedom; *viii)* Determination of the coverage factor; *ix)* Evaluation of the expanded uncertainty.

2.2.1 Measurand – (Diffraction angle of the sample)

Measurand [VIM] is the diffraction angle of the analysis of a sample, the evaluation of its measurement uncertainty can be from expression 5.

$$2\theta_{\text{sample}} \left[^0 \right] = a + b \cdot (2\theta_{1\text{analysis}}) + \delta_{\text{certificate DRX}} + \delta_{\text{repetitionanalysis}} + \delta_{\text{method}} + \delta_{\text{MRC}} \quad 5$$

Where:

$2\theta_{\text{sample}} =$	Diffraction angle of the sample; ;
$a =$	Linear coefficient of the DRX calibration curve;
$b =$	Angular coefficient of the DRX calibration curve;
$2\theta_{1\text{analysis}} =$	Diffraction angle indicated by the DRX for the analysis.
$\delta_{\text{method}} =$	Uncertainty arising from method validation;
$\delta_{\text{repetitionanalysis}} =$	Uncertainty regarding the repetition of the values of the analysis;
$\delta_{\text{MRC}} =$	Uncertainty of the certified reference material.

2.2.2 Cause-Effect Diagram

According to expression 5 the sources of uncertainty are presented in the diagram causes effect of figure 1:

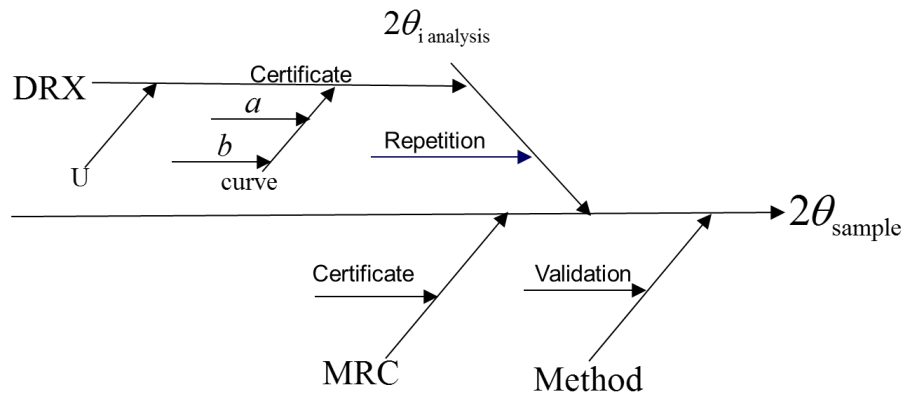


Figure 1- Diagram Cause Effect of the uncertainty evaluation of the Refraction Angle

2.2.3 Uncertainties of input quantities

2.2.3.1 $2\theta_i$ indicated by DRX

- *DRX Certificate.*

The uncertainty of the DRX certificate is evaluated by expression 6.

$$u_1 = \frac{U(2\theta_i)}{k(2\theta_i)} [^\circ] \quad 6$$

Where:

$U(2\theta_i)$ = Expanded Uncertainty Declared on DRX Calibration Certificate

$k(2\theta_i)$ = Coverage factor declared on the certificate.

- *Calibration curve*

Uncertainty regarding the calibration curve according to [3] and [19] can be evaluated, respectively, by the expressions 7a or 7b.

$$u_2 = u_{2\theta_{\text{sample}}}(a, b) = S_e \cdot \sqrt{\frac{1}{n_t} + \frac{(2\theta_{i \text{ analysis}} - \overline{2\theta_i})^2}{\sum (2\theta_i - \overline{2\theta_i})^2}} \quad [0] \quad 7a$$

$$u_2 = u_{2\theta_{\text{sample}}}(a, b) = \sqrt{1^2 \cdot S_a^2 + (2\theta_{i \text{ analysis}})^2 \cdot S_b^2 + 2 \cdot 1 \cdot 2\theta_{i \text{ analysis}} \cdot u_a \cdot u_b \cdot r_{a,b}} \quad [0] \quad 7b$$

Where:

- S_e = DRX Calibration Curve Dispersion;
- n_t = Number of Points DRX Certificate;
- $2\theta_{i \text{ analysis}}$ = Value indicated by DR X of the analysis angle;
- $\overline{2\theta_i}$ = Mean of the indicated angle values in the DRX certificate;
- $2\theta_i$ = DRX Certificate Angle Indicator Values.

The dispersion of the calibration curve of the DRX (S_e), variances s_a^2 and s_b^2 ; uncertainties u_a and u_b , the correlation coefficient between a and b are calculated respectively by the expressions 8, 9, 10, 11, 12 and 13.

$$S_e = \sqrt{\frac{[2\theta_{\text{ref}} - (a + b \cdot (2\theta_i))]^2}{n_t - 2}} \quad [0] \quad 8$$

$$s_a^2 = \frac{s_e^2 \sum (2\theta_i)^2}{D} \quad [0] \quad 9$$

$$s_b^2 = n_t \frac{s_e^2}{D} \quad 10$$

$$u_a = \sqrt{s_a^2} \quad [0] \quad 11$$

$$u_b = \sqrt{s_b^2} \quad 12$$

$$r_{a,b} = -\frac{\sum (2\theta_i)}{\sqrt{n_t \sum (2\theta_i)^2}} \quad 13$$

Where:

D is the determinant of the matrix (D) which is calculated by the expression 14.

$$D = n_i \sum (2\theta_i)^2 - \sum [(2\theta_i)]^2 \quad (0)^2 \quad 14$$

- *Repetitions of the analysis measurements*

The uncertainty concerning the repetition of measurements of the analysis is calculated by the expression 15.

$$u_3 = \frac{s}{\sqrt{n}} \quad [0] \quad 15$$

Where:

$s =$ Standard deviation of measurements repetitions analysis,

$n =$ Number of repetitions of measurements.

2.2.3.2 Validation of the method

If $C_{cal} < C_{tab}$ the uncertainty regarding the validation of the method is evaluated by the expression 16

$$u_4 = \frac{S_p}{\sqrt{N}} = \frac{\sqrt{\frac{\sum v_i S_i^2}{\sum v_i}}}{\sqrt{N}} \quad 16$$

Where: N is the total number of measurements of the K sequences.

2.2.3.3 Certified Reference Material

The evaluation of the uncertainty of the certificate of reference material is by expression 17:

$$u_s = \frac{U(MRC)}{k(MRC)} \quad 17$$

Where:

$U(MRC)$ = Expanded uncertainty stated in MRC certificate ;
 $k(MRC)$ = Coverage factor declared on the certificate.

2.2.4 Sensitivity Coefficients

From equation 5, the coefficients of sensitivity of the measurand are calculated by equations 18, 19, 20, 21 and 22.

$$\frac{\partial(2\theta_{\text{sample}})}{\partial(\text{DRX})} = 1 \quad 18$$

$$\frac{\partial(2\theta_{\text{sample}})}{\partial(a,b)} = 1 \quad 19$$

$$\frac{\partial(2\theta_{\text{sample}})}{\partial(\text{repetition})} = 1 \quad 20$$

$$\frac{\partial(2\theta_{\text{sample}})}{\partial(\text{method})} = 1 \quad 21$$

$$\frac{\partial(2\theta_{\text{sample}})}{\partial(MRC)} = - \quad 22$$

The sensitivity coefficient for the DRX calibration curve is equal to 1 because equations 7a and 7b already evaluate sample uncertainty in the unit of $2\theta_{\text{sample}}$.

2.2.5 Components of uncertainties

The uncertainty components of the sample analysis are defined by expression 23 to 27

$$u_{2\theta_{\text{sample}}}(\text{DRX}) = 1 \cdot u_1 \quad 23$$

$$u_{2\theta_{\text{sample}}}(a, b) = 1 \cdot u_2 \quad 24$$

$$u_{2\theta_{\text{sample}}}(\text{repetition}) = 1 \cdot u_3 \quad 25$$

$$u_{2\theta_{\text{sample}}}(\text{method}) = 1 \cdot u_4 \quad 26$$

$$u_{2\theta_{\text{sample}}}(\text{MRC}) = - \quad 27$$

2.2.6 Components of relative uncertainties

The relative uncertainty components of the analysis are defined by expression 28 to 32.

$$\frac{u_{2\theta_{\text{sample}}}(\text{DRX})}{2\theta_{\text{sample}}} = u_5 \quad 28$$

$$\frac{u_{2\theta_{\text{sample}}}(a, b)}{2\theta_{\text{sample}}} = u_6 \quad 29$$

$$\frac{u_{2\theta_{\text{sample}}}(\text{repetition})}{2\theta_{\text{sample}}} = u_7 \quad 30$$

$$\frac{u_{2\theta_{\text{sample}}}(\text{method})}{2\theta_{\text{sample}}} = u_8 \quad 31$$

$$\frac{u_{\text{MRC}}}{\text{MRC}} = u_9 \quad 32$$

2.2.7 Combined relative uncertainty

The combined relative uncertainty of the diffraction angle of the sample is defined by equation 33

$$\frac{u_{2\theta_{\text{sample}}}}{2\theta_{\text{sample}}} = \sqrt{u_5^2 + u_6^2 + u_7^2 + u_8^2 + u_9^2} \quad 33$$

2.2.8 Combined uncertainty

From equation 33 the combined uncertainty of the of the sample analysis is evaluated by expression 34.

$$u_c(2\theta_{\text{sample}}) = \sqrt{u_5^2 + u_6^2 + u_7^2 + u_8^2 + u_9^2} \cdot 2\theta_{\text{sample}} \left[^\circ \right] \quad 34$$

2.2.9 Number of effective degrees of freedom

Depending on the case, the number of effective degrees of freedom is defined by equations 34a and 34b.

$$v_{\text{eff}} = \frac{u_c^4(2\theta_{\text{sample}})}{\frac{u_2^4}{(n_t - 2)} + \frac{u_3^4}{(n - 1)}} \quad \text{if } C_{\text{cal}} > C_{\text{tab}} \quad 34a$$

$$v_{\text{eff}} = \frac{u_c^4(2\theta_{\text{sample}})}{\frac{u_2^4}{(n_t - 2)} + \frac{u_3^4}{(n - 1)} + \frac{u_4^4}{(n_1 + n_2 + \dots n_K - K)}} \quad \text{if } C_{\text{cal}} \leq C_{\text{tab}} \quad 34b$$

Where:

- n_i = Number of replicates of the sample analysis;
- n_t = Number of points DRX calibration certificate
- K = Number of sequences.

2.2.10 Coverage factor

Depending on the case ($C_{\text{cal}} > C_{\text{table}}$ or $C_{\text{cal}} < C_{\text{table}}$), the coverage factor (k) is defined according to the number of effective degrees of freedom calculated by equations 34a or 34b, for a 95% coverage probability.

2.2.11 Expanded uncertainty

The expanded uncertainty is then evaluated by 35:

$$U = k \cdot u_c(2\theta_{\text{sample}}) \left[^\circ\right] \quad 35$$

2.3 Measurand (Distance of the crystalline structure of a sample)

In the evaluation of the uncertainty of the interplanar distance of the structure of a sample by XRD the measurand [3] can be described by expression 36 .

$$d_{\text{sample}} = \frac{\lambda(\text{nm})}{2\text{sen}\left(\frac{\pi}{180} \cdot \frac{2\theta_{\text{sample}}}{2}\right)} \cdot 10(A^\circ) \quad 36$$

Onde:

λ = wavelength of incident ray (nm);

$2\theta_{\text{sample}}$ = diffraction angle of the sample ($^\circ$).

2.3.1 Diagram cause-effect

According to equation 30, the sources of uncertainty are presented in the diagram cause-effect of figure 2.

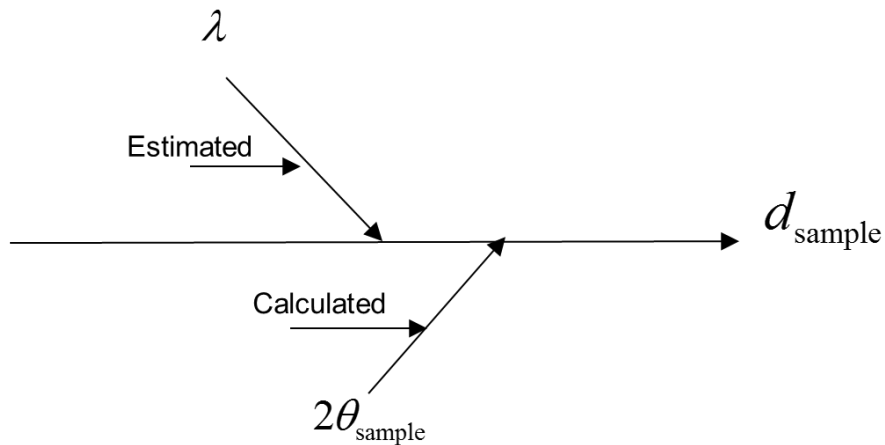


Figure 2 - Diagram cause-effect of uncertainty evaluation of interplanar distance

2.3.2 Uncertainties of input quantities

- *Wave-length*

The uncertainty of the wavelength is given by 37.

$$u_{\lambda} = \frac{a(\lambda)}{\sqrt{3}} \quad 37$$

$a(\lambda)$ = Estimation of wavelength uncertainty;

- *Angle $2\theta_{sample}$*

The uncertainty of the diffraction angle of the sample was evaluated according to equation 34.

2.3.3 Sensitivity Coefficients

From equation 36, the coefficients of sensitivity of the measurand are calculated by equations 38 and 39.

$$\frac{\partial d_{sample}}{\partial \lambda} = c_i(\lambda) = \frac{10}{2\text{sen}\left(\frac{\pi}{180} \cdot \frac{2\theta_{sample}}{2}\right)} \quad 38$$

$$\frac{\partial d_{sample}}{\partial 2\theta_{sample}} = c_i(\theta_{sample}) = \frac{\lambda \cos\left(\frac{\pi}{180} \cdot \frac{2\theta_{sample}}{2}\right)}{2\text{sen}\left(\frac{\pi}{180} \cdot \frac{2\theta_{sample}}{2}\right)^2} \cdot 10 \quad 39$$

2.3.4 Components of uncertainties

The uncertainty components of the interplanar distance are defined by expressions 40 and 41 .

$$u_{d_{sample}}(\lambda) = c_i(\lambda) \cdot u_\lambda = u_1(A^\circ) \quad 40$$

$$u_{d_{sample}}(2\theta_{sample}) = c_i(2\theta_{sample}) \cdot u_{2\theta_{sample}} \frac{\pi}{180} = u_2(A^\circ) \quad 41$$

2.3.5 Combined uncertainty

The combined uncertainty of the interplanar distance is defined by the expression 42.

$$u_c(d_{sample}) = \sqrt{u_{d_{sample}}^2(\lambda) + u_{d_{sample}}^2(2\theta_{sample})} [A^\circ] \quad 42$$

2.3.6 Number of effective degrees of freedom

The number of effective degrees of freedom is defined by equation 43.

$$v_{eff} = \frac{u_c^4(d_{sample})}{\left[u_c(2\theta_{sample}) \right]^4} \quad 43$$

Depending on the case ($C_{cal} > C_{tab}$ or $C_{cal} < C_{tab}$) the number of effective degrees of freedom of the angle is calculated according to equations 34 a or 34 b.

2.3.7 Coverage factor

The coverage factor (k) is defined according to the number of effective degrees of freedom calculated by equation 43 for a 95% coverage probability.

2.3.8 Expanded uncertainty

The expanded uncertainty is then evaluated by 44:

$$U = k \cdot u_c(d_{\text{sample}}) [A^0] \quad 44$$

3. X-ray data

The data measurements was performed in the PANalytical X'Pert PRO X-ray diffractometer with the Data Collector program, Cu $K\alpha 1$ tube and PIXcell1D detector. The integration peaks and qualitative analysis of the raw material was make through the HighScore Plus program. The diffractometer used displays calibration certificate and was tested and verified by Certified Reference Material, silicon powder NIST, whose lattice parameter in **cubic system** is $0.5431179 \text{ nm} \pm 0.000008 \text{ nm}$ at $22.5 \text{ }^\circ\text{C}$ where the expanded uncertainty stated in the MRC certificate has coverage factor (k) of 2.

3.1 DRX Calibration Curve

Table 2 shows the values indicated (2θ) with the respective $2\theta_{\text{ref}}$ of the DRX calibration certificate used in the measurement.

Table 2 - Values of $2\theta_i$ with respective values $2\theta_{\text{ref}}$. of the DRX calibration certificate

n	Angle ($^\circ$)	
	$2\theta_i$	$2\theta_{\text{ref}}$
1	28,443	28,441

2	47,300	47,300
3	56,120	56,120
4	69,130	69,126
5	76,376	76,372
6	88,028	88,025
7	94,956	94,947
8	106,707	106,701
9	114,094	114,084
10	127,545	127,534
11	136,899	136,880

From the values of $2\theta_i$ and $2\theta_{ref}$. From the table 2 characteristic values of its calibration curve are presented in table 3.

Table 3.- Characteristic values of the DRX calibration curve

Interception (a)	0,0063(°)
Slope (b)	0,999855
S^2_a	0,0000062 (°) ²
u_a	0,0025(°)
S^2_b	7,3E-10
u_b	0,000027
$r_{a,b}$	-0,93

Where:

S^2_a	Variance of the linear coefficient
u_a	Linear coefficient uncertainty
S^2_b	Variance of the angular coefficient
u_b	Angular coefficient uncertainty
$r_{a,b}$	Correlation coefficient

3.2 Validation of measurement methodology

For validation, at each of the reference angles ($2\theta_{ref}$), six sequences (K) of 10 replicates measured by the ($2\theta_{ref}$) DRX of the Certified Reference Material were performed. The values of the angles ($2\theta_{ref}$) were: 28.441 °, 47.300 °, 56.120 ° and 69.126 °. The values indicated $2\theta_i$ by the DRX for each respective $2\theta_{ref}$ were corrected by applying the XRD calibration curve. For each corrected indicated value $2\theta_i$, the error values (e_i) of the DRX were calculated.

- *Control chart of the standard deviations*

Table 4 shows the standard deviations of the DRX error values of the six sequences for the four values of $2\theta_{ref}$ angle

Table 4 - Standard deviations of the errors of the six sequences for each angle value.

$2\theta_{ref}$	Sequence					
	1	2	3	4	5	6
28,441°	8,3E-04	3,2E-03	3,2E-03	1,8E-03	1,1E-03	6,7E-04
47,300°	1,3E-03	2,3E-03	2,4E-03	2,0E-03	1,3E-03	9,5E-04
56,120°	1,8E-03	2,1E-03	1,8E-03	2,7E-03	1,8E-03	1,3E-03
69,126°	3,0E-03	1,7E-03	1,6E-03	3,0E-03	2,3E-03	2,1E-03

The graph of Figure 3 shows the control chart for the standard deviations of the measurement sequences performed by the DRX of the $2\theta_{ref}$ angles, where all the standard deviation values are contained in the range whose tolerance limit of 0,05°.

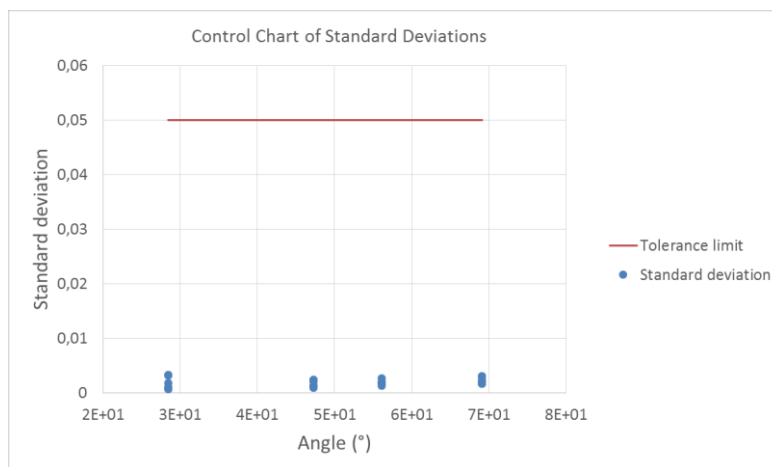


Figure 3 – Control Chart of Standard Deviations

- *Cochran Test*

To the values of variance of the standard deviation presented in table 4 the Cochran parametric test was also applied in the validation, since the test is more robust than the control chart study. Following the steps cited for application of the test, according to item 2.1.2.1, in table 5 the values for application of the Cochran test are presented.

Table 5 Cochran test values

S^2_{\max}	6,0E-06
ΣS^2_i	5,6E-05
C_{cal}	0,108
$C_{table (95\%)}$	0,151

As the C_{cal} value is lower than C_{table} , it is concluded that the variances of all the sequences studied for the DRX measurement method are statistically homogeneous for the 95% probability. The homogeneous variances are calculated as the weighted standard deviation (0.0018°), according to equation 4, of all sequences studied.

-Control chart of the averages

Table 6 shows the mean values of the DRX error values of the six sequences for the four values of $2\theta_{ref}$ angle.

Table 4 - Mean of the errors of the six sequences for each angle value.

$2\theta_{ref}$	Sequence					
	1	2	3	4	5	6
	Means					
28,441°	5,72E-03	2,41E-03	2,33E-03	4,26E-03	6,00E-03	6,22E-03
47,300°	7,76E-03	4,30E-03	4,32E-03	6,18E-03	7,59E-03	7,84E-03
56,120°	1,09E-02	7,23E-03	7,13E-03	8,58E-03	1,02E-02	1,05E-02
69,126°	1,26E-02	9,54E-03	9,25E-03	1,19E-02	1,24E-02	1,25E-02

In the graph of figure 4 the control chart of the error means of the measurement sequences made by the DRX of the $2\theta_{ref}$ angles is presented, where all the mean values are contained in the interval whose tolerance limit of 0.05° .

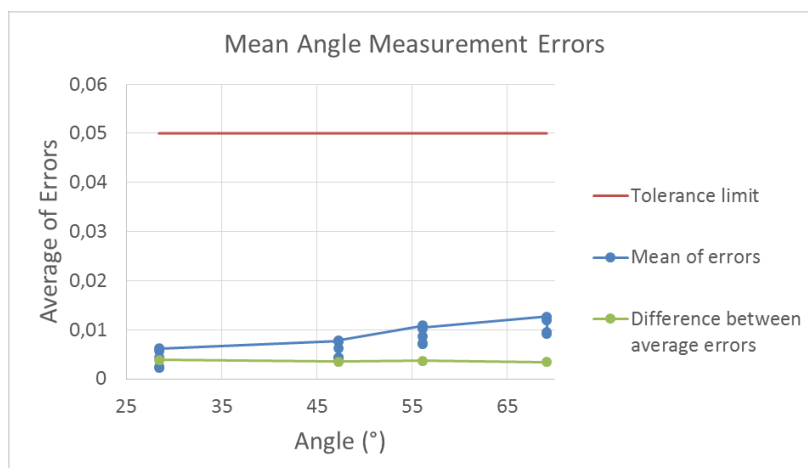


Figure 4 - Control chart of error averages

3.3 Resultados das medições da análise da amostra

The analyzed sample was in powder form and its preparation followed the recommendations "XRD Sample Holders & Preparation User's Guide, 4th Ed., PANalytical B. V". The analysis was performed by an X-ray diffractometer from manufacturer PANalytical, model X'Pert PRO MPD. The ambient conditions of the measurement were in the temperature of 23.3 °C and 60 UR. The measurements were performed under the scanning conditions mentioned in table 5.

Table 5 - Scan Conditions

Range: 2θ (°): 4,0-42,0	Slit DS (°): 1/8
Step (°): 0,026	Mask (mm): 10
Time for step (s): 150	Slit AS (mm): 6,4

Three measurements were taken at each nominal angle of analysis and the results of corrected by the calibration curve of the diffractometer are presented in table 6.

Table 6- Values of measurements performed.

No	Replicate			No	Replicate		
	2θ _{ref} (°)				2θ _{ref} (°)		
1	5,94543	5,96542	5,92743	33	22,18307	22,15308	22,19307
2	6,05741	6,05641	6,04941	34	22,34305	22,24306	22,33305
3	9,01498	8,96499	9,00498	35	22,46303	22,39304	22,46303
4	9,06697	9,06597	9,07497	36	22,59701	22,58601	22,60101
5	9,36593	9,34493	9,36493	37	23,20392	23,19792	23,19892
6	9,48091	9,45292	9,46892	38	23,60287	23,59987	23,59987
7	10,17781	10,17481	10,17581	39	24,29377	24,29776	24,29876
8	10,46877	10,46577	10,46477	40	24,56373	24,58172	24,55773
9	11,74859	11,74659	11,74759	41	25,07365	25,08865	25,09565
10	12,09953	12,10053	12,09953	42	26,24648	26,25048	26,23848
11	13,80929	13,82029	13,80929	43	27,23234	27,21234	27,20234
12	15,11110	15,10810	15,11110	44	27,38732	27,38132	27,38232
13	16,19394	16,19394	16,18394	45	27,81525	27,84225	27,83225
14	16,77586	16,76586	16,76286	46	28,18620	28,18120	28,18020
15	16,95583	16,95283	16,95183	47	28,71912	28,72812	28,71212
16	17,15180	17,14980	17,14980	48	29,06207	29,05208	29,04208
17	17,35377	17,35377	17,38377	49	30,18191	30,21191	30,18191
18	17,51375	17,54374	17,54374	50	30,50186	30,51186	30,50186
18	17,81771	17,82570	17,82270	51	30,79182	30,78182	30,79182
19	18,21265	18,20265	18,20865	52	31,76268	31,76668	31,76168
20	18,61359	18,64359	18,59359	53	32,74154	32,76154	32,75154
21	18,77357	18,75357	18,75357	54	33,12948	33,14848	33,14148
22	18,96354	18,81356	18,97354	55	33,87538	33,87438	33,86238
23	19,30349	19,31349	19,26350	56	35,00121	35,05120	35,05120
24	19,38148	19,38348	19,36748	57	36,11305	36,13105	36,11905

25	19,78142	19,78142	19,78442	58	36,81095	36,78095	36,65097
26	20,05338	20,08138	20,08238	59	37,14090	37,15090	37,12090
27	20,49332	20,50132	20,49132	60	37,51085	37,59084	37,61083
28	21,05324	20,99324	21,05324	61	38,10076	38,08076	38,08076
29	21,22621	21,22321	21,22821	62	38,42072	38,39072	38,42072
30	21,50817	21,50717	21,50817	63	39,20960	39,22360	39,20960
31	21,85812	21,84012	21,86012				
32	5,94543	5,96542	5,92743				

3.3.1 Evaluation of the uncertainty of the diffraction angle of the sample

The relative uncertainty values of each source and the combined relative uncertainty of the diffraction angle measurement of the sample are shown in table 7.

Table 7 - Relative and combined uncertainties of the sources of the diffraction angle of the sample

No	$u_{\text{certificate}}/2\theta_{\text{sample}}$	$u_{\text{curve}}/2\theta_{\text{sample}}$	$u_{\text{repetition}}/2\theta_{\text{sample}}$	$u_{\text{Method}}/2\theta_{\text{sample}}$	$u_{\text{MRC}}/\text{MRC}$	$u_{c2\theta_{\text{sample}}}/2\theta_{\text{sample}}$
1	1,94E-04	3,94E-04	1,85E-03	2,20E-05	7,36E-06	1,90E-03
2	1,91E-04	3,86E-04	4,16E-04	2,17E-05	7,36E-06	5,99E-04
3	1,28E-04	2,52E-04	1,70E-03	1,46E-05	7,36E-06	1,72E-03
4	1,27E-04	2,50E-04	3,14E-04	1,45E-05	7,36E-06	4,21E-04
5	1,23E-04	2,41E-04	7,31E-04	1,40E-05	7,36E-06	7,79E-04
6	1,22E-04	2,38E-04	8,56E-04	1,38E-05	7,36E-06	8,97E-04
7	1,13E-04	2,20E-04	8,66E-05	1,29E-05	7,36E-06	2,63E-04
8	1,10E-04	2,13E-04	1,15E-04	1,25E-05	7,36E-06	2,66E-04
9	9,83E-05	1,87E-04	4,91E-05	1,12E-05	7,36E-06	2,17E-04
10	9,54E-05	1,81E-04	2,75E-05	1,08E-05	7,36E-06	2,07E-04
11	8,36E-05	1,55E-04	2,65E-04	9,49E-06	7,36E-06	3,19E-04
12	7,64E-05	1,40E-04	6,62E-05	8,68E-06	7,36E-06	1,73E-04
13	7,13E-05	1,29E-04	2,06E-04	8,10E-06	7,36E-06	2,53E-04
14	6,89E-05	1,24E-04	2,34E-04	7,82E-06	7,36E-06	2,74E-04
15	6,81E-05	1,22E-04	7,09E-05	7,73E-06	7,36E-06	1,57E-04
16	6,73E-05	1,20E-04	3,89E-05	7,64E-06	7,36E-06	1,44E-04
17	6,65E-05	1,19E-04	5,76E-04	7,55E-06	7,36E-06	5,92E-04
18	6,59E-05	1,17E-04	5,70E-04	7,48E-06	7,36E-06	5,86E-04
19	6,48E-05	1,15E-04	1,31E-04	7,36E-06	7,36E-06	1,86E-04
20	6,34E-05	1,12E-04	1,60E-04	7,20E-06	7,36E-06	2,05E-04
21	6,20E-05	1,09E-04	7,80E-04	7,04E-06	7,36E-06	7,90E-04
22	6,16E-05	1,08E-04	3,55E-04	6,99E-06	7,36E-06	3,77E-04
23	6,10E-05	1,07E-04	2,73E-03	6,93E-06	7,36E-06	2,74E-03

24	5,98E-05	1,04E-04	7,92E-04	6,80E-06	7,36E-06	8,01E-04
25	5,96E-05	1,04E-04	2,60E-04	6,77E-06	7,36E-06	2,86E-04
26	5,84E-05	1,01E-04	5,05E-05	6,63E-06	7,36E-06	1,28E-04
27	5,75E-05	9,93E-05	4,73E-04	6,53E-06	7,36E-06	4,87E-04
28	5,63E-05	9,67E-05	1,49E-04	6,40E-06	7,36E-06	1,87E-04
29	5,49E-05	9,37E-05	9,51E-04	6,23E-06	7,36E-06	9,57E-04
30	5,44E-05	9,26E-05	6,84E-05	6,18E-06	7,36E-06	1,28E-04
31	5,37E-05	9,11E-05	1,55E-05	6,10E-06	7,36E-06	1,07E-04
32	5,28E-05	8,92E-05	2,91E-04	6,00E-06	7,36E-06	3,09E-04
33	5,21E-05	8,76E-05	5,42E-04	5,91E-06	7,36E-06	5,51E-04
34	5,18E-05	8,69E-05	1,43E-03	5,88E-06	7,36E-06	1,43E-03
35	5,15E-05	8,63E-05	1,04E-03	5,84E-06	7,36E-06	1,04E-03
36	5,11E-05	8,55E-05	1,98E-04	5,80E-06	7,36E-06	2,22E-04
37	4,98E-05	8,27E-05	8,00E-05	5,65E-06	7,36E-06	1,26E-04
38	4,89E-05	8,08E-05	4,24E-05	5,56E-06	7,36E-06	1,04E-04
39	4,75E-05	7,78E-05	6,29E-05	5,40E-06	7,36E-06	1,11E-04
40	4,70E-05	7,67E-05	2,93E-04	5,34E-06	7,36E-06	3,07E-04
41	4,60E-05	7,46E-05	2,59E-04	5,23E-06	7,36E-06	2,73E-04
42	4,40E-05	7,03E-05	1,34E-04	5,00E-06	7,36E-06	1,58E-04
43	4,24E-05	6,69E-05	3,24E-04	4,82E-06	7,36E-06	3,34E-04
44	4,22E-05	6,64E-05	6,78E-05	4,79E-06	7,36E-06	1,04E-04
45	4,15E-05	6,49E-05	2,83E-04	4,71E-06	7,36E-06	2,94E-04
46	4,10E-05	6,38E-05	6,58E-05	4,65E-06	7,36E-06	1,01E-04
47	4,02E-05	6,22E-05	1,61E-04	4,56E-06	7,36E-06	1,78E-04
48	3,97E-05	6,12E-05	1,99E-04	4,51E-06	7,36E-06	2,12E-04
49	3,82E-05	5,80E-05	3,31E-04	4,34E-06	7,36E-06	3,38E-04
50	3,79E-05	5,72E-05	1,09E-04	4,30E-06	7,36E-06	1,29E-04
51	3,75E-05	5,64E-05	1,08E-04	4,26E-06	7,36E-06	1,28E-04
52	3,64E-05	5,40E-05	4,81E-05	4,13E-06	7,36E-06	8,14E-05
53	3,53E-05	5,17E-05	1,76E-04	4,00E-06	7,36E-06	1,87E-04
54	3,48E-05	5,08E-05	1,67E-04	3,96E-06	7,36E-06	1,79E-04
55	3,41E-05	4,92E-05	1,23E-04	3,87E-06	7,36E-06	1,37E-04
56	3,30E-05	4,68E-05	4,76E-04	3,74E-06	7,36E-06	4,79E-04
57	3,20E-05	4,47E-05	1,46E-04	3,63E-06	7,36E-06	1,57E-04
58	3,14E-05	4,36E-05	1,34E-03	3,57E-06	7,36E-06	1,34E-03
59	3,11E-05	4,29E-05	2,37E-04	3,53E-06	7,36E-06	2,43E-04
60	3,07E-05	4,21E-05	8,13E-04	3,49E-06	7,36E-06	8,15E-04
61	3,03E-05	4,12E-05	1,75E-04	3,44E-06	7,36E-06	1,82E-04
62	3,01E-05	4,07E-05	2,60E-04	3,41E-06	7,36E-06	2,65E-04
63	2,94E-05	3,94E-05	1,19E-04	3,34E-06	7,36E-06	1,29E-04

The percent contributions of each source to the combined uncertainty of the diffraction angle of the sample are shown in table 8.

Table 8- Percentage contributions of each source to the combined uncertainty of the diffraction angle of the sample

No	$u_{\text{certificate}}/2\theta_{\text{sample}}$	$u_{\text{curve}}/2\theta_{\text{sample}}$	$u_{\text{repetition}}/2\theta_{\text{sample}}$	$u_{\text{Method}}/2\theta_{\text{sample}}$	$u_{\text{MRC}}/\text{MRC}$
%					
1	1,0	4,3	94,6	0,01	0,002
2	10,1	41,6	48,1	0,13	0,015
3	0,6	2,1	97,3	0,01	0,002
4	9,1	35,1	55,6	0,12	0,031
5	2,5	9,6	87,9	0,03	0,009
6	1,8	7,0	91,1	0,02	0,007
7	18,7	70,1	10,9	0,24	0,079
8	17,2	64,0	18,6	0,22	0,076
9	20,4	74,1	5,1	0,26	0,115
10	21,3	76,5	1,8	0,27	0,127
11	6,9	23,7	69,2	0,09	0,053
12	19,5	65,4	14,6	0,25	0,181
13	7,9	25,9	66,0	0,10	0,084
14	6,3	20,4	73,1	0,08	0,072
15	18,8	60,4	20,4	0,24	0,220
16	22,0	70,2	7,3	0,28	0,263
17	1,3	4,0	94,7	0,02	0,015
18	1,3	4,0	94,7	0,02	0,016
19	12,1	38,1	49,5	0,16	0,157
20	9,5	29,8	60,4	0,12	0,129
21	0,6	1,9	97,5	0,01	0,009
22	2,7	8,2	89,0	0,03	0,038
23	0,0	0,2	99,8	0,00	0,001
24	0,6	1,7	97,7	0,01	0,008
25	4,3	13,1	82,4	0,06	0,066
26	20,9	62,8	15,7	0,27	0,333
27	1,4	4,2	94,4	0,02	0,023
28	9,1	26,9	63,7	0,12	0,156
29	0,3	1,0	98,7	0,00	0,006
30	18,1	52,6	28,7	0,23	0,333
31	25,1	72,1	2,1	0,32	0,471
32	2,9	8,3	88,6	0,04	0,057
33	0,9	2,5	96,6	0,01	0,018
34	0,1	0,4	99,5	0,00	0,003
35	0,2	0,7	99,1	0,00	0,005
36	5,3	14,8	79,7	0,07	0,110
37	15,7	43,3	40,5	0,20	0,344
38	22,1	60,5	16,6	0,29	0,502
39	18,3	49,1	32,0	0,24	0,439
40	2,3	6,2	91,3	0,03	0,058

41	2,8	7,5	89,6	0,04	0,073
42	7,7	19,8	72,2	0,10	0,217
43	1,6	4,0	94,3	0,02	0,049
44	16,4	40,6	42,3	0,21	0,500
45	2,0	4,9	93,0	0,03	0,063
46	16,5	40,1	42,6	0,21	0,534
47	5,1	12,3	82,4	0,07	0,172
48	3,5	8,3	88,0	0,05	0,121
49	1,3	2,9	95,7	0,02	0,047
50	8,6	19,6	71,4	0,11	0,325
51	8,6	19,4	71,5	0,11	0,331
52	20,0	44,0	34,9	0,26	0,819
53	3,5	7,6	88,6	0,05	0,155
54	3,8	8,1	87,9	0,05	0,170
55	6,2	12,8	80,6	0,08	0,288
56	0,5	1,0	98,5	0,01	0,024
57	4,2	8,1	87,4	0,05	0,221
58	0,1	0,1	99,8	0,00	0,003
59	1,6	3,1	95,2	0,02	0,092
60	0,1	0,3	99,6	0,00	0,008
61	2,8	5,1	91,9	0,04	0,163
62	1,3	2,4	96,3	0,02	0,077
63	5,2	9,3	85,1	0,07	0,326

3.3.2 Results of measurements of diffraction angles

Table 9 presents the values of the diffraction angles of the analyzes of the studied sample with their expanded uncertainty and coverage factors.

Table 9- Results of measurements of the diffraction angles of the sample, expanded uncertainties and coverage factors

No	$2\theta_{ref}$ (°)	U (°)	k (95%)
1	5,946	0,049	4,30
2	6,0544	0,0086	2,36
3	8,995	0,067	4,30
4	9,0693	0,0098	2,57
5	9,359	0,031	4,30
6	9,468	0,037	4,30

7	10,1761	0,0057	2,12
8	10,4664	0,0059	2,13
9	11,7476	0,0054	2,12
10	12,0999	0,0053	2,13
11	13,813	0,012	2,78
12	15,1101	0,0055	2,11
13	16,191	0,011	2,78
14	16,768	0,015	3,18
15	16,9535	0,0056	2,12
16	17,1505	0,0052	2,11
17	17,364	0,044	4,30
18	17,534	0,044	4,30
19	17,8220	0,0078	2,36
20	18,2080	0,0096	2,57
21	18,617	0,063	4,30
22	18,760	0,030	4,30
23	18,9169	0,22	4,30
24	19,293	0,066	4,30
25	19,377	0,024	4,30
26	19,7824	0,0053	2,11
27	20,072	0,042	4,30
28	20,495	0,011	2,78
29	21,033	0,087	4,30
30	21,2259	0,0059	2,16
31	21,5078	0,0049	2,11
32	21,853	0,029	4,30
33	22,176	0,053	4,30
34	22,3064	0,14	4,30
35	22,4397	0,10	4,30
36	22,595	0,016	3,18
37	23,2003	0,0066	2,26
38	23,6009	0,0052	2,10
39	24,2968	0,0059	2,18
40	24,568	0,032	4,30
41	25,0860	0,029	4,30
42	26,2451	0,013	3,18
43	27,2157	0,039	4,30
44	27,3837	0,0065	2,26
45	27,8299	0,035	4,30
46	28,1825	0,0064	2,26
47	28,7198	0,022	4,30
48	29,0521	0,026	4,30
49	30,1919	0,044	4,30
50	30,5052	0,013	3,18

51	30,7885	0,013	3,18
52	31,7637	0,0056	2,18
53	32,7515	0,026	4,30
54	33,1398	0,025	4,30
55	33,8707	0,015	3,18
56	35,0345	0,072	4,30
57	36,1210	0,024	4,30
58	36,7476	0,21	4,30
59	37,1376	0,039	4,30
60	37,5708	0,13	4,30
61	38,0874	0,030	4,30
62	38,4107	0,044	4,30
63	39,2143	0,022	4,30

The graph of figure 5 shows the behavior of the expanded uncertainty of the diffraction angle according to all measurement points of the sample analysis.

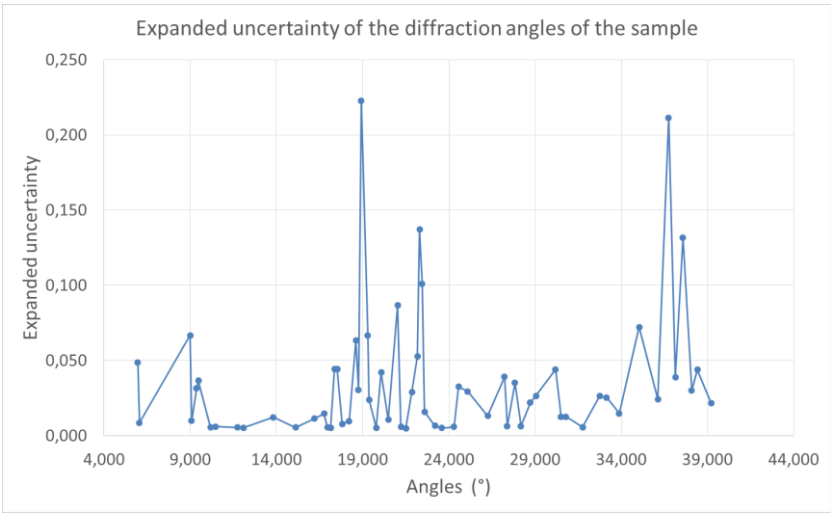


Figure 5 - Expanded uncertainty of the diffraction angle of the sample

3.3.3 Uncertainty evaluation of the interplanar distance of the sample

The values of the uncertainties of each source and the combined uncertainty of measuring the interplanar distance of the sample are shown in table 10.

No	$u\lambda$ (Å)	$u2\theta$ (radian)	$c_i\lambda$ (radian ⁻¹)	$c_i2\theta$ (Å/radian)	$u_{dsample}(\lambda)$ [Å]	$u_{dsample}(2\theta)$ [Å]	$u_c(d_{sample})$ [Å]
1	5,77E-05	1,97E-04	9,640	-286,0	5,57E-04	5,63E-02	5,63E-02
2	5,77E-05	6,33E-05	9,468	-275,8	5,47E-04	1,75E-02	1,75E-02
3	5,77E-05	2,70E-04	6,376	-124,9	3,68E-04	3,37E-02	3,37E-02
4	5,77E-05	6,67E-05	6,324	-122,8	3,65E-04	8,19E-03	8,20E-03
5	5,77E-05	1,27E-04	6,129	-115,4	3,54E-04	1,47E-02	1,47E-02
6	5,77E-05	1,48E-04	6,059	-112,7	3,50E-04	1,67E-02	1,67E-02
7	5,77E-05	4,66E-05	5,638	-97,5	3,25E-04	4,55E-03	4,56E-03
8	5,77E-05	4,86E-05	5,482	-92,2	3,16E-04	4,48E-03	4,50E-03
9	5,77E-05	4,45E-05	4,886	-73,2	2,82E-04	3,26E-03	3,27E-03
10	5,77E-05	4,36E-05	4,744	-69,0	2,74E-04	3,01E-03	3,02E-03
11	5,77E-05	7,69E-05	4,158	-52,9	2,40E-04	4,06E-03	4,07E-03
12	5,77E-05	4,56E-05	3,803	-44,2	2,20E-04	2,01E-03	2,03E-03
13	5,77E-05	7,16E-05	3,551	-38,5	2,05E-04	2,75E-03	2,76E-03
14	5,77E-05	8,01E-05	3,429	-35,8	1,98E-04	2,87E-03	2,88E-03
15	5,77E-05	4,64E-05	3,392	-35,1	1,96E-04	1,63E-03	1,64E-03
16	5,77E-05	4,30E-05	3,353	-34,3	1,94E-04	1,47E-03	1,48E-03
17	5,77E-05	1,79E-04	3,312	-33,4	1,91E-04	5,99E-03	5,99E-03
18	5,77E-05	1,79E-04	3,281	-32,8	1,89E-04	5,88E-03	5,88E-03
19	5,77E-05	5,78E-05	3,228	-31,7	1,86E-04	1,83E-03	1,84E-03
20	5,77E-05	6,52E-05	3,160	-30,4	1,82E-04	1,98E-03	1,99E-03
21	5,77E-05	2,57E-04	3,091	-29,1	1,78E-04	7,46E-03	7,46E-03
22	5,77E-05	1,23E-04	3,068	-28,6	1,77E-04	3,53E-03	3,53E-03
23	5,77E-05	9,04E-04	3,043	-28,1	1,76E-04	2,54E-02	2,54E-02
24	5,77E-05	2,70E-04	2,984	-27,0	1,72E-04	7,29E-03	7,29E-03
25	5,77E-05	9,67E-05	2,971	-26,8	1,72E-04	2,59E-03	2,60E-03
26	5,77E-05	4,40E-05	2,911	-25,7	1,68E-04	1,13E-03	1,14E-03
27	5,77E-05	1,71E-04	2,869	-25,0	1,66E-04	4,26E-03	4,27E-03
28	5,77E-05	6,67E-05	2,811	-24,0	1,62E-04	1,60E-03	1,61E-03
29	5,77E-05	3,51E-04	2,739	-22,7	1,58E-04	7,99E-03	7,99E-03
30	5,77E-05	4,72E-05	2,715	-22,3	1,57E-04	1,05E-03	1,07E-03
31	5,77E-05	4,02E-05	2,680	-21,7	1,55E-04	8,73E-04	8,87E-04
32	5,77E-05	1,18E-04	2,638	-21,1	1,52E-04	2,48E-03	2,49E-03
33	5,77E-05	2,13E-04	2,600	-20,4	1,50E-04	4,36E-03	4,36E-03
34	5,77E-05	5,56E-04	2,585	-20,2	1,49E-04	1,12E-02	1,12E-02
35	5,77E-05	4,09E-04	2,570	-20,0	1,48E-04	8,16E-03	8,16E-03
36	5,77E-05	8,76E-05	2,552	-19,7	1,47E-04	1,72E-03	1,73E-03
37	5,77E-05	5,08E-05	2,487	-18,7	1,44E-04	9,48E-04	9,59E-04
38	5,77E-05	4,27E-05	2,445	-18,0	1,41E-04	7,70E-04	7,83E-04
39	5,77E-05	4,70E-05	2,376	-17,0	1,37E-04	8,00E-04	8,11E-04
40	5,77E-05	1,32E-04	2,350	-16,6	1,36E-04	2,19E-03	2,19E-03
41	5,77E-05	1,20E-04	2,302	-15,9	1,33E-04	1,91E-03	1,91E-03
42	5,77E-05	7,24E-05	2,202	-14,6	1,27E-04	1,05E-03	1,06E-03
43	5,77E-05	1,58E-04	2,125	-13,5	1,23E-04	2,14E-03	2,15E-03

44	5,77E-05	4,97E-05	2,112	-13,4	1,22E-04	6,63E-04	6,75E-04
45	5,77E-05	1,43E-04	2,079	-12,9	1,20E-04	1,84E-03	1,85E-03
46	5,77E-05	4,95E-05	2,054	-12,6	1,19E-04	6,23E-04	6,34E-04
47	5,77E-05	8,89E-05	2,016	-12,1	1,16E-04	1,08E-03	1,09E-03
48	5,77E-05	1,07E-04	1,993	-11,9	1,15E-04	1,27E-03	1,28E-03
49	5,77E-05	1,78E-04	1,920	-11,0	1,11E-04	1,96E-03	1,96E-03
50	5,77E-05	6,87E-05	1,901	-10,7	1,10E-04	7,38E-04	7,46E-04
51	5,77E-05	6,87E-05	1,884	-10,5	1,09E-04	7,24E-04	7,32E-04
52	5,77E-05	4,49E-05	1,827	-9,9	1,05E-04	4,44E-04	4,57E-04
53	5,77E-05	1,07E-04	1,773	-9,3	1,02E-04	9,94E-04	9,99E-04
54	5,77E-05	1,03E-04	1,753	-9,1	1,01E-04	9,37E-04	9,42E-04
55	5,77E-05	8,10E-05	1,716	-8,7	9,91E-05	7,04E-04	7,11E-04
56	5,77E-05	2,93E-04	1,661	-8,1	9,59E-05	2,37E-03	2,38E-03
57	5,77E-05	9,86E-05	1,613	-7,6	9,31E-05	7,52E-04	7,57E-04
58	5,77E-05	8,57E-04	1,586	-7,4	9,16E-05	6,31E-03	6,31E-03
59	5,77E-05	1,58E-04	1,570	-7,2	9,07E-05	1,14E-03	1,14E-03
60	5,77E-05	5,34E-04	1,553	-7,0	8,96E-05	3,76E-03	3,76E-03
61	5,77E-05	1,21E-04	1,532	-6,8	8,85E-05	8,29E-04	8,34E-04
62	5,77E-05	1,78E-04	1,520	-6,7	8,78E-05	1,19E-03	1,20E-03
63	5,77E-05	8,81E-05	1,490	-6,4	8,60E-05	5,68E-04	5,74E-04

The percentage contributions of each source to the combined uncertainty of measuring the interplanar distance of the sample are shown in table 11.

Table 11 - Percent contributions from each source to the combined uncertainty of the interplanar distance of the sample

No	$ud_{\text{sample}}(\lambda)$	$ud_{\text{sample}}(2\theta)$
%		
1	0,01	99,99
2	0,10	99,90
3	0,01	99,99
4	0,20	99,80
5	0,06	99,94
6	0,04	99,96
7	0,51	99,49
8	0,50	99,50
9	0,74	99,26
10	0,82	99,18
11	0,35	99,65
12	1,17	98,83

13	0,55	99,45
14	0,47	99,53
15	1,43	98,57
16	1,70	98,30
17	0,10	99,90
18	0,10	99,90
19	1,02	98,98
20	0,84	99,16
21	0,06	99,94
22	0,25	99,75
23	0,00	100,00
24	0,06	99,94
25	0,44	99,56
26	2,16	97,84
27	0,15	99,85
28	1,02	98,98
29	0,04	99,96
30	2,16	97,84
31	3,04	96,96
32	0,38	99,62
33	0,12	99,88
34	0,02	99,98
35	0,03	99,97
36	0,73	99,27
37	2,24	97,76
38	3,25	96,75
39	2,86	97,14
40	0,38	99,62
41	0,48	99,52
42	1,44	98,56
43	0,33	99,67
44	3,27	96,73
45	0,42	99,58
46	3,49	96,51
47	1,15	98,85
48	0,81	99,19
49	0,32	99,68
50	2,16	97,84
51	2,21	97,79
52	5,33	94,67
53	1,05	98,95
54	1,15	98,85
55	1,94	98,06
56	0,16	99,84

57	1,51	98,49
58	0,02	99,98
59	0,63	99,37
60	0,06	99,94
61	1,13	98,87
62	0,54	99,46
63	2,24	97,76

3.3.4 Results of measurements of interplanar distances

Table 12 presents the values of the interplanar distances of the analyzes of the studied sample with their expanded uncertainty and coverage factors.

Table 12- Interplanar distances values, expanded uncertainty and coverage factor of the sample analysis

No	d_{sample} [Å]	U[Å]	k(95%)
1	14,85	0,24	4,303
2	14,586	0,041	2,365
3	9,82	0,15	4,303
4	9,743	0,021	2,571
5	9,442	0,063	4,303
6	9,334	0,072	4,303
7	8,6856	0,0097	2,120
8	8,4453	0,0096	2,131
9	7,5270	0,0069	2,120
10	7,3087	0,0064	2,131
11	6,406	0,011	2,776
12	5,8587	0,0043	2,110
13	5,4701	0,0077	2,776
14	5,2829	0,0092	3,182
15	5,2256	0,0035	2,120
16	5,1660	0,0031	2,110
17	5,103	0,026	4,303
18	5,054	0,025	4,303
19	4,9729	0,0044	2,365
20	4,8683	0,0051	2,571
21	4,762	0,032	4,303
22	4,726	0,015	4,303
23	4,69	0,11	4,303
24	4,597	0,031	4,303
25	4,577	0,011	4,303

26	4,4843	0,0024	2,110
27	4,420	0,018	4,303
28	4,3299	0,0045	2,776
29	4,220	0,034	4,303
30	4,1825	0,0023	2,160
31	4,1283	0,0019	2,101
32	4,064	0,011	4,303
33	4,005	0,019	4,303
34	3,982	0,048	4,303
35	3,959	0,035	4,303
36	3,9321	0,0055	3,182
37	3,8308	0,0022	2,262
38	3,7667	0,0016	2,093
39	3,6603	0,0018	2,179
40	3,6206	0,0094	4,303
41	3,5469	0,0082	4,303
42	3,3929	0,0034	3,182
43	3,2740	0,0092	4,303
44	3,2543	0,0015	2,262
45	3,2031	0,0079	4,303
46	3,1639	0,0014	2,262
47	3,1059	0,0047	4,303
48	3,0711	0,0055	4,303
49	2,9577	0,0084	4,303
50	2,9280	0,0024	3,182
51	2,9017	0,0023	3,182
52	2,8149	0,00099	2,160
53	2,7322	0,0043	4,303
54	2,7010	0,0041	4,303
55	2,6444	0,0023	3,182
56	2,559	0,010	4,303
57	2,4847	0,0033	4,303
58	2,4437	0,027	4,303
59	2,4189	0,0049	4,303
60	2,3920	0,016	4,303
61	2,3608	0,0036	4,303
62	2,3416	0,0052	4,303
63	2,2955	0,0025	4,303

The graph of figure 6 shows the behavior of the expanded uncertainty of the interplanar distance according to all measurement points of the sample analysis.

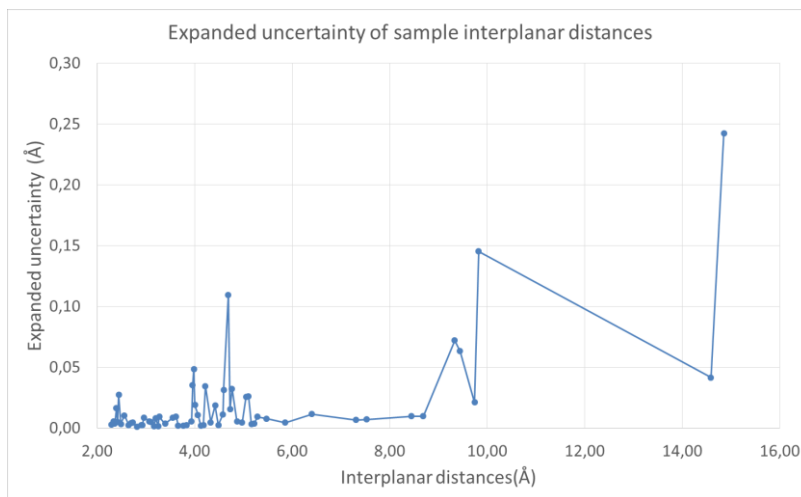


Figure 6 - Expanded uncertainty of the interplanar distances of the sample

CONCLUSIONS

In applying the control chart to validate the standard deviations of the diffraction angle errors of the measurements made by the DRX, it is concluded that the values of the dispersions of all the sequences are homogeneous because they are all within the tolerance range adopted from 0.05° . Additionally it was concluded that the tolerance of 0.05° is oversized in relation to the dispersion values of all measurement sequences, ranging from 15 to 75 times greater than the values of the calculated standard deviations.

Applying the Crochran test for the validation of the standard deviations of the DRX errors, it was also concluded that the values of the dispersions of all the sequences are homogeneous, giving the weighted standard deviation of 0.0018° . In this case, the value of the weighted standard deviation induces that the initially stipulated tolerance may be less than 0.05° .

The mean error values of the diffraction angles of the measurements performed by the DRX are also homogeneous with each other, because they are all within the tolerance range of 0.05° . The difference between the mean error values of each sequence was approximately 0.004° , which also induces, in the case studied, which tolerance initially adopted (0.05°) can be revised.

In the combined uncertainty of all the values of the diffraction angle of the studied sample, in most measured points, the uncertainty regarding the repeatability of the measurements is the one that most impacts, reaching the maximum value of 99.8%.

In the measurements made in the study sample of this article, in some points of measurement of the diffraction angle the uncertainty source referring to the calibration curve of the DRX has greater impact and in some cases the impact is in the order of approximately 76%.

In the sample studied, the uncertainty regarding the validation of the methodology was at most about 0.3% of the combined uncertainty.

The uncertainty from the Certified Reference Material used in the measurement showed the maximum impact of 0.8% on the combined uncertainty, where it can be concluded that the MRC could present a greater uncertainty that will in a way reduce the cost of the analysis.

The expanded uncertainty of the diffraction angle of the sample studied ranged from 0.0049° to 0.22° .

In the combined uncertainty of all the values of the interplanar distance of the studied sample, the uncertainty regarding the diffraction angle is the one that most impacts, reaching the general maximum value of approximately 100%.

In the case studied, due to the low impact of wavelength uncertainty on the combined uncertainty of the interplanar distance, it is suggested that the length can be measured with a worse accuracy

The expanded uncertainty of the interplanar distance of the sample studied ranges from 0.0010 \AA to 0.24 \AA .

Annex

Critical Values for Cochran's C at the 5% significance level

K	$n_i=2$	$n_i=3$	$n_i=4$	$n_i=5$	$n_i=6$
2	-	0,975	0,939	0,906	0,877
3	0,967	0,871	0,798	0,746	0,707
4	0,906	0,768	0,684	0,629	0,59
5	0,841	0,684	0,598	0,544	0,506
6	0,781	0,616	0,532	0,48	0,445
7	0,727	0,561	0,48	0,431	0,397
8	0,68	0,516	0,438	0,391	0,36
9	0,638	0,478	0,403	0,358	0,329
10	0,602	0,445	0,373	0,331	0,303
11	0,57	0,417	0,348	0,308	0,281
12	0,541	0,392	0,326	0,288	0,262
13	0,515	0,371	0,307	0,271	0,243
14	0,492	0,352	0,291	0,255	0,223
15	0,471	0,335	0,276	0,242	0,22
16	0,452	0,319	0,262	0,23	0,208
17	0,434	0,305	0,25	0,219	0,198
18	0,418	0,293	0,24	0,209	0,189
19	0,403	0,281	0,23	0,2	0,181
20	0,389	0,27	0,22	0,192	0,174
21	0,377	0,261	0,212	0,185	0,167
22	0,365	0,252	0,204	0,178	0,16
23	0,354	0,243	0,197	0,172	0,155
24	0,343	0,235	0,191	0,166	0,149
25	0,334	0,228	0,185	0,16	0,144
26	0,325	0,221	0,179	0,155	0,14
27	0,316	0,215	0,173	0,15	0,135
28	0,308	0,209	0,168	0,146	0,131
29	0,3	0,203	0,164	0,142	0,127
30	0,293	0,198	0,159	0,138	0,124

Source [18]

REFERENCES:

1. ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. Requisitos gerais para a competência de laboratórios de ensaio e calibração. Rio de Janeiro: ABNT, 2005. 31p. (ABNT ISO/IEC 17025:2005)
2. Couto, Paulo Roberto Guimarães Estimativa da incerteza da massa específica da gasolina pelo ISO GUM 95 e método de Monte Carlo e seu impacto na transferência de custódia. Rio de Janeiro, 2006. xxi,122 f. il. color. Dissertação (Mestrado em Tecnologia de Processos Químicos e Bioquímicos) – Universidade Federal do Rio de Janeiro, Escola de Química, 2006.
3. JCGM 100:2008 GUM 1995 with minor corrections - Evaluation of measurement data — Guide to the expression of uncertainty in measurement - First edition September 2008 - JCGM 2008
4. Food and Drug Administration Amendments Act (FDAAA) of 2007
5. Haleblian, J. (1975). *J. Pharm. Sci.* 64, p.1269-1288
6. McCrone, W. C. (1965). *Physics and Chemistry of the Organic Solid State*, edited by Fox, D., Labes, M. M. & Weissberger, A., v.2, pp.725–767. New York: Wiley Interscience.
7. Cooper, V. B., Pearce, G. E. S., Petts, C. R. (2003). *J. Pharm. Pharmacol.*, 55, 1323–1329.
8. Araújo, G. L. B., Pitaluga Jr., A., Antonio, S. G., Paiva-Santos, C. O., Matos, J. R. (2012). *Rev. Ciênc. Farm. Básica Apl.* 33, 27–36.
9. Brittain, H. G. (1999). Methods for the characterization of polymorphs and solvates. In: *Polymorphism in Pharmaceutical Solids*, pp. 277-278. New York: Marcel Dekker, Inc.
10. USP (2016). *United States Pharmacopeia 39th Edition, National Formulary 34 <941> Characterization of crystalline and partially crystalline solids by x-ray powder diffraction (XRPD)*. p.755-760. Berti, G. (2001). *Powder Diff.* 16, 6-15.
11. Taupin, D. (1973). *J. Appl. Cryst.* 6, 266-271.
12. Huang, T. C. (1988). *Aust. J. Phys.* 41, 201-212.

13. Schwarzenbach, D. (1995). *Acta Cryst.* A51, 565-569.
14. Rebmann, H. R. et al. (1998). *Acta Cryst.* **A54**, 225-231.
15. Toraya, H. (2001). *J. Appl. Cryst.* **34**, 558-565.
16. Withers, P. J., Daymond, M. R., Johnson, M. W. (2001) *J. Appl. Cryst.* **34**, 737-743.
17. Natrella, M.G. (1963) *Experimental Statistics*. National Bureau of Standards Handbook 91. Issued August 1, 1963
18. Massart, D.L.; Vandeginste, B. M. G.; Buydens, L. M. C.; Jong, S.; Lewis, P. J.; Smeyers-Verbeke, J. *Handbook of Chemometrics and Qualimetrics. Part A. Volume 20A*. Elsevier Science B.V. 1997.867p

APÊNDICE B

CENTRO TÉCNICO CIENTÍFICO
DEPARTAMENTO DE ENGENHARIA INDUSTRIAL
MEMORANDO TÉCNICO

MEMORANDO TÉCNICO No. 02/18

**TOWARDS A SYSTEMATIC METHODOLOGY TO ASSOCIATE
RESEARCH-PROMOTING ENTITIES WITH INDUSTRIAL
SECTORS: THE NANOTECHNOLOGY CASE**

**Ronaldo Pedro da Silva
Fellipe Silva Martins
Daniel Giacometti Amaral
Antonio Marcio Tavares Thomé
Eugenio Kahn Epprecht
Adriano Koshiyama**

31 de Janeiro de 2018

PUC

Pontifícia Universidade Católica do Rio de Janeiro
Rua Marquês de São Vicente, 225 - CEP 22451-900
Rio de Janeiro - Brasil

Towards a Systematic Methodology to Associate Research-Promoting Entities with Industrial Sectors: the Nanotechnology Case

Ronaldo Pedro da Silva^{1,2}, Felipe Silva Martins³, Daniel Giacometti Amaral⁴,
Antônio Márcio Tavares Thomé², Eugenio Epprecht², Adriano Koshiyama^{5*}

¹Campus Duque de Caxias, Professor Geraldo Cidade – Federal University of Rio de Janeiro
27 Xerem Road, Duque de Caxias, 25245-390
Rio de Janeiro, Brazil

²Department of Industrial Engineering – Pontifical Catholic University of Rio de Janeiro
225 Marquês de São Vicente Street, Gávea, 22430-060
Rio de Janeiro, Brazil

³Nove de Julho University – Management Graduate School - Industrial Engineering Graduate School
61 Francisco Matarazzo Avenue, Agua Branca, 05001-000
São Paulo, Brazil

⁴NanoBusiness Information and Innovation Ltda - Xerem Technology Park
50 Nossa Senhora das Gracas Avenue, Duque de Caxias, 25250-020
Rio de Janeiro, Brazil

⁵Department of Computer Science – University College London
Gower Street, WC1E 6BT, (Telephone: +44 (0) 20 7679 2000)
London, United Kingdom

*corresponding author

ronaldo@xerem.ufjf.br, silvamartinsfelipe@gmail.com, daniel.giacometti@nanobusiness.com.br,
eke@puc-rio.br, mt@puc-rio.br, a.koshiyama@cs.ucl.ac.uk

Abstract

In this paper, we propose a method, based on Fuzzy Relations to define the degree of relevance or affinity of an RPE – Research-promoting Entities (universities, research institutes, national laboratories and public and private companies) – to each productive sector (e.g., Oil & Gas, Dentistry). We applied the method over 129 Brazilian RPEs that produced the larger number of articles in the area of nanotechnology during the period from 2000 to 2014. Our task was to associate 112 knowledge areas that such publications can be identified with one or more of the 12 productive sectors. The results were analysed using the Fuzzy C-means method. These results contribute to the decision-making process of entrepreneurs, researchers and public managers, for the search of partnerships, investment, training and use of human capital, among others.

Keywords: cluster analysis, fuzzy relations, nanotechnology, university-company interaction, public-private partnership.

1. Introduction

Scientific fields such as traditional disciplines (e.g., mathematics, physics, biology) still operate mainly on Bohr's quadrant (scientific rigour without the need of immediate, guided applicable relevance) (Stokes, 1997). Innovation, on the other hand, pushes market and research entities together in what is called Pasteur's quadrant (when research is at the same time rigorous and application-g geared). As such, collaboration may entail in rigorous science that is not entirely devoid of practical, productive intent. However, contrary to social sciences (e.g., Management) which already live in this bi-dimensional world and are used to "serve several masters" (Tushman & O'Reilly, 2007), the so-called *hard* sciences usually develop their own institutional nests, speak a different language from the market, and, finally, may suffer from failing to communicate with it.

In one hand, in developed countries, this linkage between scientific knowledge generation and innovation suffers from the same bottlenecks, but is mitigated by the fact that research-promoting entities (RPE – universities, research institutes, national laboratories, and so forth) and market entities are continuously 'rubbing shoulders'. Although the language they speak is different, there is a partial permeability of ideas, concepts and scientific discovery with eventual transitions to innovation, which is usually stimulated by national or local innovation systems that put them together. In developing countries, on the other hand, scientific research fails to become innovations because this 'edge' between these two important points in the triple helix (Etzkowitz & Leydesdorff, 2000) is largely inexistent and innovation systems are underdeveloped. A close consequence is that developing countries' research is mostly absorbed by developed countries (through scientific absorption or straight-to-market absorption) (Krauskopf et al., 2007).

Either way, in both developed and developing countries, the difficulty of characterising how closely or associated is an RPE with the potential entities in the productive sector and deciding in investing in its fit persists. This difficulty occurs because the publications generated by the researchers are usually categorised by knowledge area (e.g., Engineering, Physics, and so forth) and not by productive application sector (e.g., Oil & Gas, Dentistry, among others). Therefore, this asymmetry in the categorization process hinders the association between these entities.

As such, we propose a method to define the degree of affinity or membership of a given RPE to each productive sector. This methodology aims to collaborate in the decision-making process of entrepreneurs, researchers and public managers, for the search of partnerships, training and use of human capital, among others. Our proposal is underpinned by a technique called fuzzy relations. A fuzzy relation (Klir & Yuan, 1995; Pedrycz & Gomide, 1998) represents a degree of presence or absence of association, interaction or interconnection between elements of two or more sets. To test the method, we employed data from the 129 Brazilian RPEs which produced more scientific papers in nanotechnology from 2000 to 2014, associated with 112 knowledge areas and linked to 12 productive sectors. Publication data was extracted from the *Web of Science* (WoS) database based on a modular keyword strategy proposed by Arora et al. (2012). The knowledge areas considered in this study were based on Web of Science Research Areas. The results allow understanding that there is a natural grouping of RPEs, according to their specialisations. Broader specialisations occur due to many reasons, such as proximity to excellence centres and market entities (industries). A second reason is the clustering of RPEs in areas with high investment in research. The main causes for stricter specialisations may be linked to specific human and technical assets in specific RPEs.

The results also allow managing portfolios of which RPEs produce more research on each specific knowledge area, compare their development and organise the efforts of industries according to the academic production. As such, we take another step in bridging the gaps in innovation fostering the communication and translation between the needs of its academic and market entities.

This work is organised as follows: next section presents a broader literature review covering areas from RPEs and industry collaboration, translating logics and the very few available solutions to associate R&D dedicated entities with industry partners. The third section outlines our method, starting with a brief description of what elements we used in the association process and the final output handled to the decision-maker. Then, we show how the estimation process for the different components is met, as well as a brief description of the Fuzzy C-Means method. Section 4 presents a study case in the Nanotechnology area in Brazil, in which we exhibit our main findings by applying the proposed methodology. Finally, we close this paper with some concluding remarks.

2. Literature review

2.1. RPEs and industry collaboration

The goals or purposes of science are usually desired for their sake. That is, science is traditionally set in what is known as Bohr's quadrant – development of knowledge devoid of obligation for any practical and commercial purposes other than the development of knowledge itself (Stokes, 1997). As such, what guides research and researchers is mainly the will to understand phenomena, commonly devoid of commercial purposes. However, since commercial innovation is one the most important drivers for investments in scientific development, the transition between pure science and its innovations are increasingly being sought.

As much as the positive results from such partnerships are desired, the actual relationships between the innovation entities are problematic. Universities fear collaboration with industry because they are usually the weaker link - that is, industries enter the partnership with funding but also high expectations for actual problem solving, while universities may not manage to deliver (Lee, 1996). Industries may also fear collaboration because scientific discoveries follow different paths - it goes where the questions arise, not where industry-specific problems are located. As such, long-term partnerships are difficult to maintain (Lee, 2000). Researchers, on the other hand, are the key to significant science-to-market transition, if they have experience in both worlds (Siegel, Waldman, Atwater & Link, 2003; Perkmann & Walsh, 2008; Lam, 2010). Most researchers, however, must conform to legitimacy pre requisites in university evaluation and employment systems that do not emphasise technical and industrial development, and, as such, these activities are left aside (Lam, 2007; D'este & Perkmann, 2010; Tartari & Breschi, 2012). Nevertheless, the benefits of partnering are also countless.

Some scientific fields such as social sciences have been dealing with this 'serving two masters' situation for decades and, as such, developed a basic symbiotic relationship with the organisations and social phenomena they study, although these relationships are not perfect (McDonald, 2017). Other hard sciences, such as pharmaceutical research, have also had strong stimulation from their environment to conform to innovation (innovation-gear science).

More than that, akin to human-like entities, they diverge in mindset, behaviour and interpretation of the environment. Consequently, many opportunities to ship scientific discoveries to innovation development fail in the process, not because they are commercially useless, but because there is currently no deliberate ‘translation’ in the institutional logics, whose transitions mainly depend on adapted processes. By adapted processes, we mean that institutional arrangements, such as the development of local and national innovation systems, work to implement coordination between RPEs, governmental bodies and economic entities, but such arrangements as of now are not sufficient. The university-industry-government interactions are often described by the concept of Triple Helix systems, a conceptual framework for analyzing the dynamics of innovation within the context of global economic change. Each helix is an independent institutional sphere working in cooperation with the others in an interdependent innovation system of constant knowledge flow (Etzkowitz & Leydesdorff, 2000). The idea underneath innovation systems is that by putting these three elements together, innovation will arise.

Examples that coordination between entities in innovation systems improve the development of innovation abound. However, examples of failure (Orsenigo, 2001; Edquist, 2011) demonstrate that innovation systems work not necessarily because they are optimally managed but because they physically set RPEs and industries together (Morgan, 2004; Laursen, Masciarelli & Prencipe, 2012), so that by ‘rubbing shoulders’ at least some of the scientific developments find permeable ways to transition to commercial innovations (Rodríguez-Pose & Crezcenzi, 2008).

As such, the problem is that coordination between them do not usually include ‘translating’ one’s needs and understanding to others. A keen example of the need for translating logics in innovation systems is the European ‘pre-commercial procurement’ (PCP) program. It started as a way of fostering scientific innovation with fast tracking to industries, but quickly became a program to foster supplying late-stage, applied science to commercial innovation (Edquist & Zabala-Iturriagoitia, 2014).

2.2. Translating logics

This translation paradox is not new and appears long before the concept of the triple helix. The first instance of the problem of adaptation between entities appears in the mutual adjustment concept (Lindblom, 1965). The idea is that coordination between entities is difficult to achieve because each entity is “one who makes decisions calculated to serve his own goals, not goals presumably shared by all other decision-makers with whom he is interdependent” (Lindblom, 1965:29). Lindblom, however, suggests resolving this standoff by choosing between two strategies for innovation coordination – “synopsis” (a rational or scientific approach to hierarchically defining needs) or “centrality” (centralised coordination of the network) (Lindblom, 1965; White, 1965).

The problem associated with the first is that a scientific approach puts more emphasis on the ‘research’ point of the triple helix, whose ‘language’, goals and protocols are not shared by other entities. The problem associated with the former is that centralised processes tend to be dominated by governmental entities because of their power to rule and their lack of understanding of the goals and need of others takes its toll. It may also suffer from “dogmatism” and lack of flexibility to deal with institutional changes in the constituent parts of the triple helix (White, 1965). Therefore, the role of ‘ruling’ associated with national coordination suffers from a social lag, typical of laws been passed, in which the innovations in the innovation system itself become official much later than they are treated as *de facto* by partnerships (Magro, Navarro & Zabala-Iturriagoitia, 2014). This fact provides evidence over the inefficacy of governments to

be the central role in innovation. However, the role of governments in bringing entities together and fostering innovation (if not centralising) is paramount. This problem is better developed in the work of Willke (2002), who defined democratic intelligence, i.e., the inter-coordination among entities in a dynamic way to resolve such issues, “including intermediate forms of social associations in democratic decision-making” (Willke, 2007, p. 167). In the case of a decision-making problem, such as putting RPEs and market entities together, the overall, traditional centralised mechanisms do not work effectively, but by mediating the efforts of experts, it is possible to translate the needs of the market to RPEs and vice-versa. The role of the expert is so central that Willke defines that “the legitimacy of decisions comes to depend on expertise” (p. 44). As such, only experts whose experience and resources are in both worlds can develop a significant framework of intelligence to coordinate efforts in both academia and market needs.

2.3. Linking research to industries

As much as setting industries around RPEs has its advantages, the translating of industry needs to research capabilities is something that has been scarcely researched. Most of the theories deal with coordination between triple helix entities, mediated by experts, but the actual mechanisms of finding which RPEs are most adequately paired to which productive sectors remain a challenge.

As such, frameworks such as the ones that develop how innovations may enter academia (Siguaw, Simpson & Enz, 2006), how universities should be open to innovations (Perkmann & Walsh, 2007), and how stakeholders in industries need to understand their own needs before looking for research partnerships (Succar, 2008) abound. However, frameworks and mechanisms that bridge the gap between industries needs and which institutions produce what they need are largely non-existent.

In this sense, a challenge persists – it is not a straightforward path, but strategies to attain it need to be broken down in two (Figure 1).

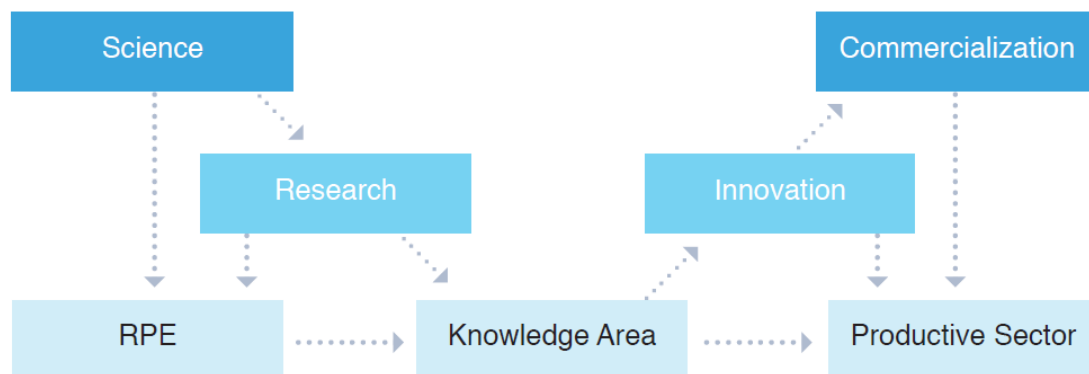


Figure 1. Microfoundational explanation for the matrices generation

That is, from research to innovation, a common ground (Knowledge area) is necessary to be defined as a bridge. As such, one perceives that the path from science to final commercialization is a macro approach, whose explanation can be better found in a micro foundational approach (Foss & Lindenberg, 2013).

As indicated above, several studies have approached the linkage between RPEs and industry with most of the research focused on understanding the complexity of the

knowledge transfer process between triple helix entities and the ex-post evaluation of interactions. Besides providing insights into key aspects of workings of these linkages, research in this area has also been used to provide various metrics, especially for university rankings, indicating that industry interaction has become of paramount importance for evaluating the performance of RPEs (Tijssen, 2012).

With research focused on an ex-post perspective, attention has been drawn to the drivers and factors behind the success of the knowledge transfer process. Among the most relevant variables related to this achievement, the concept of ‘proximity’ is usually pointed as central, assuming different shapes and forms such as: geographical (e.g. physical proximity); cognitive (e.g. knowledge base and expertise); social (e.g. friendship, shared experience, behavioral codes, common culture, mutual trust); institutional (e.g. similarities in terms of institutional frameworks and shared organizational arrangements); and organizational (e.g. degree of autonomy of partners) (Boschma, 2005).

Although this approach can provide an interesting framework of analysis for the linkages between research and industry, it only covers part of the complexity of creating a successful connection between RPEs and the industry sector since it assumes that all the actors are known to each other. The question that arises from this assumption is: how to find which RPEs are more adequately paired to a specific industry without limiting the analysis to previously known entities?

Despite literature in this area is largely nonexistent, some efforts have already been made in similar challenges as linking technological areas from patent classifications to industrial sectors (Schmoch et al., 2003). The underlying problem is finding a correlation between research areas or technological fields and industrial or productive sectors considering the existing asymmetry regarding categorization.

To fulfil this gap, this article approaches the correlation between knowledge areas (as scientific production is classified) and productive sectors based on a fuzzy relationships method (Klir & Yuan, 1995; Pedrycz & Gomide, 1998). While similar applications of fuzzy relationships have been identified for other topical areas (Guifrida & Nagi, 1998), the problem of defining degrees of affinity or pertinence of a given RPE to each productive sector has not yet been addressed.

Ghazinoory & Farnoodi (2013) have recently proposed a method based on text mining to select technology centres of excellence. Regardless of having a different focus, this work approaches the challenge of better understanding the competencies of a particular RPE using specific quantitative criteria, using information retrieved from scientific and patent international databases as well as applying fuzzy relations techniques for it. In this particular, next section presents our proposed methodology.

3. Methodology

3.1. RPEs association via Fuzzy Relations

The method proposed in this article makes use of two matrices of relevance – matrix R (RPE x Knowledge Area) and matrix A (Knowledge Area x Productive Sector) – whose information will enable you to obtain the matrix of relevance S (RPE x Productive Sector). Therefore, consider the existence of n areas of knowledge ($i = 1, \dots, n$) and K RPEs ($k = 1, \dots, K$). Let r_{ki} be the degree of affinity between the k -th RPE (universities and research centres) and the i -th knowledge area (e.g., Materials Science, Mechanical Engineering).

Each r_{ki} composes the membership matrix R , that is, the affinity matrix between RPEs and Knowledge Area:

$$R = \begin{bmatrix} r_{11} & \cdots & r_{1n} \\ \vdots & \ddots & \vdots \\ r_{K1} & \cdots & r_{Kn} \end{bmatrix}_{K \times n}, \text{ such that } r_{ki} \in [0,1]$$

Let a_{ij} be the degree of association between the knowledge area i and the productive sector j ($j = 1, \dots, J$). The membership matrix A , that is, the matrix of association between Knowledge Area and Productive Sector is given by:

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1J} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nJ} \end{bmatrix}_{n \times J}, \text{ such that } a_{ij} \in [0,1]$$

From the composition between the matrices R and A , the membership matrix S is obtained, so that each position represents the degree of affinity between the k -th RPE and the j -th productive sector:

$$S = \begin{bmatrix} s_{11} & \cdots & s_{1J} \\ \vdots & \ddots & \vdots \\ s_{K1} & \cdots & s_{KJ} \end{bmatrix}_{K \times J}, \text{ such that } s_{kj} \in [0,1]$$

Where $S = R \circ A$, where the operator \circ represents the max-min composition (Klir & Yuan, 1995), and therefore, each $s_{kj} = \max(\min(r_{k1}, a_{1j}), \dots, \min(r_{kn}, a_{nj}))$. Therefore, the decision-maker can evaluate the degree of affinity between the RPEs and the most diverse sectors, and, therefore, to form partnerships with those that have more relation with their final activity (Automotive, Agribusiness, and so forth).

However, it is necessary to establish objective criteria with respect to the estimation of the entries of the matrices of pertinence R and A , that is, the values r_{ki} and a_{ij} . The following sections discuss in more detail the established methodology for computing such quantities.

3.1.1. Estimation of the elements of RPE x knowledge area matrix

This section presents methods to estimate the items of the relevance matrix between RPE and each Knowledge Area. The original problem is to determine the degree of association between the k -th RPEs (example: Minas Gerais State Federal University, Bahia State Federal University) and the knowledge area i (example: Instrumentation, Solid-state Chemistry), which will be denoted by r_{ki} . This degree of association takes values between 0 and 1, so that the more r_{ki} is close to 1, the greater the association between Research Entities k with knowledge area i .

To do so, consider $\#P_{k \Rightarrow i}$ the number of publications of the k -th RPE in the i -th knowledge area, while $\#P_k$ represents the total number of publications of the k -th RPE. To compute these quantities, the publications of researchers linked to this entity should be retrieved, considering the publication period from 2000 to 2014, categorized by knowledge area, as is quite common in the academy. From these definitions, we propose the following estimator for r_{ki} :

$$\hat{r}_{ki} = \frac{\#P_{k \Rightarrow i}}{\#P_k} \in [0,1]$$

where \hat{r}_{ki} is the frequency of publications of the k-th RPE in the knowledge area i . The more the RPE publishes in a given knowledge area, the greater will be its affinity with it. After computing the quantities present in the membership matrix R , it becomes necessary to compute the elements of the membership matrix A to finalize the process of determining the membership matrix S .

3.1.2. Estimation of the elements of Knowledge area x Productive Sector matrix

This section presents a mechanism to estimate the elements of the pertinence matrix between Knowledge Area and Production Sector. The original problem is to determine the degree of association between the knowledge area i (example: Materials Engineering, Electrical Engineering) and the productive sector j (example: Oil & Gas, Agribusiness), which will be denoted by a_{ij} . This degree of association takes values between 0 and 1, so that the more a_{ij} is close to 1, the greater the degree of association between knowledge area i and sector j .

There are two ways to estimate a_{ij} . First, an objective approach, in which one verifies the number of publications in the knowledge area i that shares the same application with, or proposes it as a field of implementation of the productive sector j . Denote by $\#P_{i \Rightarrow j}$ the number of publications associating knowledge area i with sector j , whereas $\#P_i$ is the number of publications surveyed in the knowledge area i . Therefore, the estimator is:

$$\hat{a}_{ij} = \frac{\#P_{i \Rightarrow j}}{\#P_i} \in [0,1]$$

It is possible to see that \hat{a}_{ij} is the relative frequency of publications that associate knowledge area i with sector j . Therefore, if the entire spectrum of publications is researched a_{ij} converges to the probability of applying the publication in the knowledge area i in the productive sector j . Note that in this case relative frequency is being considered as an association between the knowledge area and the productive sector.

A second approach would be a subjective one, in which we propose an expert committee, whose combined views and experience can form an estimate for a_{ij} . The specialist will therefore serve as a bridge between the knowledge created in the knowledge area i and its potential use in the productive sector j . Consider for $x_{i \Rightarrow j}^{(d)}$ the opinion of the d -th specialist ($d = 1, \dots, D$) on the association between knowledge area i and productive sector j . We suggest the following estimator for a_{ij} :

$$\ddot{a}_{ij} = \sum_{d=1}^D w^{(d)} x_{i \Rightarrow j}^{(d)} \in [0,1], \text{ com } w^{(d)} \geq 0 \text{ e } \sum_{d=1}^D w^{(d)} = 1$$

where $w^{(d)}$ represents the weight or importance of the specialist d in the body of opinions of the specialists (defined from a greater performance in the area, titration, publications, among others). If all the specialists have equal relevance, we can choose $w^{(d)} = \frac{1}{D}$ for all d , that is, we have that \ddot{a}_{ij} is the arithmetic mean of each specialist's opinions.

Another point of analysis is how to collect the information $x_{i \Rightarrow j}^{(d)}$ of each specialist. It is easy to see how costly and ambiguous to ask the specialist a value between 0 and 1 for $x_{i \Rightarrow j}^{(d)}$. As this procedure is not feasible, the Likert Scale procedure was used, commonly used in psychometric studies (interest in products, degree of affinity between ideas, among others). A Likert item is just a statement to which the respondent responds with the level of agreement or non-agreement to the statement. Five levels of responses are usually used, although some researchers prefer to use seven or even nine levels. The typical format of a Likert item is:

- I absolutely do not agree – $x_{i \Rightarrow j}^{(d)} = 0$
- I partially do not agree – $x_{i \Rightarrow j}^{(d)} = 0,25$
- I am indifferent – $x_{i \Rightarrow j}^{(d)} = 0,50$
- I partially agree – $x_{i \Rightarrow j}^{(d)} = 0,75$
- I completely agree – $x_{i \Rightarrow j}^{(d)} = 1,00$

Where the statement made to the specialist is: "The Knowledge Area has a full affinity with this Productive Sector." From this statement, the specialist assigns his level of agreement. If this does not fully agree, we assign $x_{i \Rightarrow j}^{(d)}$ to the null value, that is, degree of affinity 0, otherwise a positive degree of affinity is established at $x_{i \Rightarrow j}^{(d)}$. After presenting all the statements (in all are n for each j -th productive sector) to each of the experts, the estimator presented previously is computed. This process is performed for all productive sectors, so that matrix A is formed from the elements \ddot{a}_{ij} .

Finally, it is emphasized that the ideal approach would be to merge objective and subjective information since it would add both a diagnosis about the past and present (objective form) and new boundaries (subjective perception) in the relationship between the knowledge areas and the productive sectors. However, the objective approach is complex to perform normally, given the number of publications in the various areas of knowledge and productive sectors encompassed in this research. With a committee of experts, the subjective estimation path will be used to determine the values that make up the matrix of relation Knowledge Area x Productive Sector. Next section provides the results and discussions of the presented methodology.

3.2. Fuzzy C-means Algorithm

Fuzzy C-means algorithm (Bezdek and Pal, 1992) works pretty similarly with the traditional k-means method, but instead of providing hard/crispy partitions it generates soft partitions – that is, overlapping sets. Hence, instead of an element belonging to a single group, it pertains to many groups but with different membership degrees. More specifically, consider a database with N objects $X = \{x_1, \dots, x_N\}$, the final result is a fuzzy partition matrix of these objects in a given number of groups k . This partition matrix is represented by the matrix U ($k \times N$) whose elements U_{ij} represent the membership degree of the j -th object to the i -th fuzzy group.

In general, C-Means algorithm seeks to minimise some variation or extension of the following objective function:

$$\text{Min} : J(v_j) = \sum_{i=1}^N \sum_{j=1}^k (U_{ij})^m D_{ij}^2$$

s.a.

$$\sum_{j=1}^k U_{ij} = 1$$

$$1 \leq i \leq N$$

The exponent m is a fuzzification parameter (usually equal to 2), D_{ij} is a given distance between the j -th object and the i -th group and J is a measure of intra-group dissimilarity. The choice of the number of groups is subjective and problem-specific, but it is possible to take into account grouping quality metrics (Sum of Quadratic Deviations, Centrality and Silhouette).

4. Case Study: Nanotechnology area in Brazil

4.1. Experimental Setup

We selected 129 Brazilian RPEs that most produced articles in the area of nanotechnology during the period from 2000 to 2014. We present their total number of publications during this time in each of the 112 areas of knowledge in Appendix A. We have set a minimum volume threshold of 7 works to be included in the listing. It should be noted that the lower the number of publications of an RPE, the more challenging and uncertain be to categorise an RPE in a productive sector, or even infer about its predominant knowledge area. However, this low value was chosen so that a larger spectrum of Brazilian RPEs were considered in the study.

The group of experts was composed of researchers with proven performance in some of the 112 areas of knowledge. This group forms a diverse committee of opinions, with some of these more specialised for a certain set of areas of knowledge (related to Medicine, Materials Engineering), while others are more transversal (Chemistry, Physics). All were given the same weight for the composition of the Matrix of Relation Area of Knowledge x Productive Sector.

Before presenting the results showing the degree of association of each RPE to the 12 productive sectors considered (Automobile, Electronic, Cosmetics, Textiles, Paints, Coatings, Aerospace, Dentistry, Oil and Gas, Agribusiness, Environment and Medicine), some exploratory analyses were performed considering the information presented. These analyses were divided into two stages:

- Clustering of the RPEs based on total publications – comparing its volume of works in a knowledge area to others;
- Clustering of the RPEs based on its focus on a specific area – comparing its frequency of publications in an field to others RPEs.

The first analysis aims to identify the RPEs that dominate some subset of the 112 areas of knowledge. The dominant ones tended to group, leaving the other set of RPEs with less dominance. This analysis is necessary because of the large number of areas of knowledge, which makes an area-of-knowledge analysis costly. The second analysis aims

to identify RPEs that are more transversal, that is, have publications in a significant number of areas of knowledge, or even those that focus on certain areas. This information is relevant since an RPE with a high level of transversality will tend to have high affinity with several sectors.

It is worth emphasising again two points of view about the methodology used:

- The estimator used to measure the compatibility of an RPE with a knowledge area only uses information of the RPE itself. If an RPE has a degree of relevance superior to another in a particular sector, this does not mean that it has a comparative advantage concerning others. A degree of importance of an RPE to one sector must be compared to those obtained in other sectors. Therefore, if RPE X has a higher degree of pertinence in the Oil and Gas Sector than in Dentistry, this means that it has a greater focus in this area, and not that it has more competence than a lesser degree of pertinence. Hence, the ordering must be performed in an absolute and non-comparative perspective. The smaller the number of publications, the more difficult is the process of assessing the compatibility of an RPE with a knowledge area. This difficulty causes a greater risk in the modelling that follows this, requiring a post-processing step to check for potential outliers.

4.2. Exploratory Analysis

The first analysis focuses on the volume of papers published by each RPE per knowledge area. Through the results of the cluster analysis by C-means, one observes the formation of two groups: RPEs with a high number of publications in some areas of knowledge (14 RPEs in 129 - group 2), while those in group 1 (115 RPEs) are characterized by relatively less publication in some knowledge areas than group 2. The RPEs belonging to group 2 are: USP (3694), UNICAMP (2117), UNESP (1513), UFSCar (1373), UFRGS (1081), UFMG (998), UFRJ (993), UFPE (567), UFSC (522), UFPR (517), PUC-Rio (278), UNIFESP (224), FIOCRUZ (71) and FURG (36), where values in parentheses are the total of publications of these RPEs during the study period.

There are seven major RPEs regarding publication (USP, UNICAMP, UNESP, UFSCar, UFRGS, UFMG and UFRJ), as well as others that have a high number of publications but are not the first in rank (UFPE, UFSC, UFPR, PUC-Rio, UNIFESP, FIOCRUZ and FURG). These have comparative advantages to other universities in some areas of knowledge in which the largest RPEs do not have, for example, PUC-Rio in the field of Computer Science and Mathematics, FIOCRUZ in Public Health and FURG in Zoology and Marine Animals.

Figure 2 shows a visualisation for the groups formed by the C-means algorithm, using the Principal Components Analysis method (Johnson and Wichern, 2002). Given the projection in two components of the information utilised by the grouping method, one has the visualisation of the elements of groups 1 and 2. Each point in the graph represents a university. It is possible to notice that the elements of group 1 (red dots) are quite close, forming a cohesive nucleus, while the elements of group 2 (dots in blue) distance themselves from this group. The most distant point (approximately 0.5 in the 1st component and 1.5 in the 2nd component of scores) is USP, due to its a high number of publications (3694), almost double that of UNICAMP (second largest regarding Number of publications). Finally, the RPEs that were closer to the centroid of each group were CETENE (group 1) and UFSC (group 2). It means that if a more detailed study of the reasons for this asymmetry is necessary (possibly due to the number of teachers, courses,

students), it would be very relevant to select these two RPEs and others similar to these since they represent the profile closer to each group.

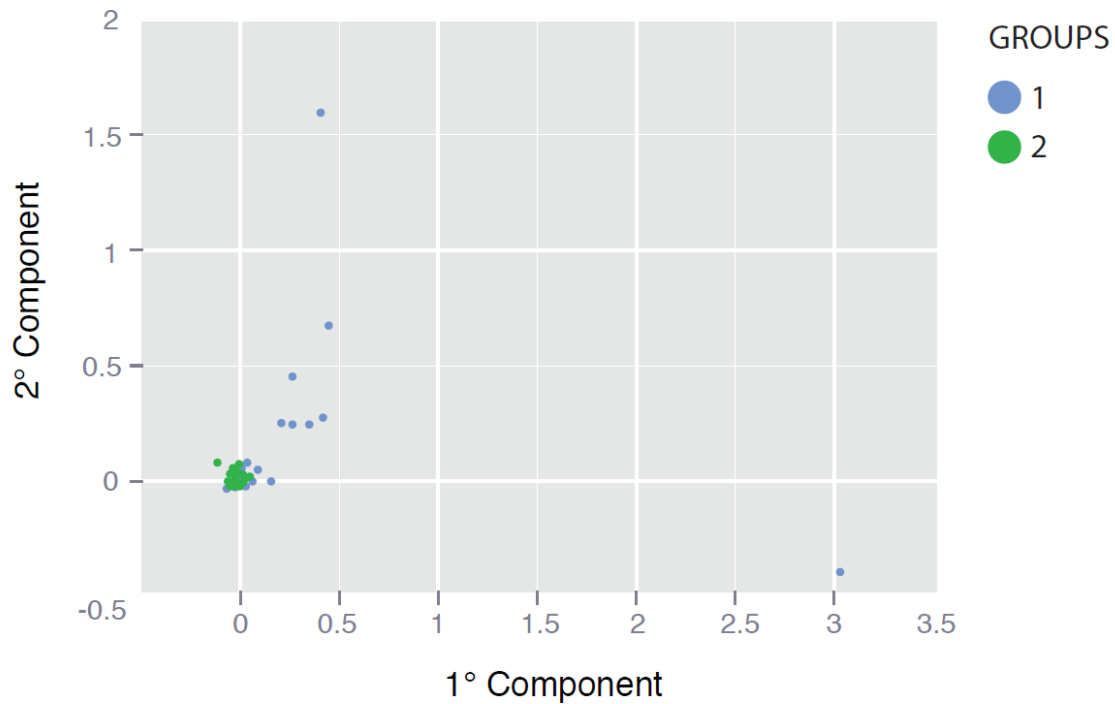


Figure 2. Projection in two components by Analysis of Principal Components, considering the total number of publications by Area of Knowledge of each RPE.

As a second analysis, the RPEs were grouped according to their frequency of publication, considering the 112 knowledge areas, considering frequency of publication, in addition to the results of the grouping. It is possible to observe the formation of three groups:

- **Group 1** (Focuses on some areas): This group is formed by 37% of the listed RPEs, having as its profile representing RPE UFRGS. This group aims at the publication with greater prevalence in the areas of Physics, Materials Science, Chemistry, Science and Technology and Engineering. In this category belong the seven largest RPEs in terms of publication, to say: USP, UNICAMP, UNESP, UFSCar, UFRGS, UFMG and UFRJ.
- **Group 2** (Transversal): This group is formed by RPEs that have a balance in the volume of their publications, being composed of 38% of the analysed ones. The typical RPE of this group is UFSC. In this group are located the average RPEs in term of publication, such as UFSC, UNIFESP, EMBRAPA, IPEN, among others.
- **Group 3** (Focus on 1 area): This group is formed by 25% of the RPEs analysed in the study, having as a typical element the UNB. In general, this group is formed by RPEs that are more closely related to physics, such as UNB, LNLS, CBPF, PUC-Rio, to quote a few.

Figure 3 illustrates the arrangement of the elements of each group, using a lower dimensionality projection via Principal Component Analysis. We also point out the prototype RPEs of each of the group. Unlike Figure 2, the projection does not present the groups arranged cohesively; such issue is the result of the projection method used - this enables an approximate solution to reduce the original dimension to 2 dimensions, which, of course, is difficulty done without loss.

EXPERT SYSTEMS WITH APPLICATIONS

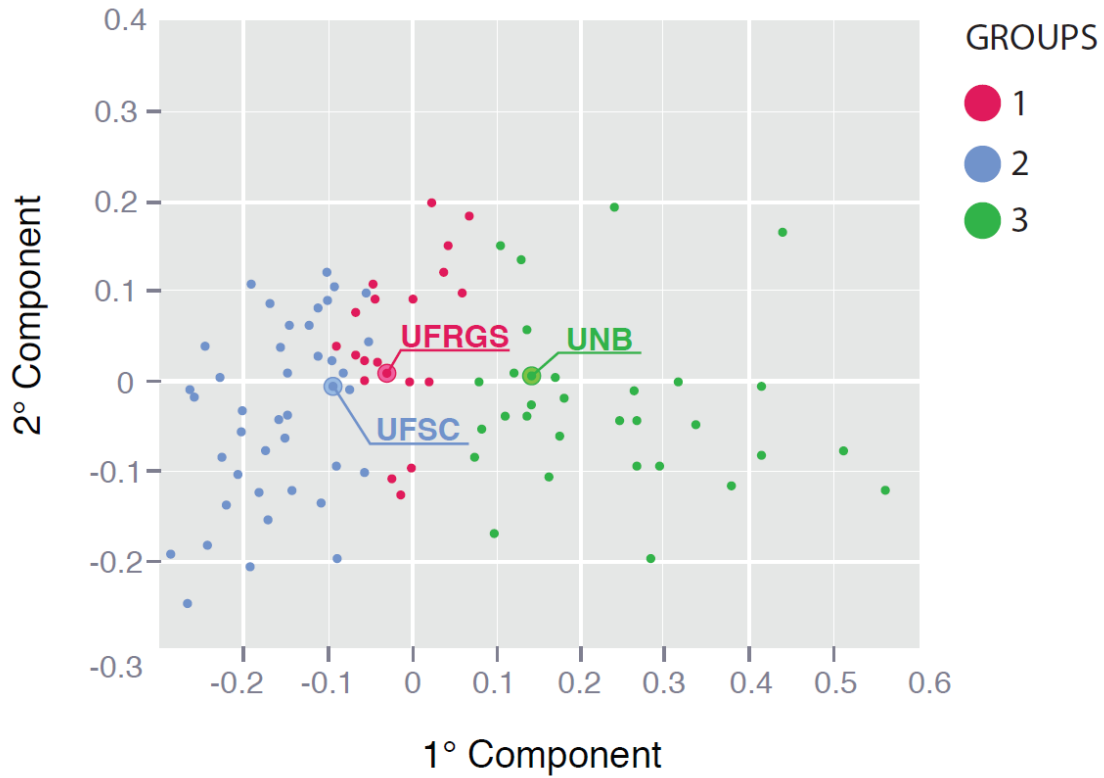


Figure 3. Projection in two components by Analysis of Principal Components, of the Publication Proportions in each Knowledge Area by RPE.

Finally, the analysis was performed for each RPE per sector, according to the previously presented methodology. With the results obtained for each analysed RPE, as well as the cluster analysis by C-means, three groups were formed, whose centroids are shown in Table 1.

EXPERT SYSTEMS WITH APPLICATIONS

Table 1. Centrality of the groups formed by the membership degrees of each RPE to the productive sectors under study.

GROUPS	AUTOPARTS	ELETRONIC	PHPC	TEXTILE
1	0,98	0,92	0,65	0,80
2	0,96	0,87	0,84	0,80
3	0,66	0,58	0,86	0,60
GROUPS	AEROSPACE	DENTISTRY	PAINTS	COATINGS
1	0,94	0,90	0,52	0,93
2	0,85	0,84	0,69	0,93
3	0,60	0,77	0,55	0,65
GROUPS	OIL & GAS	AGRIBUSINESS	ENVIRONMENT	MEDICINE
1	0,96	0,66	0,59	0,86
2	0,97	0,83	0,78	0,89
3	0,70	0,69	0,71	0,94

Following there are some descriptions of the profiles of each group, observed from the information described in Table 1.

- **Group 1** (Focus on Some Sectors): It is formed by 36 of the 129 RPEs analysed. The characteristic element of this group is the LNLS, which is the closest to the centroid of the group. This group has a greater affinity, that is, the intense membership degree to the sectors of Electronics, Automobile and Aerospace. However, it also operates, to a lesser extent, in the Oil and Gas, Coatings and Dentistry sectors.
- **Group 2** (Transverse): This cohort consists of 78 of the 129 RPEs under study. It is characterised by greater transversality among the productive sectors considered because although it does not always have the highest degree of pertinence in all sectors, these have a membership degree that is now the largest in the sector or the second largest. The typical RPEs of this sector, i.e., the closest to the centroid was EMU.
- **Group 3** (Little Affinity with the Productive Sector or Focus on Medicine): This group is the smallest of all, formed by only 15 RPEs. The closest element to the centroid is FIOCRUZ. This group is characterised by the low affinity to the productive sectors, being the sectors of minor relevance the Electronic and Paints. However, in the sectors of Environment and Medicine, this sector has a high degree of importance to these sectors. Table 2 presents the membership degree of the RPEs closest to the centroids of each group to the productive sectors, while Figure 4 shows the projection, via Principal Components, of this result.

EXPERT SYSTEMS WITH APPLICATIONS

Table 2. RPEs closest to the centroids of each group and their degree of pertinence to each productive sector considered in the study.

RPE	AUTOPARTS	ELETRONIC	PHPC	TEXTILE
LNLS – 1	1,00	0,99	0,66	0,84
UEM – 2	0,97	0,84	0,88	0,82
FIOCRUZ – 3	0,58	0,54	0,88	0,56
RPE	AEROSPACE	DENTISTRY	PAINTS	COATINGS
LNLS – 1	0,54	0,96	0,95	0,91
UEM – 2	0,69	0,94	0,85	0,88
FIOCRUZ – 3	0,58	0,60	0,56	0,80
RPE	OIL & GAS	AGRIBUSINESS	ENVIRONMENT	MEDICINE
LNLS – 1	0,98	0,68	0,58	0,89
UEM – 2	1,00	0,85	0,80	0,89
FIOCRUZ – 3	0,64	0,78	0,78	1,00

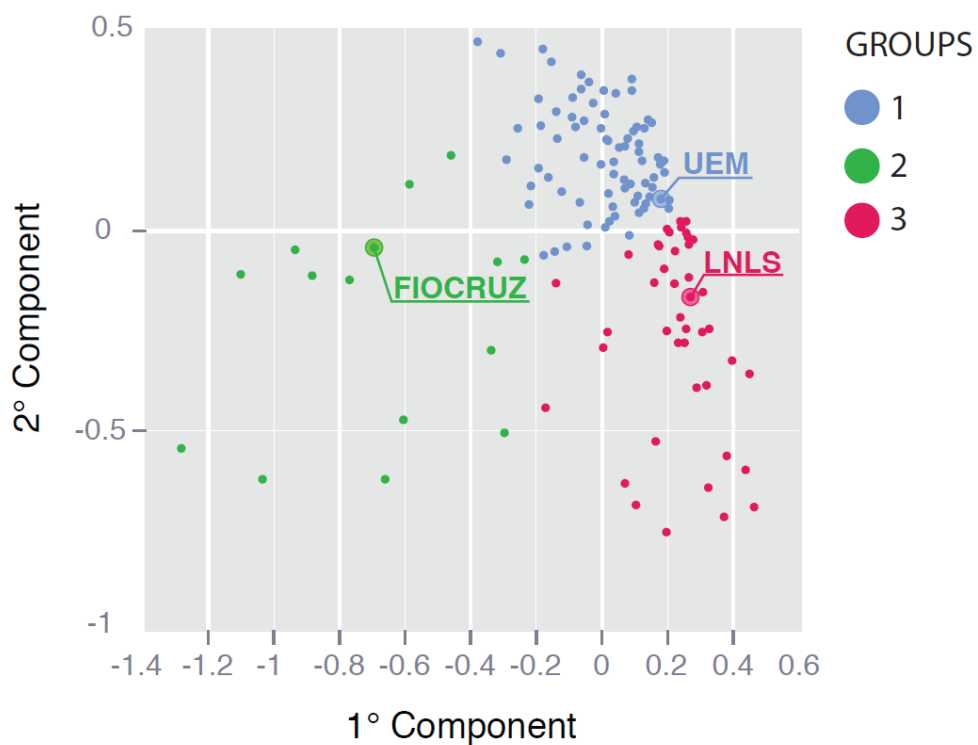


Figure 4. Projection in two components by Analysis of Principal Components, considering the degree of relevance of each RPE to the productive sectors.

4.3. Modelling Results

The following are the ten most relevant RPEs in each productive sector under analysis, from Table 3 to 5. Other modelled results not presented are available at Appendix A. We present in detail two sectors: automotive; and In general, the automotive sector is one of those with the highest number of RPEs with a relevance degree of 1 (fully associated) with this sector, making a total of 57 RPEs. This feature is due to the high transversality that the automotive sector has, since publications in the area of applied physics, materials science, engineering, chemistry, metallurgy, robotics, among others, have some degree of applicability for the production of parts, simulators, new materials and coatings, among others.

Table 3 presents the ten RPEs with the highest degree of relevance to the Automotive Sector, given the tie-breaking criterion of number of publications (values located in parentheses). It is observed the presence of five universities in the State of São Paulo (USP, UNICAMP, UNESP, UFSCar and UFABC), whose greater presence is justified by the association of courses in mechanical engineering, control and automation and mechatronics, as well as the state highly present automotive industry in São Paulo, mainly concentrated in the region known as ABC Paulista. For example, USP has the Automotive Engineering Center (CEA), responsible for the training of masters and doctors in the sector, as well as research and development projects with large companies, such as the study on diesel engines for SCANIA. Another case of a partnership is the UFPE with Fiat Chrysler, forming the Center for Research, Development, Innovation and Automotive Engineering in Recife. The centre is the second of the automaker in the country (another is located in Betim-MG) and the fourth in the world. Also, it has a partnership with the University of Turin for the training of automotive engineers, via visiting students exchange program.

EXPERT SYSTEMS WITH APPLICATIONS

Table 3. Ten RPEs with greater degree of relevance to the Automotive Sector, given the tie-breaking criterion of number of publications.

RPE (# Publications)	Relevance degree <i>Automobile</i>
USP (3694)	1,00
UNICAMP (2117)	1,00
UNESP (1513)	1,00
UFSCar (1373)	1,00
UFRGS (1081)	1,00
UFMG (998)	1,00
UFPE (567)	1,00
UFPR (517)	1,00
UFU (411)	1,00
UFABC (354)	1,00

Table 4 lists the ten RPEs with the highest degree of relevance to the Aerospace Sector, given the tie-breaking criterion of number of publications (whose value is presented in parentheses). It is possible to observe the presence of four RPEs from the State of São Paulo (UNICAMP, CTA, INPE and EEL / USP), the largest being the Aerospace Technological Center (CTA), which has the Aeronautical Technological Institute (ITA) Institute of Aeronautics and Space (IAE) and Institute of Advanced Studies (IEAV). This conglomerate of institutes enables the generation of undergraduate and postgraduate knowledge in the areas of hypersonic, aero-thermodynamics, mechanical engineering and aeronautical infrastructure, applied physics, among others, all of them focused on the Aerospace Sector. There is still lots of interaction between ITA and EMBRAER, one of the world's leading producers of commercial aircraft, executive jets, supply planes, combat aircraft and agricultural aviation.

The National Institute for Space Research (INPE) appears as a link to teaching, research and extension very similar to the ITA model. However, INPE seeks to develop solutions in the field of mapping the Brazilian territory, space information, construction and launching of satellites (and nanosatellites), among others, in addition to its partnerships with other RPEs, such as UNICAMP in the production of research And patents on georeferencing and radar mapping, for example. The other EPPs have this high degree of relevance to the aerospace sector, due mostly to graduate and post-graduate courses in mechanical engineering or control and automation (and in some cases, UNICAMP, aeronautics). These courses greatly encourage the interest of students and teachers, both in undergraduate studies (as well as the training of air development teams - PUC-Rio and UPM), as well as in graduate studies (research, Work, among others). It is also worth

EXPERT SYSTEMS WITH APPLICATIONS

noting the presence of CBPF in Experimental Physics and Low Energy research, whose applications are direct in the aerospace industry.

Table 4. Ten RPEs with higher degree of relevance to the Aerospace Sector, given the tie-breaking criterion of number of publications.

RPE (# Publications)	Aerospace relevance
UNICAMP (2117)	1,00
PUC-Rio (278)	1,00
INPE (256)	1,00
CBPF (169)	1,00
CTA – ITA/IEAV/IAE (156)	1,00
IME (32)	1,00
PUC-Goiás (24)	1,00
UPM (18)	1,00
EEL/USP (10)	1,00
UNISINOS (8)	1,00

In the dentistry sector, a high number of RPEs with degree of relevance 1 were not observed (in total there were 11 entities). Table 10 shows the ten RPEs with the highest degree of relevance to the dentistry sector, given the tie-breaking criterion of number of publications (presented in parentheses). It is possible to observe a range of universities and research centres across the country, being three of the ten belonging to the State of Rio de Janeiro. The presence of the CBPF is due to the researchers in the area of experimental physics, mainly in photoacoustic evaluations of resins, with the focus on application in the dental industry.

The publications in the areas of physics, materials science, image science, spectroscopy, microscopy and dentistry, oral surgery and medicine, provided that UFMG, UFRJ, UNB and UFF have a high degree of affinity with both the Dentistry, as well as with the medical sector. Finally, it is noteworthy that both UFMG and UFRJ have one of the oldest dentistry courses in the country.

Table 5. Ten RPEs with greater degree of relevance to the Dentistry Sector, given the tie-breaking criterion of number of publications

RPE (# Publications)	Relevance degree Dentistry
UFMG (998)	1,00
UFRJ (993)	1,00
UNB (617)	1,00
UFF (337)	1,00
CBPF (169)	1,00
UEPG (112)	1,00
UFES (96)	1,00
PUC-Goias (24)	1,00
PUC-PR (21)	1,00
UPF (15)	1,00

5. Final remarks

This work considers the importance of the role of the RPEs to improve industries' competitiveness through knowledge transfer. Previous research successfully addresses metrics (Tijssen, 2012) and drivers and factors (Boschma, 2005) that intend to analyse the complex linkage between RPEs and industry. However, there was still a gap to fulfill, regarding the need to understand the RPEs that are more adequate to a specific industry. This work proposes a method to correlate knowledge areas from RPEs, through the scientific production, and industry sectors based on fuzzy relationship methods. This contribute to previous efforts in doing so through patents (Schmoch et al., 2003; Ghazinoory & Farnoodi, 2013), as it enables to understand the competences of the RPE through objective quantitative criteria regarding a specific industry. This also contributes for practice, as it can be considered as a management tool to improve the relationships between RPEs and industry.

This work presents the methodology and the results of the test on 129 Brazilian RPEs, scientific production in nanotechnology, from 2000 to 2014. It associated 112 knowledge areas and linked to 12 productive sectors. The works in more detail the results for 4 productive sectors and the results for the remaining 8 sectors (Appendix A). The results show the grouping of RPEs considering their competences. It confirms the link of specialization and proximity to the industry (Morgan, 2004; Laursen, Masciarelli &

EXPERT SYSTEMS WITH APPLICATIONS

Prencipe, 2012; Rodríguez-Pose & Crezcenzi, 2008), areas with higher investment in research and access to human and technical assets.

As any other work, it has some limitations. Future works should test the method in other specialties and institutional contexts. For example, the Chinese context has a strong financial support for certain Universities. As well as it should be tested in developed countries.

It is proposed as future work to improve the part of the definition of degrees of affinity (to discretize the area of knowledge in more slices) and to insert uncertainties, corresponding to the standard deviation of the proportion of publications of the institution. With the objective of interpreting the results it is suggested to normalize in the horizontal direction ($\max = 1$, $\min = 0$), as well as to take into account the volume information of publications, for comparative advantages.

EXPERT SYSTEMS WITH APPLICATIONS

Appendix A: Complementary sector results

Table A1. Ten RPEs with higher degree of relevance to the Electronic Sector, given the tie-breaking criterion of number of publications.

RPE (# Publications)	Relevance degree <i>Electronics</i>
UFMG (998)	1,00
UFG (284)	1,00
PUC-Rio (278)	1,00
LNLS (340)	0,99
CBPF (169)	0,99
USP (3694)	0,98
CTA – ITA/IEAV/IAE (156)	0,98
LNNano (17)	0,98
UEFS	0,97
INPE	0,96

EXPERT SYSTEMS WITH APPLICATIONS

Table A2. Ten RPEs with greater degree of importance to the PHPC Sector, given the tie-breaking criterion of number of publications.

RPE (# Publications)	Relevance degree PHPC
UNIFESP (224)	1,00
UFOP (104)	1,00
UCB (15)	1,00
Hospital Israelita Albert Einstein (14)	1,00
UEPB (11)	1,00
Fund. Centro Tecnológico de Minas Gerais (26)	0,97
UFRGS (1081)	0,97
Instituto Federal de Minas Gerais (27)	0,97
PUC-PR (21)	0,96
UDESC (57)	0,96

EXPERT SYSTEMS WITH APPLICATIONS

Table A3. Ten RPEs with higher degree of relevance to the Textile Sector, given the tie-breaking criterion of number of publications.

RPE (# Publications)	Relevance degree Textile
UFU (411)	1,00
UFJF (140)	1,00
PUC-Goiás (24)	1,00
UNISINOS (8)	1,00
Inst. Federal do Rio de Janeiro (IFRJ) (9)	0,92
PUC-Campinas (38)	0,92
UEL (117)	0,92
UFMA (66)	0,91
UFAL (92)	0,91
UECE (29)	0,91

EXPERT SYSTEMS WITH APPLICATIONS

Table A4. Ten RPEs with greater degree of relevance to the Paints Sector, given the tie-breaking criterion of number of publications.

RPE (# Publications)	Relevance degree Paints
Braskem SA (22)	0,85
UFSC (522)	0,84
UFRJ (993)	0,83
UEMS (11)	0,83
UNIVALI (14)	0,82
UFF (337)	0,81
INT (8)	0,81
UECE (29)	0,81
UFOP (104)	0,80
UFMS (55)	0,80

EXPERT SYSTEMS WITH APPLICATIONS

Table A5. Ten RPEs with greater degree of relevance to the Coating Sector, given the tie-breaking criterion of number of publications.

RPE (# Publications)	Relevance degree Coatings
UFSCar (1373)	1,00
UFC (574)	1,00
UFPE (567)	1,00
IME (32)	1,00
FEI (31)	1,00
PUC-Goiás (24)	1,00
LNNano (17)	1,00
UNIFAP (13)	1,00
EEL/USP (10)	1,00
UFAC (9)	1,00

EXPERT SYSTEMS WITH APPLICATIONS

Table A6. Ten RPEs with greater degree of relevance to the Oil and Gas Sector, given the tie-breaking criterion of number of publications.

RPE (# Publications)	Relevance degree Oil & Gas
USP (3694)	1,00
UNICAMP (2117)	1,00
UNESP (1513)	1,00
UFSCar (1373)	1,00
UFMG (998)	1,00
UFRJ (993)	1,00
UFC (574)	1,00
UFSC (522)	1,00
UFABC (354)	1,00
UFG (284)	1,00

EXPERT SYSTEMS WITH APPLICATIONS

Table A7. Ten RPEs with greater degree of relevance to the Agribusiness Sector, given the tie-breaking criterion of number of publications.

RPE (# Publications)	Relevance degree <i>Agribusiness</i>
EMBRAPA (220)	1,00
UFV (137)	1,00
UFLA (61)	1,00
UNIOESTE (8)	0,99
UEL (117)	0,99
UFGD (11)	0,97
Instituto Nacional de Ciência e Tecnologia de Bioanalítica (19)	0,97
UNIVALI (14)	0,97
UECE (29)	0,95
UFMA (66)	0,95

EXPERT SYSTEMS WITH APPLICATIONS

Table A8. Ten RPEs with greater degree of relevance to the Environment Sector, given the tie-breaking criterion of number of publications.

RPE (# Publications)	Relevance degree Environment
UDESC (57)	1,00
UNIFRAN (36)	1,00
Universidade Tiradentes (20)	1,00
IPT (16)	1,00
UNESC (15)	1,00
UNIVALI (14)	1,00
UFGD (11)	1,00
FURG (36)	0,98
Instituto Nacional de Ciência e Tecnologia de Bioanalítica (19)	0,97
UNIOESTE (8)	0,94

EXPERT SYSTEMS WITH APPLICATIONS

Table A9. Ten RPEs with higher degree of relevance to the Medicine Sector, given the tie-breaking criterion of number of publications.

RPE (# Publications)	Relevance degree <i>Medicine</i>
UFRJ (993)	1,00
UFSM (242)	1,00
UNIFESP (224)	1,00
EMBRAPA (220)	1,00
UFPB (203)	1,00
INMETRO (139)	1,00
UFV (137)	1,00
UEL (117)	1,00
UERJ (115)	1,00
UFPeI (98)	1,00

References

Arora, S.; Porter, A.; Youtie, J.; Shapira, P (2012). Capturing new developments in an emerging technology: an updated search strategy for identifying nanotechnology research outputs. *Scientometrics*, 95(1), 351-370.

Bochma, R. (2015) Towards an evolutionary perspective on regional resilience. *Regional Studies*, 49(5), 733–51.

D’Este, P., & Perkmann, M. (2011). Why do academics engage with industry? The entrepreneurial university and individual motivations. *The Journal of Technology Transfer*, 36(3), 316-339.

Edquist, C. (2011). Design of innovation policy through diagnostic analysis: identification of systemic problems (or failures). *Industrial and Corporate Change*, 11, 1-29.

Edquist, C., & Zabala-Iturriagoitia, J. M. (2015). Pre-commercial procurement: a demand or supply policy instrument in relation to innovation? *R&D Management*, 45(2), 147-160.

Etzkowitz, H., & Leydesdorff, L. (2000). The dynamics of innovation: from National Systems and “Mode 2” to a Triple Helix of university–industry–government relations. *Research policy*, 29(2), 109-123.

Guiffrida A., & Nagi R. (1998). Fuzzy set theory applications in production management research: a literature survey. *Journal of Intelligent Manufacturing*, 9, 39-56.

Hohmann, L. (2016). To what Extent Is the Triple-Helix-Model of Etzkowitz & Leydesdorff of Use for the Implementation of Smart Governance? – an Analysis Referring on Implemented Triple Helix-Constellations. *Glocality*, 2(1), 2pp.

Klir, George; Yuan, Bo. Fuzzy sets and fuzzy logic. New Jersey: Prentice Hall, 1995.

Krauskopf, M., Krauskopf, E., & Méndez, B. (2007). Low awareness of the link between science and innovation affects public policies in developing countries: The Chilean case. *Scientometrics*, 72(1), 93-103.

Lam, A. (2007). Knowledge networks and careers: Academic scientists in industry–university links. *Journal of management studies*, 44(6), 993-1016.

EXPERT SYSTEMS WITH APPLICATIONS

Lam, A. (2010). From 'ivory tower traditionalists' to 'entrepreneurial scientists'? Academic scientists in fuzzy university-industry boundaries. *Social Studies of Science*, 40(2), 307-340.

Laursen, K., Masciarelli, F., & Prencipe, A. (2012). Regions matter: how localized social capital affects innovation and external knowledge acquisition. *Organization Science*, 23(1), 177-193.

Lee, Y. S. (1996). 'Technology transfer' and the research university: a search for the boundaries of university-industry collaboration. *Research policy*, 25(6), 843-863.

Lee, Y. S. (2000). The sustainability of university-industry research collaboration: An empirical assessment. *The journal of Technology transfer*, 25(2), 111-133.

Lindblom, C. E. (1965). *The intelligence of democracy: Decision making through mutual adjustment*. New York: Free Press.

MATLAB. MATLAB R2010a. Massachusetts: Mathworks, 2010.

McDonald, D. (2017). *The golden passport – Harvard Business School, the limits of capitalism and the moral failure of the MBA elite*. New York: Harper Collins.

Morgan, K. (2004). The exaggerated death of geography: learning, proximity and territorial innovation systems. *Journal of economic geography*, 4(1), 3-21.

Orsenigo, L. (2001). The (failed) development of a biotechnology cluster: The case of Lombardy. *Small Business Economics*, 17(1), 77-92.

Pedrycz, W., & Gomide, F. (1998). *An introduction to fuzzy sets: analysis and design*. Cambridge: Mit Press.

Perkmann, M., & Walsh, K. (2008). Engaging the scholar: Three types of academic consulting and their impact on universities and industry. *Research Policy*, 37(10), 1884-1891.

Rodríguez-Pose, A., & Crescenzi, R. (2008). Mountains in a flat world: why proximity still matters for the location of economic activity. *Cambridge Journal of Regions, Economy and Society*, 1(3), 371-388.

Siegel, D. S., Waldman, D. A., Atwater, L. E., & Link, A. N. (2003). Commercial knowledge transfers from universities to firms: improving the effectiveness of university-industry collaboration. *The Journal of High Technology Management Research*, 14(1), 111-133.

Schmoch, U., Laville, F., Patel, P., Frietsch, R. (2003). *Linking technology areas to industrial sectors. Final Report European Comissions*. DG Research.

Stokes, D. E. 1997. *Pasteur's quadrant: Basic science and technological innovation*. Washington, DC: Brookings Institution Press.

Tartari, V., & Breschi, S. (2012). Set them free: scientists' evaluations of the benefits and costs of university–industry research collaboration. *Industrial and Corporate Change*, 21(5), 1117-1147.

Tijssen, R. (2012). Joint research publications: A performance indicator of university-industry collaboration. *Assessment & Evaluation in Higher Education*, 5(2), 19-40.

Tushman, M., & O'Reilly, C. (2007). Research and relevance: Implications of Pasteur's quadrant for doctoral programs and faculty development. *Academy of Management Journal*, 50(4), 769-774.

White Jr., O. Review: The intelligence of democracy: decision making through mutual adjustment by Charles E. Lindblom. *The western political quarterly*, 18(4), 1965.

Willke, H. (2007). *Smart governance: governing the global knowledge society*. Campus Verlag.

Zadeh, L. (1965). A. Fuzzy sets. *Information and control*. 8(3), 338-353.

APÊNDICE C

Trabalhos e projetos realizados e em andamento relacionados à tese

1. Projeto conceitual de uma base de dados sobre o efeito de polimorfismo na toxicidade e eficácia de Insumos Farmacêuticos Ativos, pertencentes a RENAME

O projeto, apoiado pelo CNPq/Anvsia, consiste em um levantamento bibliográfico, a partir de diferentes fontes de informação, de todos os aspectos relacionados ao polimorfismo em fármacos da Relação Nacional de Medicamentos Essenciais (RENAME), principalmente aqueles que afetam a eficácia e toxicidade. A organização da informação levantada, sua sistematização e desenvolvimento de matrizes de avaliação de riscos associados ao polimorfismo na forma de um sistema gerenciador de banco de dados é proposta como produto final do projeto. Visto que o entendimento do comportamento no estado sólido de todos os componentes de um medicamento (IFAs e excipientes) é essencial para o desenvolvimento de medicamento e para a farmacovigilância, a inovação proposta neste projeto é proporcionar aos especialistas oriundos das mais diversas áreas o acesso rápido, fácil e mais completo possível às informações necessárias e pertinentes para a tomada de decisão relacionada ao polimorfismo e as propriedades pertinentes à avaliação do risco sanitário para os IFAs que constam na RENAME.

2. Comparação Interlaboratorial para Análise de Fases Polimórficas por Difração de Raios X (1ª Rodada) - Policristais de Insumos Farmacêuticos Ativos Tibolona

A motivação deste projeto envolve o interesse do Inmetro em avaliar o desempenho de laboratórios de ensaios que usam o método de difração de raios-X na análise quantitativa de fases polimórficas em materiais policristalinos, fornecendo confiabilidade aos resultados obtidos e uma validação adicional para as metodologias experimentais e de análise utilizadas. Os objetivos específicos incluem: Verificar o desempenho de laboratórios na análise quantitativa de polimorfos em insumos farmacêuticos ativos (IFA); Prover os laboratórios com ensaios de proficiência para atender os requisitos da norma ABNT NBR ISO/IEC 17025 quanto a garantia da qualidade dos resultados de medição e Contribuir para aumentar a confiança nos resultados das medições dos laboratórios.

3. Comparação piloto para validação de método de análise quantitativa de fases por DRX em IFAs: caso tibolona

Neste projeto são investigados diferentes modelos estruturais e microestruturais aplicados a análise de dados de difração de raios X (DRX) de pó obtidos para misturas de concentrações conhecidas das fases polimórficas do IFA tibolona. Os dados de DRX obtidos em diferentes equipamentos são analisados via método Rietveld aplicando o mesmo modelo analítico. São avaliados os fatores instrumentais, de coleta de dados e de modelo teórico de análise assumido que afetam o resultado final de medição.

4. Desenvolvimento de protocolo de caracterização de padrões de difração de raios-X em monocristal

Seguindo as diretrizes do FDA e da ANVISA, as indústrias farmoquímicas e farmacêuticas devem monitorar a presença de polimorfismo em toda a cadeia produtiva do medicamento, ou seja, desde a aquisição do insumo farmacêutico ativo até o armazenamento do produto acabado em prateleira. A caracterização de uma estrutura cristalina por difração de raios-X em monocristal é considerada a técnica definitiva para a evidência de polimorfismo. Outra técnica eficiente é a difração de raios-X por policristais, a qual é considerada uma prova inequívoca da existência de polimorfismo e é, reconhecidamente, uma das referências mundiais para a identificação de polimorfismo em insumos farmacêuticos ativos sólidos. Em diversas demandas de serviços recebidas pela NanoBusiness foi verificado a inexistência de informações cristalográficas acerca de numerosas formas polimórficas de diversas substâncias, o que dificultou o controle de qualidade adequado a partir das técnicas de difração de raios-X. Este projeto visou sanar essas lacunas existentes na literatura, a partir da obtenção e caracterização de novas estruturas cristalinas de polimorfos, contribuindo para a complementação do Banco de Dados e-Diffraction Pharma e para o aumento da qualidade do serviço de análise qualitativa e quantitativa de polimorfismo oferecido pela NanoBusiness.

5. Desenvolvimento de novos candidatos a fármacos antitumorais" **(Edital: E_07/2016 - Apoio ao Empreendedorismo e Formação de START-UPS EM SAÚDE HUMANA do RJ**

O presente projeto visa o desenvolvimento de novos compostos com potencial atividade antitumoral como alternativa terapêutica mais seletiva e menos tóxica, tendo em vista os reconhecidos efeitos colaterais que usualmente aparecem em pacientes submetidos a quimioterapia do câncer. Neste contexto, este projeto propõe o desenvolvimento, caracterização e avaliação farmacológica de novas moléculas que potencializem a ação antitumoral através da síntese de complexos de coordenação de Cobre e Platina com ligantes multidentados a base de nitrogênio e oxigênio. Em especial, pretende-se desenvolver moléculas que sejam uma alternativa terapêutica mais seletiva e menos tóxica do que aquelas atualmente disponíveis no mercado em função dos reconhecidos efeitos colaterais que usualmente aparecem em pacientes submetidos a quimioterapia do câncer