Universidade Estadual de Goiás
Campus Anápolis de Ciências Exatas e Tecnológicas
Programa de Pós-Graduação *Stricto Sensu* em Recursos Naturais do Cerrado

WESLEY FONSECA VAZ

DESENVOLVIMENTO DE MATERIAIS CRISTALINOS PARA APLICAÇÕES AGROQUÍMICAS: DIHIDROQUINOLINAS COM POTENCIAL FUNGICIDA E CARACTERIZAÇÃO DE PRODUTOS DE DEGRADAÇÃO DE S-TRIAZINAS

Anápolis
2021

WESLEY FONSECA VAZ

DESENVOLVIMENTO DE MATERIAIS CRISTALINOS PARA APLICAÇÕES AGROQUÍMICAS: DIHIDROQUINOLINAS COM POTENCIAL FUNGICIDA E CARACTERIZAÇÃO DE PRODUTOS DE DEGRADAÇÄO DE S-TRIAZINAS

Tese apresentada ao Programa de Pós-Graduação *Stricto Sensu* em Recursos Naturais do Cerrado, da Universidade Estadual de Goiás para a obtenção do título de Doutor em Recursos Naturais do Cerrado.
Orientador: Prof. Dr. Hamilton Barbosa Napolitano.
Coorientador: Prof. Dr. Paulo de Sousa Carvalho Jr.

Anápolis
2021

TERMO DE AUTORIZAÇÃO PARA PUBLICAÇÃO DE TESES E DISSERTAÇÕES NA BIBLIOTECA DIGITAL (BDTD)

Na qualidade de titular dos direitos de autor, autorizo a Universidade Estadual de Goiás a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UEG), regulamentada pela Resolução, **CsA n.1087/2019 sem** ressarcimento dos direitos autorais, de acordo com a Lei nº 9610/98, o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou *download*, a título de divulgação da produção científica brasileira, a partir desta data.

Dados do autor (a)

Nome Completo: WESLEY FONSECA VAZ

E-mail: wesfonseca@gmail.com

Dados do trabalho

Título: Desenvolvimento de materiais cristalinos para aplicações agroquímicas: Dihidroquinolinas com potencial fungicida e caracterização de produtos de degradação de s-triazinas

Tipo

(x) Tese ( ) Dissertação

Curso/Programa: RECURSOS NATURAIS DO CERRADO

Concorda com a liberação documento [x] SIM [ ] NÃO[1]

ANÁPOLIS, 03/03/2021
Local Data

Wesley Fonseca Vaz
Assinatura do autor (a)

Assinatura do orientador (a)

[1] **Casos de impedimento:**
- Período de embargo é de um ano a partir da data de defesa
- Solicitação de registro de patente;
- Submissão de artigo em revista científica;
- Publicação como capítulo de livro;
- Publicação da dissertação/tese em livro.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da UEG
com os dados fornecidos pelo(a) autor(a).

V393d Vaz, Wesley Fonseca
Desenvolvimento de materiais cristalinos para aplicações agroquímicas: dihidroquinolinas com potencial fungicida e caracterização de produtos de degradação de s-triazinas / Wesley Fonseca Vaz; orientador Hamilton Barbosa Napolitano; co-orientador Paulo de Sousa Carvalho Júnior. -- Anápolis, 2021.
102 p.

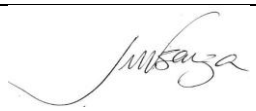
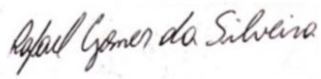
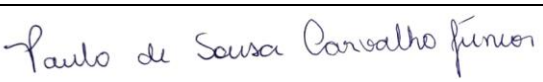
Tese (Doutorado - Programa de Pós-Graduação Doutorado em Recursos Naturais do Cerrado (RENAC)) -- Câmpus Central - Sede: Anápolis - CET, Universidade Estadual de Goiás, 2021.

1. dihidroquinolinas. 2. machine learning. 3. herbicidas . 4. cristalografia. I. Napolitano, Hamilton Barbosa, orient. II. Júnior, Paulo de Sousa Carvalho, co-orient. III. Título.

PROGRAMA DE PÓS-GRADUAÇÃO *STRICTO SENSU* EM RECURSOS NATURAIS DO CERRADO

ATA DE DEFESA PÚBLICA DE DISSERTAÇÃO Nº 95/2021

Aos 26 dias do mês de fevereiro do ano de 2021, às 8h00min, reuniu-se (Google Meet), a banca Examinadora composta pelos: Prof. Dr. Hamilton Barbosa Napolitano (Universidade Estadual de Goiás), Prof. Dr. Fabiano Guimarães Silva (IF Goiano), Prof. Dr. João Maurício Fernandes Souza (UniEvangélica), Prof. Dr. Rafael Gomes da Silveira (IF Goiano), Prof. Dr. Rodrigo de Souza Correa (Universidade Federal de Ouro Preto) e Prof. Dr. Paulo de Sousa Carvalho Jr (Universidade Federal de Mato Grosso do Sul) para sob a presidência do primeiro, procederem à "defesa de Doutorado" intitulada: *Desenvolvimento de materiais cristalinos para aplicações agroquímicas: Dihidroquinolinas com potencial fungicida e caracterização de produtos de degradação de s-triazinas*, de Wesley Fonseca Vaz, discente do PPG Recursos Naturais do Cerrado, nível doutorado. Foi realizada a avaliação oral no sistema de apresentação e defesa de tese de autoria do discente. Terminada a avaliação oral, a Banca Examinadora reuniu-se emitindo os seguintes pareceres mediante as justificativas e sugestões abaixo:

Membro da Banca	Parecer (Aprovado/Reprovado)	Assinatura
Dr. Hamilton Barbosa Napolitano	Aprovado	
Dr. Fabiano Guimarães Silva	Aprovado	
Dr. João Maurício Fernandes Souza	Aprovado	
Dr. Rafael Gomes da Silveira	Aprovado	
Dr. Rodrigo de Souza Correa	Aprovado	
Dr. Paulo de Sousa Carvalho Jr (coorientador)	Aprovado	

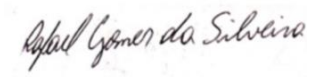
Justificativas e sugestões: *O discente respondeu satisfatoriamente as perguntas da banca*.

Após avaliação foi APROVADO na defesa. Às **11h30min** horas, o Prof. Dr. Hamilton Barbosa Napolitano, presidente da Banca Examinadora deu por encerrada a sessão e, para constar, lavrou a presente Ata:

Prof. Dr. Hamilton Barbosa Napolitano

Prof. Dr. Fabiano Guimarães Silva

Prof. Dr. João Maurício Fernandes Souza



Prof. Dr. Rafael Gomes da Silveira

Prof. Dr. Rodrigo de Souza Correa

Prof. Dr. Paulo de Sousa Carvalho Jr

AGRADECIMENTOS

Inicialmente a Deus, que me concedeu o dom da vida e a oportunidade de desfrutar deste momento.

À minha família, minha eterna gratidão, amor e confiança. Vocês foram a base sólida para que eu pudesse chegar até aqui.

Ao Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso, por me conceder o afastamento integral de minhas atividades laborais para que eu pudesse me dedicar exclusivamente ao doutorado.

Ao Programa de Pós-Graduação *Stricto Sensu* em Recursos Naturais do Cerrado, da Universidade Estadual de Goiás, pela oportunidade oferecida.

Ao meu orientador, Prof. Dr. Hamilton Barbosa Napolitano; ao meu coorientador Prof. Dr. Paulo de Sousa Carvalho Jr; ao meu orientador nos EUA, Prof. Dr. Allen Grayson Oliver e ao meu orientador na Alemanha, Prof. Dr. Hongbin Zhang. A todos vocês, pela dedicação, compreensão, incentivo à produção científica acadêmica, conhecimentos transmitidos e confiança depositada na realização deste trabalho. A humildade e paciência em ensinar e colaborar estarão sempre em minha memória.

À Prof[a]. Dr[a] Caridad Noda Perez, da Universidade Federal de Goiás, por ceder as amostras de dihidroquinolinas que foram estudadas nessa tese.

Aos colegas do laboratório de química teórica e estrutural de Anápolis (QTEA), pelas experiências trocadas e pela ótima convivência, bem como a todos os coautores nos artigos publicados durante esse período.

Às agências de fomento: Fundação de Amparo à Pesquisa do Estado de Goiás (FAPEG); Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e Serviço Alemão de Intercâmbio Acadêmico (DAAD) por acreditar no potencial desse estudo. À primeira pela concessão de bolsa para o desenvolvimento dessa pesquisa; à segunda pela bolsa no Programa de Doutorado Sanduíche no Exterior (*University of Notre Dame* – EUA); e à terceira pela bolsa no edital *Co-financed Short-Term Research Grant Brazil, 2020* (*Technische Universität Darmstadt* - Alemanha).

“Um menino caminha e caminhando chega no muro
E ali logo em frente, a esperar pela gente, o futuro está
E o futuro é uma astronave que tentamos pilotar
Não tem tempo, nem piedade, nem tem hora de chegar
Sem pedir licença, muda a nossa vida depois convida a rir ou chorar
Nessa estrada não nos cabe conhecer ou ver o que virá
O fim dela ninguém sabe bem ao certo onde vai dar
Vamos todos, numa linda passarela
De uma aquarela que um dia, enfim,
Descolorirá”

AQUARELA
Toquinho, Vinicius de Moraes, Maurizio Fabrizio

RESUMO

A demanda por novos agentes agroquímicos gera uma busca, cada vez maior, pelo desenvolvimento ou aperfeiçoamento de moléculas que possibilitem uma redução no número de doses aplicadas, sejam alternativas para os problemas de resistência e que possam combater mais de uma praga para uma mesma lavoura. Dessa maneira, o que se procura é o equilíbrio entre a proteção dos cultivos, o aumento da produtividade e o menor impacto sobre o ambiente. Da relação estrutura atividade, obtida por modelos qualitativos, sabe-se que certos grupos funcionais estão relacionados a determinadas propriedades biológicas mesmo que em espécies químicas diferentes. Frequentemente, moléculas semelhantes possuem propriedades semelhantes, sendo essa análise importante na compreensão da resposta biológica desencadeada a partir da presença de determinado grupo químico. Partindo-se dessa premissa, elucidou-se, por meio da difração dos raios X, a estrutura molecular de dez compostos sintetizados da classe das dihidroquinolinas. Em seguida, em busca do entendimento da relação estrutura-atividade, foram empregados nesse conjunto métodos de *machine learning* que predisseram essas moléculas como potenciais agentes fungicidas. Foi possível mapear as regiões moleculares e discernir quais características eram determinantes para tal atividade. Além disso, um importante metabólito inédito, fruto da degradação dos herbicida ametrina e atrazina, foram estruturalmente elucidados na forma de quatro sais, utilizando a mesma técnica empregada para as dihidroquinolinas. Desse modo essa tese aponta as dihidroquinolinas como uma classe promissora, e aberta a estudos futuros, para o desenvolvimento de novos agentes agroquímicos, além de disponibilizar toda elucidação estrutural que permite ampliar o entendimento da cinética de degradação de herbicidas, com metodologia que pode ser replicada para demais agroquímicos.

Palavras-Chave: dihidroquinolinas, *machine learning*, herbicidas, cristalografia.

ABSTRACT

The demand for new agrochemicals has driven an increasing search for the development or improvement of molecules that enable a reduction in the number of doses applied, as alternatives to resistance problems and that can reduce more than one pest for the same crop. In this way, what is sought is the balance between the protection of crops, the increase in productivity, and the reduction of environmental impacts. From the structure-activity relationship, obtained by qualitative models, it is known that certain functional groups are related to specific biological properties even in different chemical species. Similar molecules often have similar properties, and this analysis is important in understanding the biological response caused by the presence of a specific chemical group. Based on this perspective, the molecular structure of ten synthesized dihydroquinolines was elucidated using X-ray diffraction. Then, to understand the structure-activity relationship, machine learning methods were used in this set that predicted these molecules as potentially fungicidal agents. It was possible to map the molecular regions and discern which characteristics were decisive for such activity. Also, an important new metabolic, resulting from the degradation of the herbicides ametrine and atrazine, was structurally elucidated in the form of four salts, using the same technique used for dihydroquinolines. In this way, this thesis points to dihydroquinolines as a promising class, and open to future studies, for the development of new agrochemical agents, in addition to providing structural insights that allow broadening the understanding of herbicide degradation kinetics, with a methodology that can be replicated for other agrochemicals.

Keywords: dihidroquinolines, *machine learning*, herbicides, crystallography

LISTA DE FIGURAS

SUMÁRIO

1 INTRODUÇÃO

A investigação dos processos ambientais através da relação entre biodiversidade e compostos químicos, por meio da Modelagem Molecular[1], é uma metodologia capaz de ampliar o entendimento acerca das propriedades dos compostos e direcionar a obtenção de novos produtos (YURIEV; COOTE, 2011). Dentro da biodiversidade, nota-se uma diversidade molecular, que consiste nas diferenças estruturais observadas para moléculas distintas, embora da mesma classe (SINGH; PELAEZ, 2008). Especificamente, a relação entre características estruturais e as atividades resultantes delas nos permite aprimorar ou desenvolver novos produtos com aplicação ambiental.

O Cerrado[2] é significativamente conhecido como uma fronteira agrícola, isto é, apresenta condições adequadas para o uso agrícola. Essas condições dizem respeito a geografia física (solos, relevo, clima e precipitação) e a aspectos econômicos (sistema de transporte e necessidade de financiamento) da região. Esse bioma conglomera vários tipos de vegetação e apresenta duas estações bem definidas: uma seca de inverno e outra úmida de verão e com distribuição concentrada das chuvas em toda a região. Visto que o relevo é, em geral, plano ou de ondulações suaves, a região apresenta boas possibilidades para o emprego de práticas agrícolas mecanizadas (PEIXOTO *et al*., 2019; REIS *et al*., 1973)

Contudo, o Cerrado possuía características que impossibilitavam seu desenvolvimento agrícola, tais como: baixa fertilidade natural dos solos, devido à sua acidez, alto teor de alumínio, baixas concentrações de cálcio e magnésio, além de ser pobre em fósforo assimilável. No entanto, estas características foram superadas com

[1] A modelagem molecular é uma coleção de técnicas baseadas em computador para representar e manipular estruturas moleculares numericamente e simular seu comportamento utilizando para isso as equações da física quântica e clássica. Essas simulações podem ser bastante úteis como guias para o projeto e interpretação de experimentos e para o entendimento das propriedades que dependem dessas estruturas, levando a novos *insights* da dinâmica e função moleculares (GÁSPÁRI; PERCZEL, 2010; GU; LI, 2011; PIMENTEL; GUIMARÃES; MILLER, 2013)

[2] O termo Cerrado é comumente utilizado para designar o conjunto de ecossistemas (savanas, matas, campos e matas de galeria) que ocorrem no Brasil Central (KLINK *et al*., 2005). As culturas consideradas adaptadas aos solos de Cerrado são: arroz de sequeiro, soja, mandioca e abacaxi; e em áreas com condições climáticas favoráveis, trigo, amendoim, milho e feijão. Avanços posteriores permitiram aumentar a lista de culturas com ampliação da produção de café, algodão, sorgo, frutas, hortaliças e legumes.

a correção química do solo, que viabilizou a incorporação de insumos químicos, superando os problemas de fertilidade. Assim, o Cerrado tornou-se a área com as melhores condições para a expansão e diversificação da produção agropecuária brasileira (REIS *et al*., 1973)

Assim, este uso demanda na busca de tecnologias que possam proteger, favorecer ou incrementar a produção agrícola, principalmente envolvendo os fungos, bactérias e plantas (NAPOLITANO *et al*., 2016). Uma das maneiras de incrementar essa produção vêm do uso de pesticidas[3], que são substâncias utilizadas para proteger as plantas agrícolas da competição de plantas daninhas e de destruições causadas por fungos, insetos, ácaros e roedores (FREEDMAN, 1995). Porém, o uso prolongado, intensivo e indiscriminado de agroquímicos afeta a biodiversidade do solo, a sustentabilidade agrícola e a segurança alimentar, trazendo efeitos nocivos de longo prazo à segurança nutricional e à saúde humana e animal (MEENA *et al*., 2020).

Recentemente, boa parte das publicações sobre o desenvolvimento de novas tecnologias para proteção de cultivos aborda uma metodologia orientada à tecnologia para a inovação (SCHUT *et al*., 2014). A descoberta de agroquímicos modernos enfrenta desafios significativos em um mercado em rápida evolução. Dentre esses desafios destaca-se a resistência de agroquímicos e seu manejo, surgimento de novas pragas, mudanças na percepção pública da tecnologia agrícola, o aumento do custo da pesquisa e desenvolvimento, e requisitos regulatórios (ZHANG *et al*., 2018). Dessa forma tem-se a problemática de propor inovações nas rotas de descoberta de novas substâncias (ingredientes ativos) que possam atuar como pesticidas, ou ainda o aprimoramento daqueles que já são utilizados, numa perspectiva sustentável, considerando sua utilização para manutenção da produtividade e ainda os riscos do seu uso generalizado para a saúde (CANTRELL; DAYAN; DUKE, 2012).

No intuito de propor soluções a essas demandas, pesquisadores têm usado diferentes tipos de abordagem. Uma delas, através da nanotecnologia, é melhorar as estratégias para o controle de sistemas de entrega para liberação lenta e sustentada de agroquímicos, através do desenvolvimento de novos agentes carreadores em formulações de pesticidas (GRILLO; ABHILASH; FRACETO, 2016; SINHA; GHOSH; SIL, 2017; YUSOFF; KAMARI; ALJAFREE, 2016). O desenvolvimento

[3] Uma discussão mais aprofundada considerando a divergência em torno da nomenclatura agrotóxico e pesticida pode ser encontrada em um artigo de Roberto Leiser Baronas (BARONAS, 2019)

de pesticidas para a agricultura sustentável é um foco importante na indústria agroquímica e no meio acadêmico.

Outra abordagem envolve a seleção de compostos químicos de origem natural, dado seu considerável potencial para o manejo de vetores. Esses produtos tendem a ser mais seguros do que os inseticidas convencionais devido à sua rápida biodegradação e baixa toxicidade para inimigos naturais, humanos e outros mamíferos (MISHRA et al., 2020; S. MANN; E. KAUFMAN, 2012). No entanto, a quantidade de compostos bioativos no material vegetal é pequena, o que torna a extração e uso de compostos naturais para controle de insetos em larga escala algo oneroso (SANTOS; VARDANEGA; DE ALMEIDA, 2014).

Há também as pesquisas no campo do *design* auxiliado por computador (*computer-aided design*), que buscam orientar o projeto e o desenvolvimento de inseticidas mais seletivos e ambientalmente menos agressivos (JITONNOM, 2014). A abordagem computacional foi utilizada, por exemplo, para discriminar as proteínas resistentes a inseticidas de proteínas não resistentes. Essa resistência é conferida por proteínas codificadas por certas classes de genes dos insetos. Uma maneira de contornar essa situação foi o desenvolvimento de um servidor de predição online (DIRProt[4]) para prognóstico computacional de proteínas que causam resistência a inseticidas (MEHER *et al*., 2017).

Ainda nessa perspectiva, foi desenvolvida uma proposta de aplicação de triagem virtual para descoberta de inibidores para enzimas mediadoras de resistência. Os compostos identificados reverteram a resistência a inseticidas em duas espécies diferentes de pragas, representando uma estratégia sustentável que permite uma redução significativa do uso de inseticidas sem comprometer sua eficácia (CORREY *et al*., 2019). Outra estratégia de sucesso é a obtenção de novas formas sólidas de agroquímicos comerciais e o estudo de suas rotas de degradação. Um exemplo disso foi o desenvolvimento de duas novas formas sólidas do diclorodifeniltricloroetano (DDT). Apesar do impacto ambiental desses compostos não terem sido avaliados, eles devem ser considerados como um novo material, livres do estigma do DDT.

Além disso, as formas sólidas dessas novas composições exibem diferentes atividades em diferentes insetos, demonstrando que a manipulação da estrutura do

[4] O servidor de previsão é disponibilizado gratuitamente em http://cabgrid.res.in:8080/dirprot para usuários acadêmicos.

estado sólido é uma estratégia viável para o projeto de inseticidas mais eficazes (ZHU *et al*., 2019). Também é possível obter resultados em remediação, uma vez que a remoção de resíduos de pesticidas do meio ambiente é uma preocupação séria. Há um grande interesse na pesquisa e desenvolvimento de tecnologias que possam identificar produtos de degradação, a fim de diminuir seus efeitos prejudiciais ao meio ambiente (PERGAL *et al*., 2020).

Considerando tal contexto, esta tese propõe duas abordagens para desenvolver pesticidas, sendo a primeira, baseada na relação estrutura-atividade (SAR), investigar o potencial pesticida de dihidroquinolinas, por meio de métodos *in sílico*. Esse conhecimento permite fazer modificações estruturais na molécula, através da síntese planejada, o que pode potencializar suas atividades biológicas (STOMPOR *et al*., 2015). Na segunda estratégia, a hidróxi-atrazina, um produto de degradação da ametrina e atrazina, foi estruturalmente caracterizado por meio da difração dos raios X. Por mais que esse não seja um desenvolvimento *ab initio*, a compreensão e identificação de metabolitos tem impactos na compreensão dos mecanismos e da cinética de degradação. Assim, novos compostos (derivados e análogos que sejam mais seguros e efetivos), podem ser desenvolvidos considerando esta vertente.

A primeira abordagem nos levou a requerer um registro de pedido de patente[5] (BR 102021001396-6) uma vez que se alcançou o nível 4 de maturidade de prontidão tecnológica, (*Technology Readiness Level* - TRL[6]). Os resultados obtidos contribuem para o campo de estudo nos seguintes aspectos: no desenvolvimento de algoritmos de *machine learning* que predizem corretamente, uma vez que há validação experimental, as dihidroquinolinas em estudo como compostos com ação fungicida; a detalhada caracterização estrutural de dez moléculas e quatro produtos de degradação;

[5] O acompanhamento do Processo pode ser feito utilizando o número de Registro: BR 102021001396-6. Para isso, basta acessar o site do INPI e clicar no ícone busca na WEB. Em seguida selecione a opção Patente. Clique em continuar sem identificação e, na próxima aba digite o número do processo.

[6] A escala do sistema de prontidão tecnológica varia de TRL1 até TRL9. A pesquisa básica ou prova de conceito preliminar compreende a TRL de 1 a 3. A TRL1 refere-se à fase de ideias; a TRL2, à pesquisa exploratória baseada num conceito tecnológico e/ou ideia de aplicação, podendo ser chamada de demonstração preliminar; e a TRL3, à pesquisa sistemática baseada no mínimo de resultados favoráveis. O desenvolvimento tecnológico compreende a TRL4 e a TRL5. A TRL4 refere-se à validação dos componentes da tecnologia em ambiente de laboratório (QUINTELLA *et al*., 2019).

e, por fim, mostra que o entendimento estrutural é um poderoso instrumento de partida no planejamento racional de compostos com atividade pesticida.

Referente a organização desta tese, o tópico 2 apresenta os objetivos; o tópico 3 discute as partes centrais do referencial teórico; o tópico 4 apresenta os resultados obtidos na forma de quatro artigos (na íntegra[7]) e uma primeira versão de um manuscrito a ser submetido; o tópico 5 conclui este documento com as considerações finais e perspectivas.

[7] Aos autores em periódicos da Elsevier [https://bityli.com/pbO3V] e da Springer [https://bityli.com/6EwTJ], é permitida a reprodução de artigos autorais, na íntegra, em sua tese ou dissertação.

2 OBJETIVOS

(i) estudar a potencial atividade pesticida das dihidroquinolinas através de *machine learning* e da identificação dos aspectos estruturais do conjunto molecular, por meio da caracterização no estado sólido (difração de raios X).

(ii) avaliar um metabólito, produto de degradação dos herbicidas ametrina e atrazina, caracterizando os compostos formados nesse processo, a fim de avaliar seu impacto no meio ambiente.

3 NOVOS MATERIAIS EM AGROQUÍMICA

3.1 Ingredientes agroquímicos e a demanda por moléculas bioativas

Ao investigar os desafios e tendências atuais na descoberta de agroquímicos, Lamberth e colaboradores (2013) publicaram uma revisão destacando a evolução da química de proteção de culturas de uma ciência quase alquímica do século XIX para uma ciência de alta tecnologia, que apoia a produção sustentável de alimentos, rações e fibras para uma população em rápido crescimento. O desenvolvimento no campo do *design* e síntese de agroquímicos são fundamentais para a superação dos desafios, como: resistência a plantas daninhas e pragas, maiores margens de segurança que atendam marcos de regulamentação e maior custo das mercadorias. E isso só é possível com o desenvolvimento de ativos que sejam mais seletivos, ingredientes ambientalmente benignos e de baixa taxa de uso, aliado a um baixo custo.

Dessa maneira, pode-se inferir que a busca por moléculas com aplicações agrícolas passa pelos mesmos aspectos relacionados ao desenvolvimento de novos fármacos (BERTHON, 2020); logo, qualquer que seja a estratégia adotada, em algum momento ela irá passar pelo estado sólido, que demanda a caracterização estrutural e físico-química do composto. Nesta perspectiva existem algumas opções. Dentre elas, a busca por novos compostos ativos e eficazes. Sendo este o caso das dihidroquinolinas, uma vez há evidências de suas propriedades agroquímicas, conforme discutido na seção 3.2. Além disso, deve-se ainda investigar a efetividade dos compostos em consonância com a sustentabilidade. Daí surge a proposta de analisar e repensar[8] os herbicidas ametrina e atrazina. Para tal, a análise do estado sólido desses compostos e seus produtos é essencial.

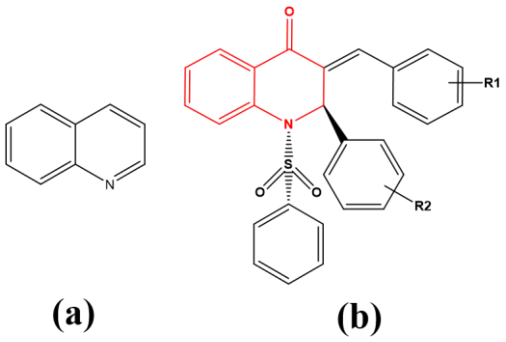
[8] Estratégia conhecida na área de fármacos, mas que também pode ser aplicada a agroquímicos, pelo termo *repurposing*, que denota a descoberta de novas aplicações para drogas existentes ou candidatos a drogas que são descobertas e refinadas (CHARLTON *et al*., 2018)

3.2 Dihidroquinolinas e s-triazinas

A classe das quinolinas (1-aza-naftaleno ou benzo[β]piridina) são caracterizadas por moléculas que possuem um núcleo aromático heterocíclico, contendo nitrogênio (Figura 1a), e ocorre em vários compostos naturais (*Cinchona Alkaloids*) estando presentes em substâncias farmacologicamente ativas exibindo uma ampla gama de atividades biológicas.

Dentre elas, podemos citar: atividade antimalárica, antibacteriana, antifúngica, anti-helmíntica, cardiotônica, anticonvulsivante, anti-inflamatória e analgésica (FELCZAK *et al.*, 2016; MARELLA *et al.*, 2013; ZENG; ZENG; LIU, 2016). Atualmente, são recorrentes os artigos que relatam a síntese de quinolinas funcionalizadas para diversas atividades biológicas (LIU *et al.*, 2017a; NEFERTITI *et al.*, 2017; ROSADO-SOLANO *et al.*, 2019; WEYESA; MULUGETA, 2020). Em diversos estudos elas são reportadas com atividade inibitória para a enzima HIV-1 integrase (SECHI *et al.*, 2009), como agente antioxidante (FILALI BABA *et al.*, 2019), como agente antituberculose (KUMAR *et al.*, 2018) e como transportadores para entrega de compostos biológicos ativos ao cérebro (FOUCOUT *et al.*, 2009)

Figura 1 - Estrutura química bidimensional de uma quinolina (a) e de uma dihidroquinolina (b) com o núcleo 1,4-dihidroquinolina destacado em vermelho



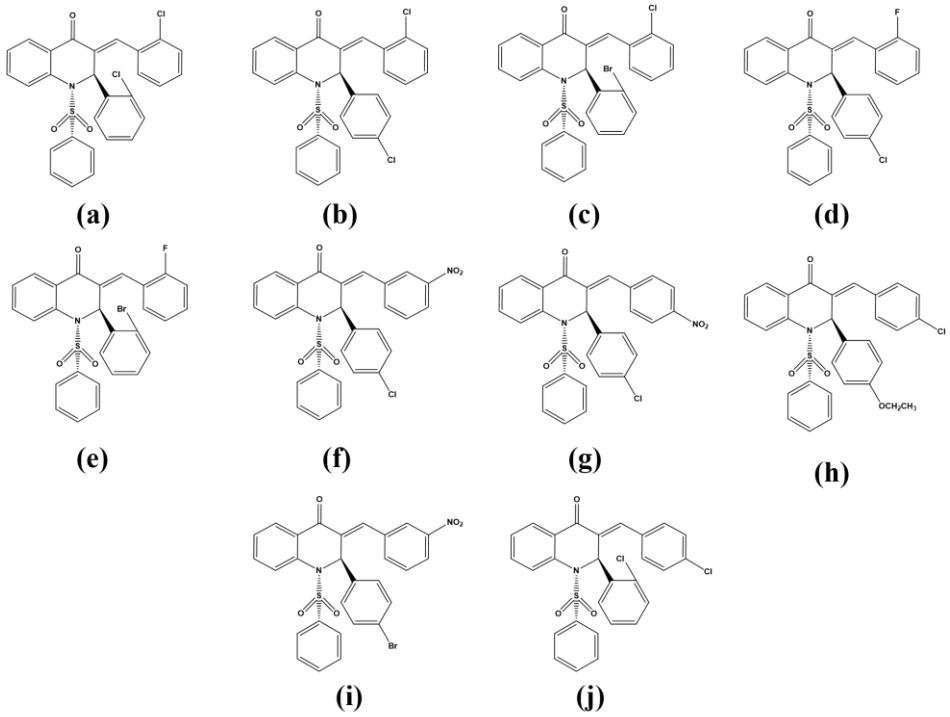
Fonte: Elaborada pelo autor.

Nosso grupo de pesquisa tinha a disposição um conjunto sintetizado de dihidroquinolinas halogenadas e já cristalizadas. A partir disso, utilizamos modelos de *machine learning* na busca pelo potencial pesticida desses compostos. Os resultados mostraram que estávamos de posse de moléculas promissora para a

aplicação alvo, e de fato estes compostos têm sido objeto de estudo em diversos processos ambientais (CANTRELL; DAYAN; DUKE, 2012).

Esses compostos são de fácil obtenção e a metodologia de síntese das moléculas reportados nessa tese (Figura 2), bem como seu mecanismo de reação, é descrita em detalhes na literatura (CASTRO *et al*., 2015; D'OLIVEIRA *et al*., 2018) e também em cada um dos respectivos artigos publicados. Em suma, as moléculas foram sintetizadas via condensação de Claisen-Schmidt[9], em que ao se reagir um aldeído aromático com uma cetona aromática, via catálise básica em meio etanólico, obtêm-se uma chalcona. Em seguida, essa chalcona sofre uma ciclização em meio ácido, gerando a dihidroquinolina que, então, reage em meio básico com um aldeído aromático produzindo os compostos em estudo. Os rendimentos para essa reação estão reportados entre 45% e 94%.

Figura 2 - Estruturas moleculares bidimensionais das dihidroquinolinas estudadas nessa tese.



Fonte: Elaborada pelo autor.

[9] A condensação de um aldeído aromático com um aldeído ou cetona alifático na presença de uma base ou um ácido para formar um aldeído ou cetona α, β-insaturado com alta quimiosseletividade é geralmente conhecida como condensação de Claisen-Schmidt (WANG, 2010)

Até o momento o banco de dados cristalográficos de pequenas moléculas (CSD)[10] (GROOM *et al*., 2016) (CSD version 5.42 – November 2020) reporta 151 estruturas com o núcleo 1,4-dihidroquinolina, um baixo número frente a outras classes semelhantes, como as chalconas (3917 estruturas), por exemplo. Esse baixo número pode ser inerente a dificuldade de cristalizar esses compostos, bem como o interesse em se pesquisar seu estado sólido. Isso demonstra que, em termos de conhecimento estrutural, há muito ainda a ser estudado sobre essas moléculas. Cabe salientar que ainda não se identificou nenhum pesticida comercial que tenha como princípio ativo uma dihidroquinolina. Sabe-se que o processo de desenvolvimento de um novo pesticida, até chegar às prateleiras, é longo (requerendo de 8 a 12 anos, em média) e dispendioso (aproximadamente US$ 250 milhões) (SPARKS, 2013).

A ametrina e a atrazina, Figura 3, são herbicidas pertencentes ao grupo das s-triazinas utilizados largamente nas áreas de solo no Cerrado, pois são recomendados para o controle de plantas daninhas na cultura do milho e da cana-de-açúcar, respectivamente (PRATA *et al*., 2001). A atrazina, banida nos países da união europeia em 2004 por causar distúrbios endócrinos, afetando o sistema hormonal (BETHSASS; COLANGELO, 2006), é o segundo herbicida mais utilizado no Brasil, possuindo, segundo a plataforma AGROFIT[11], 28 formulações disponíveis no mercado, enquanto a ametrina possui 66 formulações. O monitoramento do solo e da água nas áreas em que houve uso desses herbicidas e o conhecimento da sua mobilidade em solos no Cerrado é de extrema importância, pois pode evidenciar uma possível contaminação das águas subterrâneas da região (MARCHI, 2011).

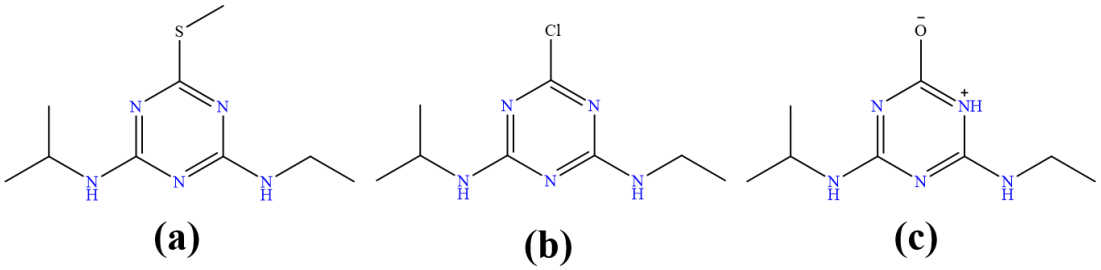
Essa preocupação não fica restrita apenas ao bioma Cerrado, uma vez que a degradação das s-triazinas no solo, água e sedimentos, com ênfase particular nas substâncias formadas ao longo das vias degradativas e aos aspectos de biodegradação, têm recebido considerável atenção há pelo menos três décadas (ERICKSON; LEE; SUMNER, 1989). A estrutura cristalográfica da ametrina (CARVALHO *et al*., 2019)

[10] O *Cambridge Structural Database* (CSD) contém um registro completo de todas as estruturas cristalinas de pequenas moléculas orgânicas e metal-orgânicas já publicadas. O banco de dados está em operação há mais de 50 anos e continua a ser o principal meio de compartilhamento de dados de química estrutural (GROOM *et al*., 2016).

[11] O AGROFIT WEB online é uma ferramenta de consulta ao público, composta por um banco de dados de todos os produtos agrotóxicos e afins registrados no Ministério da Agricultura, Pecuária e Abastecimento (MAPA), com informações do Ministério da Saúde (ANVISA) e informações do Ministério do Meio Ambiente (IBAMA).

e atrazina (LE *et al*., 2016) já estão reportadas na literatura, porém ainda não há nenhum registro de estruturas cristalográficas de algum de seus intermediários.

Figura 3 – Estrutura bidimensional da ametrina (a), da atrazina (b) e do produto de degradação em estudo (c).

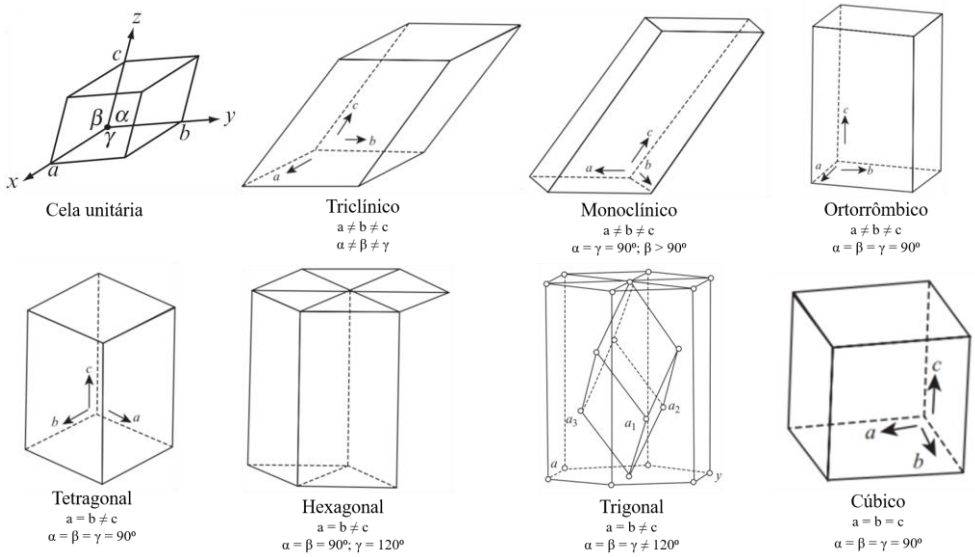


Fonte: Elaborada pelo autor.

3.3 Conhecimento estrutural a partir da difração dos raios X por pequenas moléculas

Os cristais são sólidos homogêneos que possuem ordem de longo alcance, ou seja: simetria e periodicidade. Em um cristal o arranjo dos átomos em torno de qualquer ponto é idêntico ao arranjo em outro ponto qualquer equivalente no mesmo cristal. Nessa definição desconsidera-se os possíveis defeitos locais. Uma vez que muitas propriedades físico-químicas e bioquímicas estão diretamente relacionadas à estrutura cristalina, torna-se essencial a sua elucidação, sendo a cristalografia uma das metodologias mais completas para descrição de compostos que se encontram no estado cristalino (TILLEY, 2006). Um cristal pode ser compreendido como constituído por um número extremamente grande de paralelepípedos tridimensionais imaginários chamados de cela unitária. Ela é definida por três vetores (os eixos de cristal) não-planares: **a**, **b** e **c**, que possuem entre si ângulos α (entre **b** e **c)**, β (entre **a** e **c)** e γ (entre **a** e **b**) (GLUSKER; TRUEBLOOD, 2010). A comutação entre essas dimensões, por simetria, geram sete métricas, ou sistemas cristalinos diferentes: triclínico, monoclínico, ortorrômbico, tetragonal, trigonal, hexagonal e cúbico (CULLITY, 1978) (Figura 4). Esses sete sistemas cristalinos combinados com os quatro tipos de centragem de rede (primitiva, face centrada em A, B ou C, face centrada e corpo centrado) geram então as 14 redes de Bravais (GLUSKER; TRUEBLOOD, 2010).

Figura 4 – Representação de uma cela unitária (genérica) e de seus diagramas quando aplicadas as restrições de simetria; assim são gerados os sete sistemas cristalinos.



Fonte: Elaborada pelo autor.

Os elementos de simetria (centro de simetria, planos de espelho, planos de deslizamento, eixos de rotação e eixos de parafuso) combinados com as 14 redes de Bravais geram as 230 combinações únicas dos elementos de simetria (os 230 grupos espaciais), ou seja, no estado cristalino, a matéria pode se arranjar de 230 formas distintas (MÜLLER; WONDRATSCHEK, 2005). Dadas essas características, é natural que surja a pergunta sobre qual o tipo de técnica física é útil ao estudo dos compostos cristalinos. Essa análise é comumente baseada no fenômeno de difração causado pela interação da matéria com os raios X (GIACOVAZZO *et al*., 2011).

O trabalho de identificação dos aspectos estruturais através da cristalografia de raios X envolve as seguintes etapas: (1) coleta e processamento dos dados, etapa em que uma amostra cristalina é levada ao difratômetro e submetido a um feixe monocromático de raios X, registrando o padrão de difração; (2) resolução da estrutura, onde tem-se a obtenção da fase de cada reflexão medida, de forma que seja possível a construção de mapas de densidade eletrônica. (3) refinamento, etapa em que se faz um trabalho interativo de interpretação desses mapas de densidade eletrônica até que se tenha uma estrutura preliminar; (4) validação e análise do modelo cristalográfico, onde se avalia a coerência química da estrutura e (5) depósito

da estrutura, por meio de um arquivo digital no banco de dados apropriado (GIACOVAZZO *et al*., 2011).

Os raios X são uma radiação eletromagnética (λ = aproximadamente 0,1-100 Å; para a prática cristalográfica: 0,6-3,0 Å) que podem ser produzidos, entre outras, pela desaceleração repentina de elétrons que se movem em um material alvo através do impacto desses elétrons em um alvo metálico (anodo). Dentre os muitos fatores que afetam a intensidade dos raios X no padrão de difração, aquele que depende apenas da estrutura do cristal é chamado de fator de estrutura e pode ser expresso em termos do conteúdo de uma única célula unitária como:

$$F(hkl) = \sum_{j=1}^{N} f_j \exp\left[2\pi \mathrm{i}\left(hx_j + ky_j + lz_j\right)\right] \qquad (1)$$

onde o j-ésimo átomo é dado pelas coordenadas fracionárias$\left(x_j, y_j, z_j\right)$, possuindo um fator de espalhamento f_j havendo *N* átomos na cela unitária. Cada fator de estrutura representa um feixe difratado que tem uma amplitude, $|F(hkl)|$, e uma fase relativa $\phi(hkl)$. As medidas experimentais consistem na intensidade de cada feixe e sua posição no padrão de difração. Após a aplicação de fatores de correção adequados, as quantidades registradas são h, k, l , $|F(hkl)|$ ou h, k, l , $|F(hkl)|^2$

Uma imagem da estrutura cristalina pode ser calculada a partir do padrão de difração de raios X. Como são os elétrons que espalham os raios X, são os elétrons que vemos na imagem, dando o valor da densidade eletrônica (ρ) em cada ponto de uma única célula unitária do cristal:

$$\rho(xyz) = \frac{1}{V}\sum_{hkl} F(hkl) \exp[-2\pi \mathrm{i}(hx + ky + lz)] \qquad (2)$$

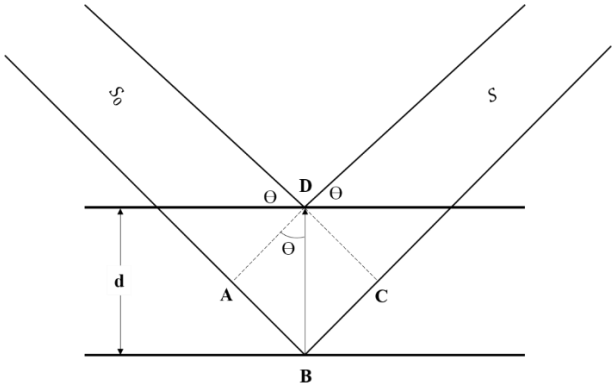
onde *V* é o volume da célula unitária. Observe que os fatores de estrutura incluem as fases $\phi(hkl)$ e não apenas as amplitudes medidas experimentalmente $|F(hkl)|$. Deste modo, se a amplitude do fator de estrutura, $|F(hkl)|$, e a fase relativa dos

ângulos, $\phi(hkl)$, (para cada *h*, *k*, *l*) fossem conhecidas, poderíamos calcular $\rho(xyz)$ para todos os valores de *x*, *y* e *z* e plotar os valores obtidos para fornecer um mapa tridimensional de densidade eletrônica. Entretanto, geralmente podemos obter apenas o $|F(hkl)|$ e não as $\phi(hkl)$ diretamente de medições experimentais. Esse é o problema da fase. Uma vez que os raios X são difratados de todo o cristal, o cálculo produz o conteúdo médio da célula unitária sobre todo o cristal e não o conteúdo de qualquer célula individual. (GLUSKER; TRUEBLOOD, 2010).

Matematicamente, existe uma relação entre as equações (1) e (2). A Eq. (1) transforma a densidade eletrônica (na forma de fatores de espalhamento atômico, f_j) nos fatores de estrutura $F(hkl)$, enquanto a Eq. (2) transforma os fatores de estrutura de volta na densidade eletrônica. Em outras palavras, o fator de estrutura é a transformada de Fourier da densidade eletrônica e a densidade eletrônica é a transformada de Fourier inversa do fator de estrutura, evidenciado pelo sinal negativo da exponencial. Esta é uma descrição matemática da formação da imagem por uma lente. A luz espalhada por um objeto (transformada de Fourier) é coletada por uma lente e focada em uma imagem (transformada inversa). No caso ótico, a imagem (real) é invertida e isso é visto matematicamente pelo aparecimento do sinal negativo no expoente da Eq. (2).

Esse fenômeno de difração de raios X por cristais pode ser considerado em termos de uma teoria análoga à da difração por grades (experimento da fenda dupla). Em 1912 W. L. Bragg descreveu as condições geométricas sob as quais um feixe difratado pode ser observado. Na Figura 5 estão representados raios difratados de planos de rede e, para obter interferência construtiva, a diferença de caminho deve ser um número inteiro de comprimentos de onda (BLAKE; CLEGG, 2009).

Figura 5 – Representação para as condições de difração dos *raios X* para dois planos da rede cristalina pertencentes a mesma família de planos (*h*,*k*,*l*). d é a distância interplanar.



Fonte: Elaborada pelo autor.

Θ é o ângulo entre o feixe primário e a família de planos de índices *hkl*. A diferença no caminho entre as ondas em D e B é igual a $AB + BC = 2dsen\Theta$. Isso leva à lei de Bragg, Eq. (3), que é expressa como:

$$2d_{hkl}sen\Theta = n\lambda \quad (3)$$

Onde *n* é um número inteiro; λ é o comprimento de onda dos raios X, elétrons, prótons e nêutrons em movimento; d é o espaçamento entre os planos na rede atômica, e θ é o ângulo entre o raio incidente e os planos de espalhamento.

3.4 Relação entre conhecimento estrutural e propriedades físico-químicas e biológicas

O estudo das relações quantitativas entre a estrutura química e a atividade biológica (QSAR) ou alguma propriedade físico-química nos permite entender e explicar o mecanismo de ação de compostos em nível molecular, auxiliando no desenvolvimento de novos materiais com propriedades biológicas desejáveis (MARTINS; FERREIRA, 2013). Assim, o uso de modelos *in silico* para prever o potencial bioativo de compostos diminui o tempo e a necessidade de testes experimentais e, consequentemente, o custo relacionado a pesquisa (BANERJEE *et al*., 2018). Nesse contexto, a construção de modelos matemáticos que abordem uma relação qualitativa entre a estrutura molecular e a presença ou ausência de uma determinada atividade biológica, ou a capacidade de modelar tal atividade, com base

na estrutura molecular de um composto, apoiada por meio de técnicas estatísticas multivariadas se mostra muito útil (CONSONNI; TODESCHINI, 2010)

Recentemente, os modelos de *machine learning* provaram ser bem-sucedidos no prognóstico de doenças humanas (KOUROU *et al*., 2015; LIBBRECHT; NOBLE, 2015), na predição de propriedades eletrônicas de compostos (MONTAVON *et al*., 2013) e na modelagem de novos materiais e moléculas (BARTÓK et al., 2017). Além disso esses métodos também são capazes de encontrar soluções para questões ambientais, como gerenciamento de resíduos sólidos municipais (ABBASI; EL HANANDEH, 2016), propriedades do solo/composto com biodisponibilidade e avaliação de risco (WU *et al*., 2013) e para modelar a biodisponibilidade de contaminantes (CIPULLO *et al*., 2019).

A toxicidade de uma substância é governada por suas propriedades, que são determinadas pela sua estrutura química, este é cerne da modelagem estrutura-toxicidade (CONSONNI; TODESCHINI, 2010). Entretanto, as estruturas químicas, geralmente, não contêm de forma explícita as informações relacionadas à atividade. Vários descritores moleculares[12], projetados racionalmente, acentuam diferentes propriedades químicas implícitas na estrutura molecular. Somente essas propriedades podem se correlacionar diretamente com a atividade, variando desde aspectos físico-químico e quânticos a características geométricas e topológicas (DUDEK; ARODZ; GALVEZ, 2006).

Um descritor molecular é o resultado de um procedimento lógico e matemático que transforma informações químicas codificadas em um número útil ou o resultado de algum experimento padronizado. Em geral, não se sabe qual descritor será melhor, então a tendência é usar um grande número deles, esperando que dentre eles existam as variáveis apropriadas e, posteriormente aplicar uma técnica de seleção de variáveis. Atualmente há mais de 3000 descritores que podem ser utilizados. Como exemplos de descritores tem-se: contagem de grupos funcionais e

[12] Os descritores moleculares são divididos em duas classes principais: *medidas experimentais*, como log P, refratividade molar, momento dipolar, polarizabilidade; *propriedades físico-químicas e descritores moleculares teóricos*, que são derivados de uma representação simbólica da molécula e podem ser classificados de acordo com os diferentes tipos de representação molecular. A diferença fundamental entre os descritores teóricos e os medidos experimentalmente é que os descritores teóricos não contêm nenhum erro estatístico devido ao ruído experimental, ao contrário das medições experimentais (CONSONNI; TODESCHINI, 2010).

fragmentos de subestruturas, características estruturais e parâmetros geométricos da molécula, energias de interações etc. (ZANNI *et al*., 2019).

Um fator importante no processo de modelagem é a escolha do conjunto molecular a ser utilizado, observando-se sempre a finalidade a qual se destina o modelo e a disponibilidade de dados experimentais. Durante muito tempo, esses modelos foram desenvolvidos para conjuntos de moléculas com uma estrutura básica comum e diferentes grupos de substituintes, logo após o interesse se estende para moléculas pertencentes a diferentes classes químicas devido a necessidade de obter relações mais gerais e explorar o enorme potencial dos grandes conjuntos de dados disponíveis (BOSC *et al*., 2019).

A literatura é repleta de estudos sobre a SAR, a seguir serão apresentados alguns exemplos desses estudos, com destaque para as características estruturais que desencadeiam atividade pesticida. Sabe-se que a introdução de halogênios em um ingrediente ativo é uma ferramenta importante para modular as propriedades de novos compostos para proteção de culturas. Uma pesquisa mostra que desde 2010, cerca de 96% dos pesticidas lançados no mercado contêm algum átomo do grupo dos halogênios. Dentre estes, os nematicidas contêm o maior número de halogênios, seguidos por inseticidas/acaricidas, herbicidas e fungicidas. Nesse contexto, fungicidas e herbicidas contêm, na maioria dos casos, átomos de flúor, enquanto nematicidas e inseticidas contêm, na maioria dos casos, átomos de halogênio "mistos", por exemplo, cloro e flúor (JESCHKE, 2017).

Em um outro exemplo, em relação a um conjunto de compostos organofosforados, foram estudadas características estruturais e quânticas que influenciam o efeito tóxico. O resultado mostrou que a toxicidade desses compostos depende fortemente da sua estabilidade e das forças de ligação entre o átomo de fósforo e seus substituintes. Quanto mais forte for a ligação do grupo de saída ao átomo de fósforo mais favorável é ao efeito tóxico, pois sua reatividade depende da carga positiva no átomo de fósforo do grupo fosforil polarizado (PAUKKU; HILL, 2012).

Por fim, El-Gamal e colaboradores (2018) relatam o desenvolvimento de um modelo QSAR para explorar os requisitos estruturais que controlam as atividades antimicrobianas de derivados de quinolinas. Os estudos de docagem molecular foram realizados usando o complexo bacteriano de DNA girase e a ciprofloxacina como

ligante cocristalizado. De modo geral, notou-se que nos derivados estudados o anel quinolínico está envolvido tanto em interações do tipo *π-stacking*[13] quanto em interações hidrofóbicas, enquanto átomos de flúor, grupo carbonil, grupo amino e grupo hidroxil formam ligações de hidrogênio em diferentes partes da enzima. Finalmente, a necessidade de dar respostas eficazes aos problemas ecológicos e de saúde impulsionam a busca em direção a novas fronteiras, onde matemática, estatística, química, biologia e suas inter-relações possam produzir novos conhecimentos (CONSONNI; TODESCHINI, 2010).

3.5 Impactos da modificação do estado sólido para o uso sustentável de agroquímicos

As descobertas e novos desenvolvimentos na área do estado sólido proporcionaram uma melhoria para esse campo do conhecimento que se deve, em grande parte pela busca por novos e melhores materiais e para que isso possa ser feito, é necessário possuir a compreensão a respeito da estrutura cristalina e das forças que mantem os cristais unidos (SMART; MOORE, 2012). Essa compreensão leva à síntese de cristais moleculares funcionais; área do conhecimento referida como engenharia de propriedade de cristais. Nas últimas três décadas, ela progrediu da análise de estruturas cristalinas, em termos de interações intermoleculares, para a otimização e construção de cristais com topologias e propriedades desejadas (DESIRAJU, 2017).

Esse conhecimento foi utilizado por pesquisadores que ao empregarem essa técnica como ferramenta de *design* obtiveram cinco novos cocristais do ativo etil-tiofanato, demonstrando que moléculas que contêm grupos tioamidas são promissoras na obtenção desses compostos (NAUHA; NISSINEN, 2011). Outra abordagem que exemplifica a aplicação dos conhecimentos do estado sólido é o reportado acerca das investigações dos mecanismos de transformação fotoquímica indireta, persistência ambiental e ecotoxicidade do fenpiclonil usando métodos de química quântica e toxicologia computacional. Esse fungicida deixa resíduos no solo e pode contaminar as águas subterrâneas, aumentado a preocupação sobre os potenciais efeitos adversos desses contaminantes à saúde humana e ao ecossistema.

[13] Termo que designa interações entre elétrons π de anéis aromáticos em diferentes moléculas (DENG *et al*., 2020).

Os cálculos da cinética de degradação mostram que as altas temperaturas são mais favoráveis para a degradação inicial do fenpiclonil. Além disso, percebeu-se que os produtos de transformação são formados durante a reação subsequente do intermediário aduto-OH IM10. A toxicidade aquática desses produtos diminui, mas não se extingue, com o processo de degradação (YANG *et al*., 2020)

Outra abordagem é a tentativa de melhorar a eficácia de agroquímicos já comercializados, como a do herbicida ametrina. Foram desenvolvidos três novos sais a partir da diversificação de sua forma sólida. Através do entendimento da estrutura cristalina desses sais percebeu-se que o sínton formado entre a s-triazina e grupos carboxila leva ao surgimento de regiões hidrofílicas no arranjo supramolecular e que isso melhora a solubilidade do herbicida de 10 a 20 vezes, quando comparado a ametrina pura. Outra vantagem que advém desses sais é sua alta estabilidade térmica o que sugere uma redução no impacto ambiental deste herbicida (CARVALHO *et al*., 2019). Além disso, os chamados líquidos iônicos (HIL) também tem sido bastante explorado. Nos HIL é possível combinar a atividade herbicida de um ânion com a atividade fungicida, antibacteriana e/ou herbicida do cátion. Dessa maneira, por exemplo, foram obtidos HIL com base no herbicida dicamba (Ácido 3,6-dicloro-2-metoxibenzoico). Esses novos compostos demonstram estabilidade térmica aumentada, volatilidade reduzida e eficácia substancialmente melhorada em relação ao ácido livre de dicamba, além de diminuírem ou eliminarem os impactos negativos, causados por escoamento, deriva de vapor e a necessidade de usar concentrações mais altas, do ácido livre de dicamba no meio ambiente, (COJOCARU *et al*., 2013)

Por fim, destaca-se que quando um agroquímico vai ser aplicado é preciso garantir uma dosagem eficiente, reduzida que traga menor impacto possível pela permanência do composto aplicado ou dos seus produtos de degradação, em sua maioria desconhecidos. Novos pesticidas, com aspectos sustentáveis, podem advir do repensar da formulação de compostos já disponíveis no mercado; o que na prática refere-se à modificação e seleção de forma sólida apropriada para o ingrediente ativo (JESCHKE, 2018).

4 RESULTADOS

Essa tese é composta por quatro artigos[14] e um manuscrito, resultantes da perspectiva abordada no tópico 3.1 (pág. 19). O primeiro deles, intitulado *Structural comparison of five new halogenated dihydroquinoline- 4(1H)-ones* (VAZ *et al*., 2020c), publicado no periódico *Journal of Molecular Structure*, apresenta a elucidação estrutural de cinco dihidroquinolinas halogenadas, bem como a discussão sobre o efeito na conformação molecular do tipo e posição do átomo de halogênio substituintes nos anéis aromáticos. O empacotamento cristalino das cinco moléculas é regido principalmente pelas interações do tipo C—H···O e C—H···halogênio que formam dímeros e cadeias. A mudança de posição e o tipo de halogênio no anel C influenciam na conformação dos compostos estudados, sendo que o empacotamento deles é mais suscetível a variações quando a posição do halogênio muda. Dado que quinolinas foram anteriormente apontadas com atividade antifúngica, a caracterização estrutural de novas quinolinas e seus análogas é imperativo para os ensaios futuros que confirmem sua efetividade como agentes pesticidas. Este artigo resulta da minha estadia de 6 meses no grupo da Prof. Allen Grayson Olliver (*University of Notre Dame*). Dessa maneira, experimentei os benefícios e os efeitos de sinergia de uma colaboração internacional e recebi um treinamento intensivo em medições de difração de raios X em pequenas moléculas. Minha contribuição para este trabalho foi a proposta e estruturação do mesmo e o estudo cristalográfico.

O segundo artigo, intitulado *Dihydroquinoline derivative as a potential anticancer agent: synthesis, crystal structure, and molecular modeling studies* (VAZ *et al*., 2020a), publicado no periódico *Molecular Diverstiy*, apresenta o composto (*E*)-2-(4-clorofenil)-3-(3-nitrobenzilideno)-1-(fenilsulfonil)-2,3-dihidroquinolin-4(1*H*)-ona, referenciado pelo acrônimo M-CNP, como um potencial agente anticâncer. Essa identificação foi obtida através de metodologia *in sílico* onde os resultados de docagem molecular mostraram uma alta afinidade dos grupos NO_2 e 4-clorofenil, bem como dos anéis benzênicos à resíduos na enzima aldeído desidrogenase (ALDH 1A1). Finalmente, testes *in vitro* foram conduzidos para três

[14] Optou-se por não incorporar os materiais suplementares nesta tese, no intuito de apresentar um texto mais enxuto. Entretanto, os mesmos podem ser obtidos na *web page* dos respectivos artigos.

linhagens de câncer: VERO (*normal human skin keratinocytes*), B16F10 (*mouse melanoma*), e MDA-MB-231 (*metastatic breast adenocarcinoma*), confirmando o M-CNP como um candidato a agente anticâncer. Esse estudo indica: o potencial biológico da dihidroquinolina; apresenta a qualidade da modelagem envolvida; a desenvoltura na discussão do estudo estrutural em pequenas moléculas e, embora o composto não tenha sido estudado para uma aplicação agroquímica, esta perspectiva será explorada posteriormente em um outro artigo. Nesse trabalho minha contribuição direta foi a sua proposta e estruturação, a descrição cristalográfica e o cálculo teórico das propriedades moleculares.

O terceiro artigo, intitulado *Machine learning prediction of the potential pesticide applicability of three dihydroquinoline derivatives: syntheses, crystal structures, and physical properties* (VAZ *et al*., 2020b) publicado no periódico *Journal of Molecular Structure*, apresenta um modelo de *machine learning* aplicado na busca de novos agentes agroquímicos e que levou a síntese de três dihidroquinolinas com potencial atividade antiviral (Vírus do mosaico do tabaco) e antifúngica (*Fusarium oxysporum*). Além dos testes *in silico* são apresentados o estudo estrutural, a caracterização espectroscópica e os cálculos de propriedades eletrônicas da (*E*)-3-(4-clorobenzilideno)-2-(4-etóxifenil)-1-(fenilsulfonil)-2,3-dihidroquinolin-4(1*H*)-ona, (*E*)-2-(4-clorofenil)-3-(3-nitrobenzilideno)-1-(fenilsulfonil)-2,3-dihidroquinolin-4(1*H*)-ona e da (*E*)-3-(4-chlorobenzylidene)-2-(2-chlorophenyl)-1-(phenylsulfonyl)-2,3-dihydroquinolin-4(1*H*)-one. Neste artigo intensifiquei minhas experiências em análise estrutural de pequenas moléculas, além de aplicar os conhecimentos em química teórica no entendimento de propriedades eletrônicas dos compostos. Nesse estudo fui responsável pela proposta e estruturação, pelo estudo cristalográfico e pelos cálculos teóricos de propriedades eletrônicas.

No quarto artigo, intitulado *In silico-driven identification and structural analysis of nitrodihydroquinolinone pesticide candidates with antifungal activity*, publicado no *Journal of Molecular Structure* (VAZ *et al*., 2021), duas moléculas (isômeras) apresentaram, através da relação quantitativa estrutura/atividade e de escaneamento virtual, potencial atividade antifúngica (*Aspergillus niger e Sclerotinia sclerotiorum*). Um desses isômeros é a molecula estudada no segundo artigo (pág. 44). A análise dos mapas de probabilidades gerados para *Aspergillus niger* mostram que todos os átomos, com exceção do grupo NO_2 contribuem para a atividade

antifúngica, enquanto para *Sclerotinia sclerotiorum* os grupos NO_2 e carbonil, juntamente de um anel aromático são os responsáveis pela maior contribuição à atividade. Os testes *in vitro* (*Aspergillus niger)* confirmam a predição do modelo *in sílico*, de que os isômeros apresentam atividade antifúngica. Nesse artigo pode-se ver o amadurecimento científico da proposta dessa tese. Começou-se com o entendimento estrutural das dihidroquinolinas, e em seguida seu potencial biológico foi testado. Posteriormente os modelos *in silico* foram aplicados para predição de potencial pesticida e, por fim, nesse artigo, todo esse entendimento foi aplicado e validado pelos testes *in vitro* que corroboram a predição teórica. Outro diferencial é que esse artigo representa uma colaboração interna no PPG-RENAC com a professora Solange Xavier dos Santos e o doutorando Lucas Leonardo da Silva que foram os responsáveis pelos ensaios citotóxicos. Neste artigo fui responsável pela proposta, estruturação e pelo estudo cristalográfico.

Por fim, numa abordagem estratégica, é realizado a caracterização estrutural do metabólito, produto de degradação dos herbicidas ametrina e atrazina. Apesar de serem estruturalmente diferentes daqueles apresentados, a investigação da hidroxi-atrazina demonstra aspectos importantes para aplicação sustentável destes compostos. Estruturalmente eles se apresentam na forma zwitteriônica[15] de sua configuração tautomérica[16] (ceto-enol). Isso permite inferir uma alta mobilidade quando presente no solo, além de uma alta estabilidade, o que pode explicar a persistência destes herbicidas no ambiente.

[15] Um zwitterion é um composto quimicamente neutro, embora seja composto por um par de íons, onde o cátion e o ânion estão covalentemente ligados, é conhecido por ser um tipo de sal (OHNO; YOSHIZAWA-FUJITA; KOHNO, 2018).

[16] Tautomeria é um fenômeno que ocorre quando duas moléculas são diretamente interconvertíveis por causa da mobilidade de um átomo de hidrogênio, que pode se mover de uma posição para outra na molécula com o rearranjo de uma ligação dupla. Um exemplo comum disso é o tautomerismo ceto-enol, pelo qual um átomo de hidrogênio ligado ao átomo de carbono de um composto de carbonila se move para o átomo de oxigênio carbonila. Os tautômeros existem em equilíbrio no estado líquido ou em solução, mas não no estado sólido, e a posição de equilíbrio para o tautomerismo depende da estrutura molecular (BENFENATI *et al*., 2007).

Journal of Molecular Structure 1219 (2020) 128559

Contents lists available at ScienceDirect

Journal of Molecular Structure

journal homepage: http://www.elsevier.com/locate/molstruc

Structural comparison of five new halogenated dihydroquinoline-4(1*H*)-ones

Wesley F. Vaz [a, b, *], Lidiane J. Michelini [c], Gerlon A.R. Oliveira [c], Luciano M. Lião [c], Caridad N. Perez [c], Allen G. Oliver [d], Hamilton B. Napolitano [a, e, **]

[a] Ciências Exatas e Tecnológicas, Universidade Estadual de Goiás, Anápolis, GO, 75001-970, Brazil
[b] Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso, Lucas do Rio Verde, MT, 78455-000, Brazil
[c] Instituto de Química, Universidade Federal de Goiás, Goiânia, GO, 74690-900, Brazil
[d] Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame, IN, 46556, USA
[e] Laboratório de Novos Materiais, Centro Universitário de Anápolis, Anápolis, GO, 75083-515, Brazil

ARTICLE INFO

Article history:
Received 3 April 2020
Received in revised form
16 May 2020
Accepted 28 May 2020
Available online 2 June 2020

Keywords:
Dihydroquinoline
Chlorine
Fluorine
Crystal structure

ABSTRACT

Compounds with dihydroquinoline-4(1*H*)-one nuclei have been reported in the literature for being important in the development of medicines due to their broad spectrum of activities. In this way, the structural knowledge of this class becomes relevant for obtaining new materials with desired biological properties. This study presents the structural elucidation of five halogenated dihydroquinolines, as well as the discussion about the effect on the molecular conformation of the type and position of halogen atom on aromatic rings. Compounds **I** and **IV** differ in halogen substitution on 2-phenyl ring, while compounds **III** and **V** differ in halogen substitution on the benzylidene ring. Moreover, compound **II** has a *para*-substituted 2-phenyl ring in their molecular structure. The crystal packing of all five molecules is mainly ruled by C—H···O and C—H···halogen interactions that form dimers and chains. The shift in position and the kind of the halogen in ring C shows a starring role in the conformation of the studied compounds, and the packaging of these compounds is more susceptible to variations when the halogen position changes.

1. Introduction

Quinolines are a class of *N*-heterocyclic compounds, also known as benzopyridines, obtained both from natural and synthetic sources. They have attracted great attention in the scientific community being used in several industrial processes, becoming increasingly important in the development of medicines, pesticides and also due to their notable biological activities [1–3]. Among the compounds of this class, we are interested in those with the dihydroquinoline-4(1*H*)-one moiety; in this way, the recent studies conducted by our research group involve structural elucidation [4–6], anti-cancer properties [7,8] and their potential application as pesticides [9]. Also, other groups have explored the cytotoxic properties of this class [10–13] and as potential Middle East respiratory syndrome coronavirus (MERS-CoV) inhibitors [14].

The biological activity of a substance is dominated by its properties, which are determined by its chemical structure. Structural elucidation is essential since it allows understanding and proposing explanations for the mechanisms of action at the molecular level, helping to design and develop new materials with desirable biological properties [15,16]. A search in the CSD version 5.41 (November 2019) database showed 132 reported structures with the dihydroquinoline-4(1*H*)-one nucleus. Furthermore, given the biological potential of these molecules, a need to enlarge elucidated structures of this class will contribute to the applications of dihydroquinoline-4(1*H*)-ones.

Thus, in this paper, we investigate and report a comprehensive single-crystal analysis of five halogenated dihydroquinoline-4(1*H*)-ones, namely (*E*)-3-(2-chlorobenzylidene)-2-(2-chlorophenyl)-1-(phenylsulfonyl)-2,3-dihydroquinolin-4(1*H*)-one (**I**), (*E*)-3-(2-chlorobenzylidene)-2-(4-chlorophenyl)-1-(phenylsulfonyl)-2,3-

* Corresponding author. Ciências Exatas e Tecnológicas, Universidade Estadual de Goiás, Anápolis, GO, 75001-970, Brazil.
** Corresponding author. Ciências Exatas e Tecnológicas, Universidade Estadual de Goiás, Anápolis, GO, 75001-970, Brazil.
E-mail addresses: wesfonseca@gmail.com (W.F. Vaz), hbnapolitano@gmail.com (H.B. Napolitano).

https://doi.org/10.1016/j.molstruc.2020.128559
0022-2860/

dihydroquinolin-4(1*H*)-one (**II**), (*E*)-2-(2-bromophenyl)-3-(2-chlorobenzylidene)-1-(phenylsulfonyl)-2,3-dihydroquinolin-4(1*H*)-one (**III**), (*E*)-2-(2-chlorophenyl)-3-(2-fluorobenzylidene)-1-(phenylsulfonyl)-2,3-dihydroquinolin-4(1*H*)-one (**IV**) and (*E*)-2-(2-bromophenyl)-3-(2-fluorobenzylidene)-1-(phenylsulfonyl)-2,3-dihydroquinolin-4(1*H*)-one (**V**). The title compounds were divided into two groups: *chlorinated* and *fluorinated*; and their molecular geometry, intermolecular interactions and crystal packing were all analyzed and discussed in terms of the effect of halogen substitution in their scaffold.

2. Experimental

2.1. Synthesis and crystallization

Compounds **I** to **V** were obtained from sulfonamide chalcones reacted with benzaldehydes in an alkaline reaction environment for 24 h (Scheme 1). The subsequent precipitates were purified by slow recrystallization from dichloromethane and ethanol (4:1), after drying at room temperature.

Nuclear magnetic resonance (NMR) spectra were acquired on a Bruker Avance III 500 spectrometer (Rheinstetten, Germany) operating at 11.75 T with a 5 mm inverse detection three-channel (^{1}H, ^{2}H and X-nucleus) BBI probe. The samples (*ca.* 10 mg) were dissolved in 600 μL of deuterated dimethylsulfoxide (DMSO-d_6), containing tetramethylsilane (TMS) as the internal standard. The unambiguous signal assignment was achieved by correlation spectroscopy (COSY), heteronuclear multiple bond correlation (HMBC), heteronuclear single quantum correlation (HSQC), and DEPT-135 experiments, in addition to ^{1}H and ^{13}C analyses (Table 1).

2.2. Crystallographic characterization

Appropriate single crystals of compounds **I**, **II**, **III**, **IV** and **V** were carefully chosen. They were mounted in a Bruker APEX II CCD diffractometer with graphite-monochromated Mo-Kα radiation (λ = 0.71073 Å), and data were measured at 120 K. Using Olex2 [17], the structure solutions were determined by Direct Methods with SHELXS [18] and refined by full-matrix least-squares on F^2 with SHELXL [19]. All the hydrogen atoms were placed in calculated positions and refined with fixed individual displacement parameters [U_{iso}(H) = 1.2U_{eq} (C) or 1.5U_{eq} (C)] according to the riding model (C—H bonds equal 0.93 Å for aromatic). Ring D in **IIIb** was found to be disordered and modeled over two equal occupancy positions. Lastly, the validation of chemical parameters were made using PARST [20] and PLATON [21]. Data collection and structure refinement details are summarized in Table 2. The structures **I** to **V** were deposited in the Cambridge Crystallographic Data Centre (CCDC) with the code number 1994320, 1994319, 1994318, 1994317, and 1994316, respectively.

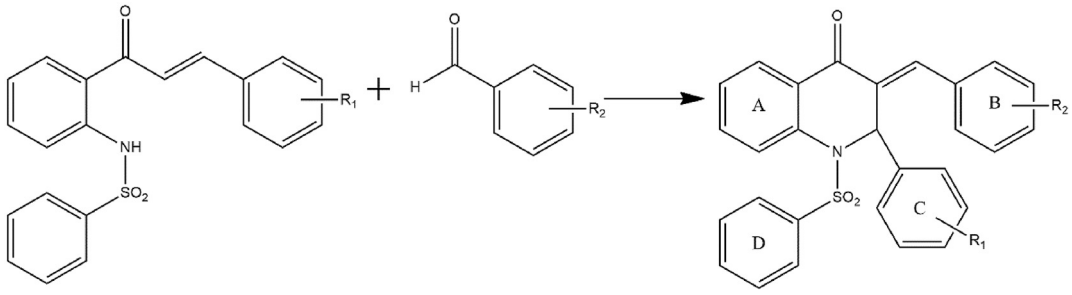
3. Results and discussion

3.1. Spectroscopy characterization

Complete assignment of the signals of ^{1}H and ^{13}C NMR spectra are presented in Table 1 and the infrared spectra are available in the Support Information. The most deshielded carbon signal, near δ 181 and assigned to carbonyl carbon (C3), was a good starting point to confirm the proposed structure and NMR signal assignments. The strong correlations with C3 in the HMBC experiment ($^3J_{CH}$) allowed the identification of hydrogens H8, H5, H10 and H1 (Fig. S1). Also, the correlations of these hydrogens with their respective carbons, by HSQC experiment ($^1J_{CH}$), enabled the assignment of the carbons C10 and C1 (Fig. S2). H10 signal is a broad singlet (Fig. S3) since it presented long-range scalar coupling with the other hydrogens of the benzylidene group. Cross peaks with C1 and H1 in the HMBC spectrum provided information for C22 and C18 assignments (Fig. S1). C9 was identified by its strong correlations with H1, H5, and H7, and weak correlation with H8 (Fig. S1). H16 (and H12 when present) was assigned based on $^3J_{CH}$ with C10 in HMBC.

The multiplet signal near to δ 7.1 in the ^{1}H NMR spectrum was assigned to H24/28, since it had an integration of 2 in all the structures and presented $^1J_{CH}$ and $^3J_{CH}$ correlations only with the carbons C24/28 and C26, respectively. H25/27 was assigned by a cross peak with H24/28 signal in the COSY experiment (Fig. S4), and their HSQC correlation allowed C25/27 assignment as the carbon signal near to δ 129. H26 and C26 correlation were observed in HSQC and this correlation was supported by correlation with H25/27 in the COSY experiment (Fig. S4). H25/27 shared a strong HMBC cross peak with a carbon signal close to δ 135, assigned as C23. C4 and C23 were differentiated by DEPT-135, in the same way as other quaternary carbons close to tertiary carbons. Thus, all the atoms of the phenyl ring attached to the sulfonamide group were assigned.

The $^3J_{CH}$ correlations of H10 and H1 were good entry points for the assignments of the rings B and C signals (C11 to C22). In structures **IV** and **V**, the presence of fluorine made the assignments of ring C easier because the neighboring fluorine carbons ($^1J_{CF}$ ~ 250 Hz; $^2J_{CF}$ ~ 22 Hz; $^3J_{CF}$ ~ 9 Hz; $^4J_{CF}$ ~ 4 Hz) were observed as doublets, due to coupling between C and F. The hydrogen H16 was assigned by cross peaks observed in HMBC experiment, especially the correlation with C10. The signal for C22 was observed near to δ 123.6 for compounds **III** and **V**, which was markedly shielded when compared to **I** and **IV** (near to δ 143.3). This shielding was due to the attached bromine on the former two compounds and chlorine on the latter two compounds. This kind of strategy allowed a



I : R_1 = *o*-Cl; R_2 = *o*-Cl **II** : R_1 = *p*-Cl; R_2 = *o*-Cl **III** : R_1 = *o*-Br; R_2 = *o*-Cl
IV : R_1 = *o*-Cl; R_2 = *o*-F **V** : R_1 = *o*-Br; R_2 = *o*-F

Scheme 1. Synthesis of compounds I to V.

Table 1
^{1}H NMR spectral data assignments for the compounds **I, II, III, IV** and **V** in DMSO-d_6.

Compound	I		II		III		IV		V	
Atom	^{1}H	^{13}C	^{1}H	^{13}C	^{1}H	^{13}C	^{1}H	^{13}C	^{1}H	^{13}C
C1	6.81 (m)	58.4	6.52 (m)	58.9	6.81 (s)	60.6	6.79 (m)	58.6	6.73 (s)	60.7
C2	–	132.1	–	131.3	–	132.3	–	132.1	–	130.0
C3	–	181.2	–	181.1	–	181.2	–	181.2	–	181.1
C4	–	129.3	–	127.4	–	129.6	–	129.3	–	129.2
C5	7.80 (ddd, 7.6; 1.6; 0.5)	127.7	7.80 (ddd, 7.6; 1.6; 0.5)	127.8	7.78 (ddd, 7.6; 1.6; 0.5)	127.6	7.82 (ddd, 7.7; 1.7; 0.4)	127.6	7.80 (dd, 7.6; 1.4)	127.3
C6	7.45 (ddd, 7.6; 7.6; 1.1)	128.5	7.42 (m)	127.8	7.39 (ddd, 7.6; 7.6; 1.1)	128.2	7.46 (m)	128.5	7.45 (ddd, 7.6; 7.6; 1.0)	128.5
C7	7.65 (ddd, 7.6; 7.6; 1.6)	135.3	7.70 (m)	135.4	7.65 (ddd, 7.6; 7.6; 1.6)	135.3	7.65 (m)	135.2	7.65 (m)	135.2
C8	7.41 (m)	128.1	7.65 (m)	127.4	7.37 (ddd, 7.6; 1.1; 0.5)	127.7	7.39 (m)	128.4	7.37 (dd, 8.2; 1.0)	128.6
C9	–	138.1	–	138.1	–	138.1	–	138.1	–	138.0
C10	7.88 (m)	136.2	7.88 (m)	136.1	7.88 (m)	136.2	7.67 (m)	132.4	7.68 (m)	132.4
C11	–	131.1	–	130.8	–	131.2	–	121.0	–	121.0
C12	–	134.5	–	134.9	–	134.5	–	160.7	–	160.5
C13	7.73 (dd, 8.0; 1.2)	130.5	7.72 (dd, 8.0; 1.4)	130.6	7.73 (dd, 8.0; 1.1)	130.5	7.45 (m)	116.5	7.46 (dd, 7.6; 1.1)	116.5
C14	7.54 (m)	132.1	7.56 (ddd, 8.0; 7.70; 1.6)	132.2	7.54 (ddd, 8.0; 7.6; 1.1)	132.1	6.85 (td, 7.7; 1.7)	129.1	6.85 (dd, 7.6; 7.6; 1.5)	129.2
C15	7.35 (m)	127.6	7.40 (m)	127.8	7.44 (ddd, 7.8; 7.6; 1.1)	128.5	7.28 (td, 7.7; 1.2)	125.0	7.28 (ddd, 7.6; 7.6; 1.1)	125.1
C16	6.72 (dd, 7.9; 1.5)	128.9	7.00 (dd, 7.9; 1.6)	128.8	6.72 (dd, 7.8; 1.7)	128.9	7.60 (m)	133.1	7.60 (m)	133.1
C17	–	135.0	–	136.3	–	136.3	–	134.7	–	135.8
C18	6.79 (dd, 7.9; 1.5)	129.8	7.35 (m)	128.9	6.77 (dd, 7.7; 1.5)	130.1	6.76 (dd, 7.7; 1.6)	129.7	6.75 (dd, 7.7; 1.5)	130.1
C19	7.14 (m)	127.5	7.42 (m)	129.5	7.17 (td, 7.7; 1.5)	128.0	7.12 (td, 7.7; 1.4)	127.4	7.15 (td, 7.7; 1.5)	127.9
C20	7.35(m)	130.8	–	133.4	7.25(ddd, 7.9; 7.7; 1.5)	131.0	7.33 (td, 7.7; 1.6)	130.8	7.23 (ddd, 7.9; 7.7; 1.5)	130.9
C21	7.60 (dd, 7.9; 1.2)	131.0	7.42 (m)	129.5	7.77 (dd, 7.9; 1.2)	134.5	7.60 (m)	131.0	7.76 (dd, 7.9; 1.2)	134.6
C22	–	133.3	7.35 (m)	128.9	–	123.6	–	133.4	–	123.7
C23	–	135.6	–	136.0	–	135.4	–	135.6	–	135.5
C24	7.13 (m)	127.4	7.16 (m)	126.6	7.13 (m)	127.6	7.05 (m)	127.7	7.04 (m)	127.5
C25	7.41 (m)	129.2	7.42 (m)	129.3	7.40 (m)	129.1	7.39 (m)	129.0	7.38 (m)	129.0
C26	7.64 (m)	134.0	7.66 (m)	134.1	7.63 (m)	134.0	7.66 (m)	134.1	7.65 (m)	134.1
C27	7.41 (m)	129.2	7.42 (m)	129.3	7.40 (m)	129.1	7.39 (m)	129.0	7.38 (m)	129.0
C28	7.13 (m)	127.4	7.16 (m)	126.6	7.13 (m)	127.6	7.05 (m)	127.7	7.04 (m)	127.5

Table 2
Crystallographic and refinement details for **I**, **II**, **III**, **IV** and **V**.

	I	**II**	**III**	**IV**	**V**
Chemical formula	$C_{28}H_{19}Cl_2NO_3S$	$C_{28}H_{19}Cl_2NO_3S$	$C_{28}H_{19}BrClNO_3S$	$C_{28}H_{19}ClFNO_3S$	$C_{28}H_{19}BrFNO_3S$
M_r	520.40	520.40	564.86	503.95	548.41
Crystal system, space group	Triclinic, *P*-1	Monoclinic, $P2_1/n$	Triclinic, *P*-1	Triclinic, *P*-1	Monoclinic, $P2_1/c$
Temperature (K)	120	120	120	120	120
a, *b*, *c* (Å)	8.0771 (2), 16.4324 (5), 19.2435 (6)	9.8048 (4), 14.2198 (6), 17.2824 (7)	8.0141 (3), 16.6545 (7), 19.3781 (8)	7.9709 (11), 21.980 (3), 24.471 (3)	29.026 (5), 16.304 (3), 15.322 (2)
α, β, γ (°)	109.114 (2), 94.537 (1), 97.435 (1)	90, 95.552 (1), 90	108.605 (1), 95.549 (1), 97.078 (1)	63.784 (2), 81.456 (2), 82.227 (2)	90, 95.775 (3), 90
V (Å^3)	2372.81 (12)	2398.25 (17)	2407.00 (17)	3791.9 (9)	7214 (2)
Z	4	4	4	6	12
μ (mm^{-1})	0.39	0.39	1.94	0.27	1.84
Crystal size (mm)	0.17 × 0.17 × 0.09	0.37 × 0.30 × 0.27	0.19 × 0.18 × 0.14	0.19 × 0.14 × 0.08	0.17 × 0.17 × 0.16
T_{min}, T_{max}	0.675, 0.745	0.683, 0.712	0.594, 0.641	0.951, 0.978	0.940, 0.984
No. of measured, independent and observed [$I > 2\sigma(I)$] reflections	35818, 9758, 6875	36054, 5954, 5242	54875, 11976, 9107	68765, 16965, 12459	194960, 18117, 13854
R_{int}	0.045	0.019	0.037	0.046	0.064
$(\sin\theta/\lambda)_{max}$ (Å^{-1})	0.627	0.668	0.668	0.644	0.670
$R[F^2 > 2\sigma(F^2)]$, $wR(F^2)$, *S*	0.038, 0.089, 1.00	0.031, 0.085, 1.04	0.043, 0.101, 1.04	0.040, 0.096, 1.03	0.043, 0.100, 1.04
No. of reflections	9758	5954	11976	16965	18117
No. of parameters	631	316	661	946	946
$\Delta\rho_{max}$, $\Delta\rho_{min}$ (e Å^{-3})	0.36, −0.42	0.39, −0.38	1.13, −2.61	0.33, −0.36	1.90, −0.84

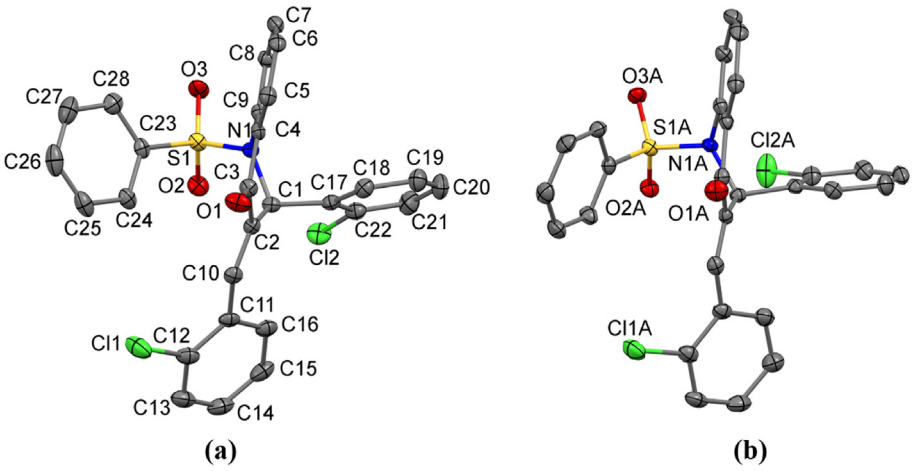


Fig. 1. The two independent molecules of **I** showing the atom-labeling scheme: (a) molecule **Ia**, (b) molecule **Ib**. To clarify, in (b) the labeling scheme shows only non-carbon atoms. The labeling scheme for C atoms in (b) follows the same way as presented in (a). Displacement ellipsoids are drawn at the 50% probability level and H atoms have been omitted.

Table 3
Torsion angles and least-square planes of aromatic rings of **I**, **II**, and **III**.

Molecule	ϕ_1	ϕ_2	ϕ_3	∠AB
Ia	−146.9(3)	145.9(2)	87.46(16)	54.40(7)
Ib	−147.4(2)	168.88(19)	77.35(16)	62.55(7)
II	137.5(2)	167.77(12)	−84.58(13)	60.12(4)
IIIa	−148.9(3)	164.83(2)	−104.43(2)	61.00(9)
IIIb	−147.7(3)	147.28(2)	−87.85(4)	55.62(9)

ϕ_1 = C2−C10−C11−C12; ϕ_2 = C2−C1−C17−C22; ϕ_3 = N1−S1−C23−C28.

complete assignment of the ^{1}H and ^{13}C nuclei (see S.I.), and agrees with our previous report [22].

3.2. Crystallographic characterization

3.2.1. Chlorine dihydroquinolinones

Compound **I** belongs to the class of dihydroquinolinones having three substituents groups in its motif. A sulfonylbenzene group attached to N atom; an *ortho*-chlorobenzene attached to C1 atom and a chloro-2-vinylbenze attached to C2 atom (Scheme 1). Compound **I** crystallizes in the triclinic system ($P\bar{1}$) with two independent molecules in the asymmetric unit (ASU), labeled as **Ia** and **Ib** (Fig. 1). To clarify, in this paper we chose arbitrarily all the molecules within the ASU have *R* configuration about the stereogenic center. To understand the geometrical differences between these molecules their structures have been overlaid using the atoms C1, C3, and C5 as anchor points (Fig. S48). The primary variances noted within these structures are related to the orientations of rings B, C and D with respect to the A ring (defined on Table 3). These variances were measured using the following parameters: the torsion angles: C2−C10−C11−C12 (ϕ1), C2−C1−C17−C22 (ϕ2) and N1−S1−C23−C28 (ϕ3), and the dihedral angle between the planes formed by ring A and ring B (∠AB) (Table 3).

The ∠AB dihedral angle is *ca* 8º larger in **Ib** than in **Ia**, evidencing that the rings in these two molecules have different orientation. This characteristic is also observed in other crystal structures of dihydroquinolinones derivatives [23–25]. Except for compound II, these characteristics are the same for all compounds studied here. The orientation of ring B (ϕ1), could be considered the same in Ia and Ib, it assumes an *anti-clinal* orientation. Furthermore, the values of ϕ2 show that ring C assumes an *anti-clinal* and an *anti-periplanar* orientation in Ia and Ib, respectively, with a difference of 23º. Finally, ϕ3 shows a *syn-clinal* orientation of ring D

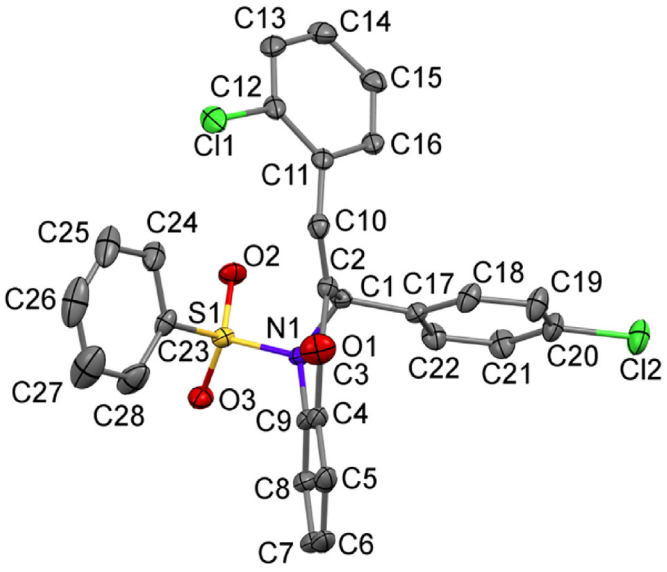


Fig. 2. The molecular structure of **II** showing the atom-labeling scheme. Displacement ellipsoids are drawn at the 50% probability level and H atoms have been omitted.

with a difference of *ca* 10º between the two molecules.

Compound **II,** (Fig. 2), is a positional isomer of **I,** having a chloro-4-vinylbenzene attached to C2 atom (Scheme 1). It is the structure with ∠AB angle smaller than **Ib** and larger than **Ia**. In its molecular structure, unlike **Ia**, ϕ_1 shows that ring B assumes an *anti-clinal* orientation. The change of position of the chlorine atom causes a significant geometric change concerning compounds **I** and **III**. In **II** the rings C and D are oppositely oriented compared with the same rings in the other compounds studied here. The values of ϕ2 indicate similarity in the molecular set and can be divided into two groups, one containing the molecules **Ib**, **II** and **IIIa** and the other containing the molecules **Ia** and **IIIb**. In compound **II**, ϕ2 and ϕ3 assume *anti-periplanar* and a *syn-clinal* orientations, respectively.

Compound **III** differs from **I** only in the presence of the bromo-2-vinylbenzene bonded to the C2 atom (Scheme 1). The two independent molecules in the ASU were labeled as **IIIa** and **IIIb** (Fig. 3). The overlay of these structures indicates a difference of 14º in ∠AB angle for **IIIa** and **IIIb** (Fig. S49). When compounds **I**, **II** and **III** are compared (∠AB angle), it is possible to distinguish two sets

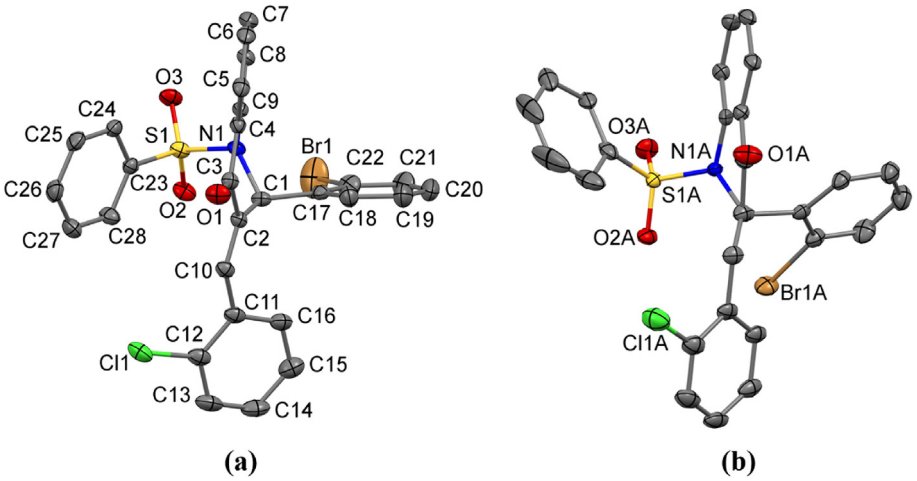


Fig. 3. The two independent molecules of compound **III**, showing the atom-labeling scheme: (a) molecule **IIIa**, (b) molecule **IIIb**. To clarify, in (b) the labeling scheme shows only non-carbon atoms. The labeling scheme for C atoms in (b) follows the same way as presented in (a). Displacement ellipsoids are drawn at the 50% probability level and H atoms have been omitted.

Table 4
Hydrogen-bond geometry (Å, °) of **I**, **II** and **III**.

D—H···*A*	*D*—H	H···*A*	*D*···*A*	*D*—H···*A*
		I		
C14—H14···O2*A*^{i}	0.95	2.46	3.374 (3)	162
C8*A*—H8*A*···O3^{ii}	0.95	2.36	3.138 (3)	138
C13*A*—H13*A*···O2^{i}	0.95	2.38	3.326 (3)	175
C27*A*—H27*A*···O3*A*^{ii}	0.95	2.50	3.175 (3)	128
		II		
C6—H6···O2^{iii}	0.95	2.44	3.3531 (16)	160
		III		
C8—H8···O3*A*^{i}	0.95	2.40	3.161 (3)	137
C13—H13···O2*A*^{iv}	0.95	2.31	3.245 (3)	169
C21—H21···O1*A*^{v}	0.95	2.55	3.458 (4)	159
C25—H25···O3^{i}	0.95	2.50	3.305 (3)	142
C14*A*—H14*A*···O2^{iv}	0.95	2.46	3.385 (3)	165
C19—H19···Br1^{iii}	0.95	2.82	3.626 (3)	143

Symmetry codes: (i) −*x*+1, −*y*+1, −*z*+1; (ii) −*x*+1, −*y*, −*z*+1; (iii) *x*+1, *y*, *z*; (iv) −*x*+1, −*y*+2, −*z*+1; (v) *x*, *y*, *z*+1.

IIIa molecules. The decreasing of ∠AB angle value regarding chlorine dihydroquinolinones is **Ia**, **IIIb**, **II**, **IIIa**, and **Ib** (Table 3). There is no significant difference between ϕ1 in the molecules of compounds **I** and **III**, but this torsion in **II** is, *ca* 10º less twisted. The values of ϕ2 show that ring C assumes an *anti-periplanar* orientation in **IIIa** while in **IIIb** it is *anti-clinal* with a difference of *ca* 17º. ϕ2 values are similar for **Ib**, **II** and **IIIa** and being *ca* 20º smaller for **Ia** and **IIIb**. The orientation of ring D (ϕ3) in **IIIa** is *anti-clinal* while in **IIIb** is *syn-clinal* with a difference *ca* 17º.

The crystal structure of these chlorine dihydroquinolinones is stabilized by C—H···O and C—H···halogen hydrogen bonds listed in Table 4. Although these interactions are weak, we are interested in how they affect the packing and if the different substituents on ring C lead to different arrangements. In the molecular packing in the unit cell of compound **I,** two **Ib** molecules are involved in C27*A*—H27*A*···O3*A* interactions leading to the formation of a $R_2^2(12)$ dimer along the *b* axis. Meanwhile, one **Ia** molecule links to

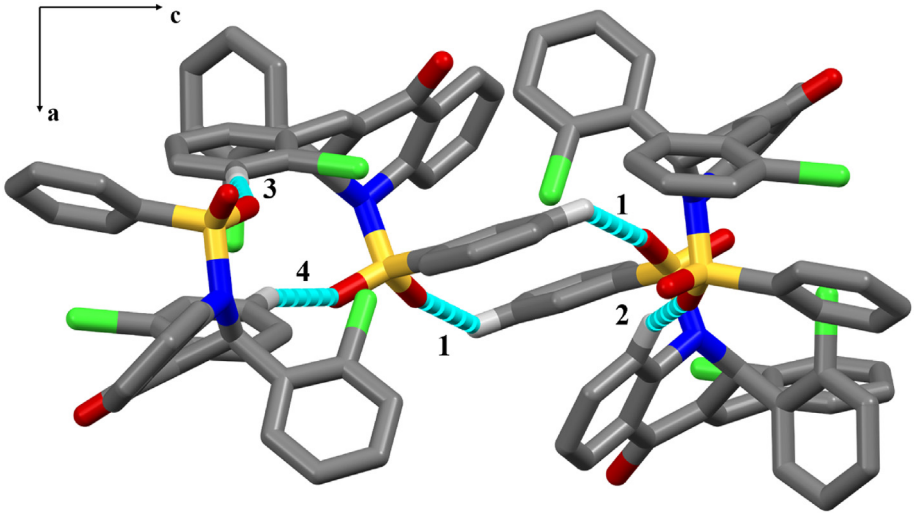


Fig. 4. A partial packing view of **I** showing the dimer formed by C27A—H27A···O3A interaction (1), and the discrete contacts C8A—H8A···O3 (2), C13A—H13A···O2 (3), and C14—H14···O2A (4). For clarity, H atoms not involved in the motif have been omitted.

with similar values: first with **IIIb** and **Ia**, and second with **Ib**, **II**, and these dimer, along the *c* axis, through three $D_1^1(2)$ discrete contacts

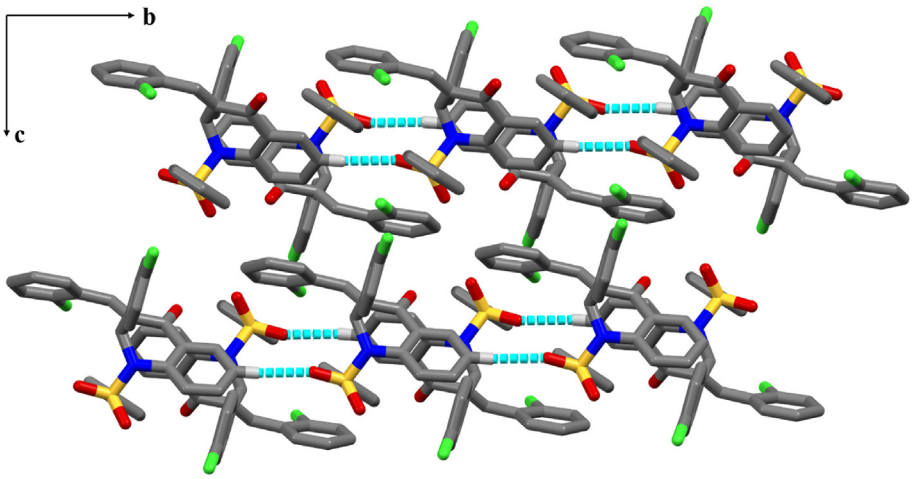


Fig. 5. A partial packing view of **II** showing the chains formed by C6—H6···O2 interaction. For clarity, H atoms not involved in the motif have been omitted.

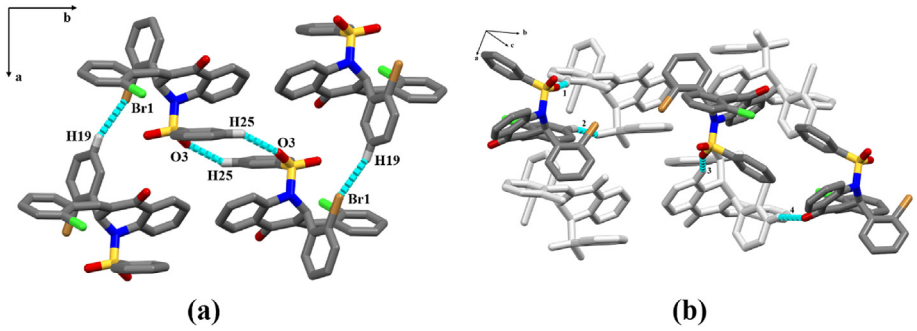


Fig. 6. A partial packing view of **III**, showing the dimer and the chain formed by C25—H25···O3 and C19—H19···Br1 interactions, respectively (a). The discrete contacts C13—H13···O2A (1), C14A—H14A···O2 (2), C8—H8···O3A (3), and C21—H21···O1A (4) join molecules of **IIIa** and **IIIb** in the packing (b); the motif in (a) is presented in (b) in light grey to better visualization. For clarity, H atoms not involved in interactions have been omitted.

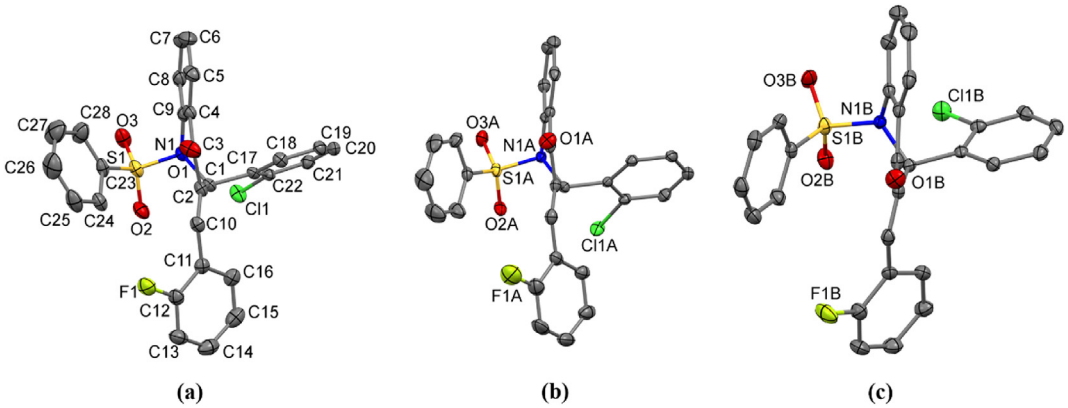


Fig. 7. The three independent molecules of **IV** showing the atom-labeling scheme: (a) molecule **IVa**, (b) molecule **IVb**, and (c) molecule **IVc**. To clarify, in (b) and (c) the labeling scheme shows only non-carbon atoms. The labeling scheme for C atoms in (b) and (c) follow the same way as presented in (a). Displacement ellipsoids are drawn at the 50% probability level and H atoms have been omitted.

[26],: C8*A*—H8*A*···O3, C13*A*—H13*A*···O2, C14—H14···O2*A* (Fig. 4).

Because this is a discrete, localized contact, there is no propagation to an extended structure through this contact. The molecular packing of compound **II** is not as varied as that of compound **I**. This arrangement can be described as $C_1^1(8)$ chains that extend along the *a* axis by molecules that are associated through a C6—H6···O2 interaction (Fig. 5). The molecular packing of compound **III** seems remarkably similar to compound **I**. The differences indicate the carbonyl groups are not related to interactions in the packing of compound **I**. Two molecules of **IIIa** are arranged by C25—H25···O3 interaction as a $R_2^2(12)$ dimer along the *b* axis, and by C19—H19···Br1 interaction as a $C_1^1(6)$ chain along the *a* axis. The molecules of **IIIb** are connected to this arrangement by discrete contacts: C13—H13···O2*A*, C14—H14*A* ···O2, and C—H8···O3*A* along

Table 5
Torsion angles and least-square planes of aromatic rings of **IV** and **V**.

Molecule	ϕ_1	ϕ_2	ϕ_3	∠AB
IVa	−141.7(2)	169.45(17)	91.06(19)	58.77
IVb	−153.9(2)	150.24(17)	72.24(16)	49.30
IVc	−150.1(2)	−178.22(16)	75.28(16)	59.71
Va	−157.7(3)	175.3(2)	78.1(2)	44.53
Vb	−144.4(3)	173.7(2)	72.1(2)	62.80
Vc	−149.7(3)	161.1(2)	82.7(2)	52.79

ϕ_1 = C2−C10−C11−C12; ϕ_2 = C2−C1−C17−C22; ϕ_3 = N1−S1−C23−C28.

The molecular packing in the unit cell of compound **IV** and **V** is stabilized by C−H···O and C−H···halogen hydrogen bonds (Table 6). In compound **IV**, the contacts C5*B*−H5*B*···O3*B*, C18*A*−H18*A*···O2*A*, and C19*A* −H19*A*···Cl1*A* form $C_1^1(7)$ and $C_1^1(6)$ chains along the *a* axis. Molecules of **IVa** are linked through C26*B*−H26*B*···O3 and C15−H15···O3*A* interactions, leading to crystal packing along the *b* axis (Fig. 9a). These motifs (molecules of **IVb** and **IVc)** are attached by the discrete contacts: C14A—H14A···O2B, C14B−H14B···O2A, C27A—H27A···O1, and

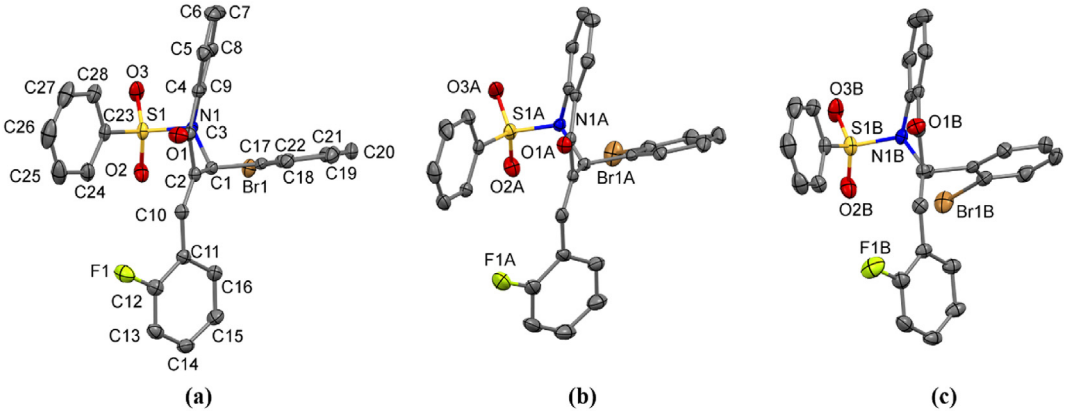


Fig. 8. The three independent molecules of compound **V**, showing the atom-labeling scheme: (a) molecule **Va**, (b) molecule **Vb**, and (c) molecule **Vc**. To clarify, in (b) and (c) the labeling scheme shows only non-carbon atoms. The labeling scheme for C atoms in (b) and (c) follow the same way as presented in (a). Displacement ellipsoids are drawn at the 50% probability level and H atoms have been omitted.

to the *b* axis, while the C21−H21···O1*A* interaction grows the packing along to the *c* axis. (Fig. 6).

3.2.2. Fluorine dihydroquinolinones

Compound **IV** has an *ortho*-chlorobenzene attached to C1 atom and a fluoro-2-vinylbenzene attached to the C2 atom (Scheme 1). The three independent molecules in ASU were labeled as **IVa**, **IVb** and **IVc** (Fig. 7). As listed in Table 5, these three structures present a variation of 10° in their ∠AB angle being the descending order equal to **IVb**, **IVa,** and **IVc** (Fig. S50). The torsion related to ring B (ϕ_1) is *syn-clinal* oriented for **IVb** and **IVc**, and *anti-clinal* oriented for **IVa**, with a variation of *ca* 12° between them. The values for ϕ_2 show that ring C has more mobility than ring B, with a *ca* 28° of variation in this torsion. The orientation of ring D (ϕ_3) varies *ca* 19°, in **IVa** it is almost perpendicular to SO_2 group while in **IVb** and **IVc** it is almost parallel to O2a and O2b atoms, respectively.

Compound **V** is like compound **IV** but has an *ortho*-bromobenzene attached to C1 atom. The three independent molecules in the ASU were labeled as **Va**, **Vb** and **Vc** (Fig. 8). These molecules present a variation of 18° in their ∠AB angle being the descending order equal to **Va**, **Vc** and **Vb** (Fig. S51). When ∠AB angle values for all fluorine dihydroquinolinones are compared, it is observed a decreasing order **Va**, **IVb**, **Vc**, **IVa**, **IVc,** and **Vb** (Table 5). Just like for chlorine dihydroquinolinones, it is not possible to figure out a clear correlation between the substitution of halogens and the AB ring orientation of these fluorine dihydroquinolinones. The torsion ϕ_1 is *anti-periplanar* oriented in **Va** while in **Vb** and **Vc** it is *anti-clinal*. The values of ϕ_2 show that ring C assumes an *anti-periplanar* orientation in all molecules of compound **V** with a difference of *ca* 14°. The torsion ϕ_3 shows ring D in *syn-clinal* orientation in all molecules of compound **V**, with a change of *ca* 10°. These values, when compared with compound **IV**, shows it having more mobility in the molecule of **IVa**, being *ca* 9° larger than in **Vc**.

Table 6
Hydrogen-bond geometry (Å, °) for compounds **IV** and **V**.

D—H···*A*	*D*—H	H···*A*	*D*···*A*	*D*—H···*A*
		IV		
C5*B*—H5*B*···O3*B*^{i}	0.95	2.48	3.104 (2)	123
C14*A*—H14*A*···O2*B*^{ii}	0.95	2.47	3.378 (3)	161
C14*B*—H14*B*···O2*A*^{iii}	0.95	2.58	3.518 (3)	168
C15—H15···O3*A*^{iv}	0.95	2.57	3.463 (3)	156
C18*A*—H18*A*···O2*A*^{iv}	0.95	2.57	3.289 (2)	133
C26*B*—H26*B*···O3	0.95	2.60	3.326 (2)	134
C27*A*—H27*A*···O1^{v}	0.95	2.56	3.238 (2)	129
C19*A*—H19*A*···Cl1*A*^{iv}	0.95	2.71	3.437 (3)	155
		V		
C7*A*—H7*A*···O1*B*	0.95	2.59	3.474 (3)	155
C8*A*—H8*A*···F1*B*	0.95	2.54	3.381 (3)	147
C18*A*—H18*A*···O2*A*^{vi}	0.95	2.56	3.283 (3)	133
C18—H18···O2^{vii}	095	2.60	3256 (2)	131
C26*A*—H26*A*···O1*B*^{viii}	0.95	2.39	3.335 (3)	171
C14*B*—H14*B*···O2	0.95	2.38	3.257 (3)	154
C18*B*—H18*B*···O2*B*8^{ix}	0.95	2.44	3.261 (3)	145
C24*A*—H24*A*···O1*A*^{vii}	0.95	2.62	3.346 (4)	133
C26—H26···O1^{x}	0.95	2.47	3.256 (4)	140
C27*B*—H27*B*···O1*A*^{viii}	0.95	2.49	3.357 (3)	152
C27—H27···F1^{x}	0.95	2.52	3.342 (3)	144

Symmetry codes: (i) *x*+1, *y*, *z*; (ii) *x*+1, *y*−1, *z*; (iii) *x*−1, *y*+1, *z*; (iv) *x*−1, *y*, *z*; (v) −*x*+1, −*y*+1, −*z*; (vi) *x*,-*y*+3/2,+*z*+1/2; (vii) *x*,-*y*+3/2,+*z*-1/2; (viii) -*x*+1,-*y*+1,-*z*+2; (ix) *x*,-*y*+1/2,+*z*+1/2; (x) -*x*,-*y*+2,-*z*+1.

C14B−H14B···O2A leading the packing to grow along to *c* axis (Fig. 9b).

In the crystal packing of compound **V,** two molecule of **Va** are arranged as dimers through C26−H26···O1, $R_2^2(22)$, and C27−H27···F1, $R_2^2(24)$, interactions. C14B−H14B···O2 interaction attaches those dimers. Along to *a* axis there is a molecular pair formed by the following interactions: C8A—H8A···F1B and C7A—H7A···O1B (Fig. 10). The interaction C18−H18···O2, and

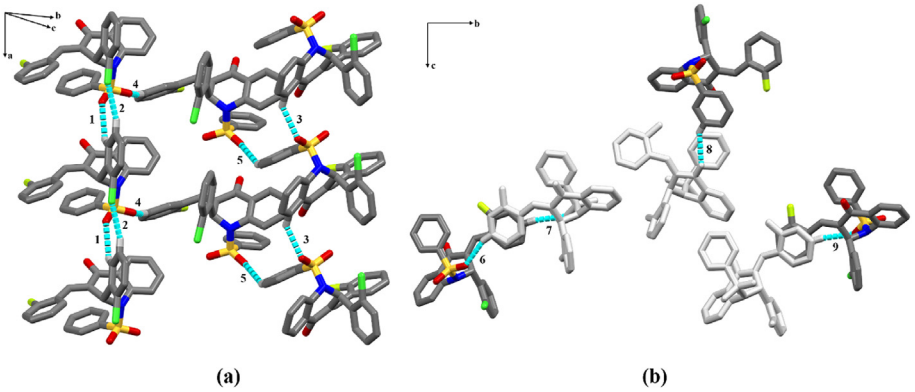


Fig. 9. A partial packing view of **IV** showing (a) the two chains formed by C18A—H18A ···O2A (1) and C19A —H19A···Cl1A (2), and C5—H5B···O3B (3) interactions. In addition, the interactions C15—H15···O3A (4) and C26B—H26B···O3 (5) responsible to connect molecules of **IVa** to the chains. In (b) are shown the discrete contacts C14A—H14A···O2B (6), C14B—H14B···O2A (7), C27A—H27A···O1 (8), C14B—H14B···O2A (9), responsible for grow the packing along the *c* axis. The motif represented in (a), is presented in light grey. For clarity, H atoms not involved in the motif have been omitted.

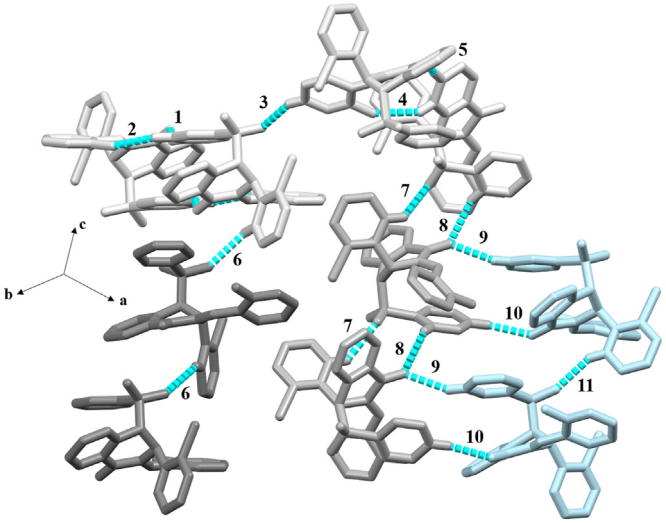


Fig. 10. A partial packing view of **V** showing the dimer and molecular pair in light grey, interactions C26—H26···O1 (1), C27—H27···F1 (2), C14B—H14B···O2 (3), C8A—H8A···F1B (4), nC7A—H7A···O1B (5). The chains formed by C18—H18···O2 (6) (grey); C18A—H18A···O2A (7) and C24A—H24A···O1A (8) (dark grey); and C18B—H18B···O2B (11) (light blue). The interactions C27B—H27B···O1A (9) and C26A—H26A···O1b (10) are discrete and connect two chains. For clarity, H atoms not involved in the motif have been omitted.

C18A—H18A···O2A create two independent $C_1^1(7)$ chains while C24A—H24A···O1A create a $C_1^1(9)$ chain, along to *c* axis. In the same orientation to these chains there is one more $C_1^1(7)$ arrangement formed by C18B—H18B···O2B interaction. This chain connects to others along to *a* axis, via C26A—H26A···O1B and C27B—H27B···O1A interactions (Fig. 10).

4. Final remarks

In this paper we briefly presented a comparison between chlorine and fluorine dihydroquinolinones. Due to the molecular similarities, we discuss the comparison of compounds **I** and **IV**, and compounds **III** and **V**. Concerning **I** and **IV,** the difference is in the halogen at the *ortho* position of ring B, chlorine, and fluorine, respectively. The decreasing order of ∠AB angle is **IVb**, **Ia**, **IVa**, **IVc** and **Ib**. The torsion ϕ_1 in **I** shows a smaller variation when compared to compound **IV**, *ca.* 0.3° and 12°, respectively. Regarding torsion ϕ_2, both compounds present similar values (23° for **I** and 19° for **IV)**. The torsion angle ϕ_3 in **IV** is larger than **I (**10° and 19°, respectively). Regarding molecular packing, the chlorine compound forms $C_1^1(7)$ and $C_1^1(6)$ chains, while the fluorine compound forms $C_1^1(7)$ chains. All those chains grow along the *a* axis. Finally, discrete contacts join these chains along the *b* and *c* axis.

In contrast, compounds **III** and **V** differ in the halogen in the *ortho* position of ring B, chlorine, and fluorine, respectively. In terms of ∠AB values, the decreasing order is **Va**, **Vc**, **IIIb**, **IIIa,** and **Vb**. The torsion ϕ_1 in **III** is smaller than **V** (1° and 13°, respectively). Regarding the torsion ϕ_2, both compounds present similar values (17° in **III** and 14° in **V)**. Moreover, the torsion ϕ_3 to **III** is larger than **V (**17° and 10°, respectively). These small alterations occur likely due to the change in radius of the halogen atoms. In terms of molecular packing both compounds have $R_2^2(12)$ dimers but in **III** it is related to SO_2 group while in the compound **V** it is related to the carbonyl group. In compound **III** these dimers are arranged along the *c* axis while in compound **V** they are along the *b* axis. Also, the molecular packing in **V** present two $C_1^1(7)$ and one $C_1^1(9)$ chain along the *c* axis. All those observations show that the change in position and kind of halogen atom attached in ring C play a significant role in the conformation of the studied compounds. On the other hand, the packing of these compounds is more susceptible to variations when the substituent position changes.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Wesley F. Vaz: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Validation, Visualization, Writing - original draft, Writing - review & editing. **Lidiane J. Michelini:** Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing - original draft, Writing - review & editing. **Gerlon A.R. Oliveira:** Data curation, Formal analysis, Investigation, Methodology, Resources, Validation, Visualization, Writing - original draft, Writing - review & editing. **Luciano M. Lião:** Data curation, Formal analysis, Investigation, Methodology, Resources, Validation, Visualization, Writing - original draft, Writing - review & editing. **Caridad N. Perez:** Data curation, Formal analysis, Investigation,

Methodology, Visualization, Resources, Writing - original draft, Writing - review & editing. **Allen G. Oliver:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Supervision, Validation, Visualization, Writing - original draft, Writing - review & editing. **Hamilton B. Napolitano:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Validation, Visualization, Writing - original draft, Writing - review & editing.

Acknowledgements

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES, Brazil); Grant ID: 88881.190472/2018-01 - Finance Code 001, Conselho Nacional de Desenvolvimento Científico e Tecnológico and Fundação de Amparo à Pesquisa do Estado de Goiás (FAPEG); Grant ID: 20170267000634.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.molstruc.2020.128559.

References

[1] S. Jain, V. Chandra, P. Kumar Jain, K. Pathak, D. Pathak, A. Vaidya, Comprehensive review on current developments of quinoline-based anticancer agents, Arab. J. Chem. 12 (2019) 4920–4946, https://doi.org/10.1016/j.arabjc.2016.10.009.

[2] A. Felczak, P. Bernat, S. Różalska, K. Lisowska, Quinoline biodegradation by filamentous fungus Cunninghamella elegans and adaptive modifications of the fungal membrane composition, Environ. Sci. Pollut. Res. 23 (2016) 8872–8880, https://doi.org/10.1007/s11356-016-6116-4.

[3] Y.-M. Zeng, H.-H. Zeng, F.-M. Liu, Synthesis and structure characterization of 1,5-benzothiazepine derivatives bearing quinoline moiety, J. Heterocycl. Chem. 53 (2016) 887–893, https://doi.org/10.1002/jhet.1519.

[4] C.A. Moreira, J.M.F. Custódio, W.F. Vaz, G.D.C. D'Oliveira, C. Noda Perez, H.B. Napolitano, A comprehensive study on crystal structure of a novel sulfonamide-dihydroquinolinone through experimental and theoretical approaches, J. Mol. Model. 25 (2019) 205, https://doi.org/10.1007/s00894-019-4091-7.

[5] L.J. Michelini, W.F. Vaz, G.D.C. D'Oliveira, C.N. Pérez, H.B. Napolitano, Analysis of two novel 1–4 quinolinone structures with bromine and nitrobenzyl ligands, J. Mol. Model. 25 (2019) 55, https://doi.org/10.1007/s00894-019-3937-3.

[6] L. Michelini, W. Vaz, L. Naves, C. Pérez, H. Napolitano, Synthesis, characterization and conformational analysis of two novel 4(1H)-Quinolinones, J. Braz. Chem. Soc. 31 (2020) 66–78, https://doi.org/10.21577/0103-5053.20190124.

[7] G.D.C. d'Oliveira, J.M.F. Custodio, A.F. Moura, H.B. Napolitano, C.N. Pérez, M.O. Moraes, L. Prókai, P. Perjési, Different reactivity to glutathione but similar tumor cell toxicity of chalcones and their quinolinone analogues, Med. Chem. Res. 28 (2019) 1448–1460, https://doi.org/10.1007/s00044-019-02384-8.

[8] W.F. Vaz, J.M.F. Custodio, G.D.C. D'Oliveira, B.J. Neves, P.S.C. Junior, J.T.M. Filho, C.H. Andrade, C.N. Perez, E.P. Silveira-Lacerda, H.B. Napolitano, Dihydroquinoline derivative as a potential anticancer agent: synthesis, crystal structure, and molecular modeling studies, Mol. Divers. (2020), https://doi.org/10.1007/s11030-019-10024-x.

[9] W.F. Vaz, G.D.C. D'Oliveira, C.N. Perez, B.J. Neves, H.B. Napolitano, Machine learning prediction of the potential pesticide applicability of three dihydroquinoline derivatives: syntheses, crystal structures and physical properties, J. Mol. Struct. 1206 (2020) 127732, https://doi.org/10.1016/j.molstruc.2020.127732.

[10] J. Jean, D.S. Farrell, A.M. Farrelly, S. Toomey, J.W. Barlow, Design, synthesis and evaluation of novel 2,2-dimethyl-2,3-dihydroquinolin-4(1H)-one based chalcones as cytotoxic agents, Heliyon 4 (2018), e00767, https://doi.org/10.1016/j.heliyon.2018.e00767.

[11] A. Pejović, J. Drabowicz, M. Cieslak, J. Kazmierczak-Baranska, K. Królewska-Golińska, Synthesis, characterization and anticancer activity of novel ferrocene containing quinolinones: 1-Allyl-2-ferrocenyl-2,3-dihydroquinolin-4(1H)-ones and 1-allyl-2-ferrocenylquinolin-4(1H)-ones, J. Organomet. Chem. 873 (2018) 78–85, https://doi.org/10.1016/j.jorganchem.2018.08.004.

[12] J. Koszuk, T. Bartosik, J. Wojciechowski, W.M. Wolf, A. Janecka, J. Drogosz, A. Długosz, U. Krajewska, M. Mirowski, T. Janecki, Synthesis of 3-Methylidene-1-tosyl-2,3-dihydroquinolin-4(1 H)-ones as potent cytotoxic agents, Chem. Biodivers. 15 (2018), e1800242, https://doi.org/10.1002/cbdv.201800242.

[13] L. Politanskaya, T. Rybalova, O. Zakharova, G. Nevinsky, E. Tretyakov, p-Toluenesulfonic acid mediated one-pot cascade synthesis and cytotoxicity evaluation of polyfluorinated 2-aryl-2,3-dihydroquinolin-4-ones and their derivatives, J. Fluor. Chem. 211 (2018) 129–140, https://doi.org/10.1016/j.jfluchem.2018.04.005.

[14] J.H. Yoon, J.Y. Lee, J. Lee, Y.S. Shin, S. Jeon, D.E. Kim, J.S. Min, J.H. Song, S. Kim, S. Kwon, Y. Jin, M.S. Jang, H.R. Kim, C.M. Park, Synthesis and biological evaluation of 3-acyl-2-phenylamino-1,4-dihydroquinolin-4(1H)-one derivatives as potential MERS-CoV inhibitors, Bioorg. Med. Chem. Lett 29 (2019) 126727, https://doi.org/10.1016/j.bmcl.2019.126727.

[15] R. Todeschini, V. Consonni, P. Gramatica, Chemometrics in QSAR, in: S. Brown, R. Tauler, R. Walczak (Eds.), Compr. Chemom. 4, Elsevier, Oxford, 2009, pp. 129–172.

[16] J.P. a Martins, M.M.C. Ferreira, Qsar modeling: a new open source computational package to generate and validate qsar models, Quim. Nova 36 (2013), https://doi.org/10.1590/S0100-40422013000400013. 554-U250.

[17] O.V. Dolomanov, L.J. Bourhis, R.J. Gildea, J.A.K. Howard, H. Puschmann, OLEX2: a complete structure solution, refinement and analysis program, J. Appl. Crystallogr. 42 (2009) 339–341, https://doi.org/10.1107/S0021889808042726.

[18] G.M. Sheldrick, A short history of *SHELX*, Acta Crystallogr. Sect. A Found. Crystallogr. 64 (2008) 112–122, https://doi.org/10.1107/S0108767307043930.

[19] G.M. Sheldrick, Crystal structure refinement with SHELXL, Acta Crystallogr. C 71 (2015) 3–8, https://doi.org/10.1107/S2053229614024218.

[20] M. Nardelli, Parst 95 – an update to PARST : a system of Fortran routines for calculating molecular structure parameters from the results of crystal structure analyses, J. Appl. Crystallogr. 28 (1995), https://doi.org/10.1107/S0021889895007138, 659–659.

[21] A.L. Spek, Single-crystal structure validation with the program PLATON, J. Appl. Crystallogr. 36 (2003) 7–13, https://doi.org/10.1107/S0021889802022112.

[22] G. d'Oliveira, A. Moura, M. de Moraes, C. Perez, L. Lião, Synthesis, characterization and evaluation of in vitro antitumor activities of novel chalcone-quinolinone hybrid compounds, J. Braz. Chem. Soc. 29 (2018) 2308–2325, https://doi.org/10.21577/0103-5053.20180108.

[23] J.H. Kim, H.W. Ryu, J.H. Shim, K.H. Park, S.G. Withers, Development of new and selective trypanosoma cruzi trans-sialidase inhibitors from sulfonamide chalcones and their derivatives, Chembiochem 10 (2009) 2475–2479, https://doi.org/10.1002/cbic.200900108.

[24] H. Ma, X. Zhou, D. Wei, J. Cao, C. Shi, Y. Fan, G. Huang, KHCO 3 - and DBU-promoted cascade reaction to synthesize 3-Benzyl-2-phenylquinolin-4(1 H)-ones, Chem. Asian J. 11 (2016) 2829–2833, https://doi.org/10.1002/asia.201600901.

[25] M.R.C. de Castro, R.F. Naves, A. Bernardes, C.C. da Silva, C.N. Perez, A.F. Moura, M.O. de Moraes, F.T. Martins, Tandem chalcone-sulfonamide hybridization, cyclization and further Claisen–Schmidt condensation: tuning molecular diversity through reaction time and order and catalyst, Arab. J. Chem. 13 (2020) 1345–1354, https://doi.org/10.1016/j.arabjc.2017.11.005.

[26] J. Grell, J. Bernstein, G. Tinhofer, Graph-set analysis of hydrogen-bond patterns: some mathematical concepts, Acta Crystallogr. Sect. B Struct. Sci. 55 (1999) 1030–1043, https://doi.org/10.1107/S0108768199007120.

Molecular Diversity
https://doi.org/10.1007/s11030-019-10024-x

ORIGINAL ARTICLE

Dihydroquinoline derivative as a potential anticancer agent: synthesis, crystal structure, and molecular modeling studies

W. F. Vaz[1,2] · J. M. F. Custodio[3] · G. D. C. D'Oliveira[3] · B. J. Neves[4] · P. S. C. Junior[5] · J. T. M. Filho[3] · C. H. Andrade[4] · C. N. Perez[3] · E. P. Silveira-Lacerda[6] · H. B. Napolitano[1]

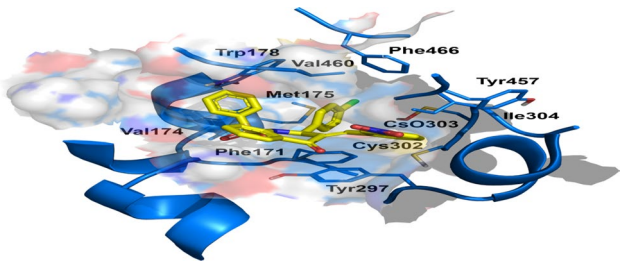
Received: 31 August 2019 / Accepted: 5 December 2019

Abstract

Cancer is one of the leading causes of death worldwide and requires intense and growing research investments from the public and private sectors. This is expected to lead to the development of new medicines. A determining factor in this process is the structural understanding of molecules with potential anticancer properties. Since the major compounds used in cancer therapies fail to encompass every spectrum of this disease, there is a clear need to research new molecules for this purpose. As it follows, we have studied the class of quinolinones that seem effective for such therapy. This paper describes the structural elucidation of a novel dihydroquinoline by single-crystal X-ray diffraction and spectroscopy characterization. Topology studies were carried through Hirshfeld surfaces analysis and molecular electrostatic potential map; electronic stability was evaluated from the calculated energy of frontier molecular orbitals. Additionally, in silico studies by molecular docking indicated that this dihydroquinoline could act as an anticancer agent due to their higher binding affinity with human aldehyde dehydrogenase 1A1 (ALDH 1A1). Tests in vitro were performed for VERO (normal human skin keratinocytes), B16F10 (mouse melanoma), and MDA-MB-231 (metastatic breast adenocarcinoma), and the results certified that compound as a potential anticancer agent.

Graphic abstract

A Dihydroquinoline derivative was tested against three cancer cell lines and the results attest that compound as potential anticancer agent.



Keywords X-ray diffraction · Molecular docking · ADMET properties · Hirshfeld surface · Anticancer activity

Introduction

Compounds derived from the fusion of two molecules acting independently can increase the pharmacological effectiveness, comparing to the sum of each individual moiety's potencies. These compounds can involve various functional parts, including small molecules or biomolecules, acting either simultaneously or sequentially [1]. In this paper, our

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11030-019-10024-x) contains supplementary material, which is available to authorized users.

✉ W. F. Vaz
wesfonseca@gmail.com

Extended author information available on the last page of the article

Published online: 03 January 2020

purpose was to study a novel hybrid compound composed by chalcone, sulfonamide, and quinolinone moieties. Chemically, chalcones provide a backbone with two aromatic rings linked by both carbonyl and olefin portions, being this molecular structure responsible by their biological potential [2–5]. In contrast, benzene sulfonamides (here called merely by sulfonamides) are compounds that contain the functional group -SO_2NH_2. The biological activities of these compounds are determined by changes in the aromatic ring [6–8]. In addition, the quinolines represent a leading class of alkaloids which contain a heteroatom in an aromatic ring, present in natural biologically active products [9].

It is well known that these compound classes play an important role in the field of pharmacological science. For example, a wide range of applications is assigned to chalcones, like inhibitors of glutathione S-transferase, potential as photo-luminescent material, besides being useful in the treatment of some cancer cells and hyperglycemia [9–14]. Not so far, sulfonamides have been available in many clinically used drugs, like diuretics (furosemide, indapamide, chlorthalidone, thiazides); carbonic anhydrase (CA) inhibitors (CAIs)—such as acetazolamide, dichlorphenamide, dorzolamide, and brinzolamide and antiepileptics (zonisamide and sulthiame) [12–14]. Quinolines also are very important heterocyclic compounds for they display interesting physiological and pharmacological activities. They are highlighted as cytotoxic agents as well as doxorubicin, against tumor and leukemic cells [14]. Regarding the latter, myeloid leukemia, the quinolinones have specific action inhibiting the tyrosine kinase 3 (FLT3) [15].

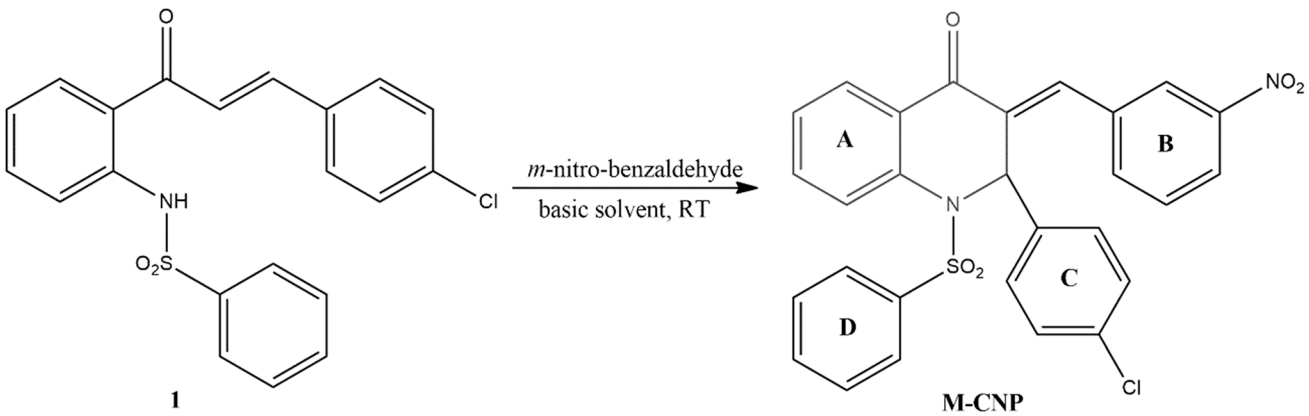
Although the chemistry of chalcones, sulfonamides, and quinolinones is well discussed in the literature, not so many data are found when these classes are fused to hybrid compounds. Recently, a series of compounds containing sulfonamide and chalcone derivatives were synthesized and tested in vitro against human liver cancer [16]. Based on this perspective, molecules that contain sulfonamide, quinoline, and chalcone moieties could lead to the new hybrid architectures with improved biological profiles. Therefore, this study aimed the synthesis of the hybrid compound (E)-2-(4-chlorophenyl)-3-(3-nitrobenzylidene)-1-(phenylsulfonyl)-2,3-dihydroquinolin-4(1H)-one (M-CNP) using (E)-*N*-(2-(3-(4-chlorophenyl)acryloyl)phenyl)benzenesulfonamide and *m*-nitro-benzaldehyde, in a basic medium as reagents. Later, its structural elucidation in terms of lengths and bond angles, intra- and intermolecular interactions, and supramolecular arrangements were studied. Finally, anticancer activities were assessed using an integrated ligand- and structure-based target. Fishing approach and in vitro tests were performed to validate this model.

Experimental and computational procedures

Synthesis and crystallization

M-CNP was obtained by reacting, respectively, *m*-nitro-benzaldehyde with precursor **1** (chalcone, which was previously synthesized based on the methodology reported by Castro et al. [17]). The first step of this reaction is the intramolecular addition of the nitrogen of sulfonamide group to β carbon of the precursor **1**, followed by the reaction of β carbonyl of the quinolinone formed by the aldehyde by Claisen–Schmidt condensation, yielding the desired compounds.

M-CNP was synthesized according to Scheme 1. Precursor **1** (1.0 mmol) and *m*-nitro-benzaldehyde (2.0 mmol) were dissolved in 15 mL of basic ethanol (56.1 mg of dissolved potassium hydroxide) and reacted (25 °C) for 48 h. The solution was filtered, and the precipitate was rinsed with 15 mL of ethanol. The precipitate was dissolved in dichloromethane (10 mL), and this solution was extracted with water. The organic phase was slowly evaporated, yielding the product. (E)-2-(4-chlorophenyl)-3-(3-nitrobenzylidene)-1-(phenylsulfonyl)-2,3-dihydroquinolin-4(1H)-one: pale



Scheme 1 Conditions for the synthesis of M-CNP

yellow crystalline solid, yield 90.4%, purity of 98.3%, mp 212–215 °C. ^{1}H NMR (500 MHz, $CDCl_3$) δ 6.49 (*s*, 1H), δ 7.09–7.11 (*m*, 2H), δ 7.24–7.30 (*m*, 4H), δ 7.32–7.37 (*m*, 3H), δ 7.36–7.39 (*m*, 1H), δ 7.55–7.61 (*m*, 3H), δ 7.62 (*s*, 1H), δ 7.70 (*dd*, 0.83 Hz, 8.18 Hz, 1H), δ 7.91 (*dd*, *J* 1.55 Hz, 7.80 Hz, 1H), δ 8.02 (*t*, *J* 1.93 Hz, 1H), δ 8.30 (*dd*, *J* 2.18 Hz, 8.28 Hz, 1H); ^{13}C NMR (126 MHz, $CDCl_3$) δ 59.4, 124.6, 125.1, 127.2, 127.7, 127.9, 128.2, 128.4, 128.9, 129.1, 129.4, 130.4, 132.8, 133.7, 134.3, 134.9, 135.1, 135.2, 135.4, 137.4, 138.8, 148.5, 182.0; IR 1674 (m), 1613 (m), 1478 (w), 1353 (s), 1305 (m), 1237 (m); HRMS calculated for $C_{28}H_{19}ClN_2O_5S$ 531.0781, found 531.0540.

Crystallographic characterization

Single-crystal X-ray data were obtained using an Agilent SuperNova diffractometer with MoKα radiation ($\lambda = 0.71073$ Å) where crystals of M-CNP were kept at room temperature (298 K). Data collection, reduction, and cell refinement were made with the CRYSALISPRO software [18]. All the final unit cell parameters were considered on the fitting of all reflection positions. Using Olex2 [19], with ShelXS [20], the structure was solved using direct methods and refined by full-matrix least-squares refinement on F^2 using SHELXL [21]. All non-hydrogen atoms were refined with anisotropic parameters. All the hydrogen atoms were placed in calculated positions and refined with fixed individual displacement parameters [$U_{iso}(H) = 1.2$ Ueq or 1.5 Ueq], according to the riding model (C–H bond lengths of 0.97 Å for aromatic and CH_3 groups). H atoms in other positions were freely refined. The structures were drawn by ORTEP-3 for windows [22] and Mercury [23] programs. The possible interactions and hydrogen bonds were checked by PLATON [24] software and analyzed using the Hirshfeld surface from Crystal Explorer software [25]. The crystallographic data were deposited at the Cambridge Crystallographic Data Centre (CCDC) [26, 27] under the Number 1838222. Copies of the data can be obtained, free of charge, via www.ccdc.cam.ac.uk.

Hirshfeld surface analysis

The M-CNP molecular packing was studied throughout the Hirshfeld surface (HS). Once it provides a three-dimensional picture of close contacts in a crystal, this methodology is a great tool to analyze intermolecular interaction, which can be summarized in a fingerprint plot. Both donor and acceptor regions are considered by the d_{norm} surface, which is based on the normalized function of d_e (distance from the point to the nearest nucleus external to the surface) and d_i (distance to the nearest nucleus internal to the surface) contacts. The d_{norm} surface is given by

$$d_{norm} = \frac{\left(d_i - r_i^{vdW}\right)}{r_i^{vdW}} + \frac{\left(d_e - r_e^{vdW}\right)}{r_e^{vdW}} \quad (1)$$

where r_i^{vdW} and r_e^{vdW} are the van der Waals radii [25]. The HS could be used to recognize hydrophobic interactions such as $\pi\cdots\pi$ and C–H$\cdots\pi$ from the curvature function named shape index surface [28]. This function depends on the principal curvatures k_1 and k_2 following the equation

$$S = \frac{2}{\pi}\left(\frac{k_1 + k_2}{k_1 - k_2}\right) \quad (2)$$

The donor and acceptor regions of these hydrophobic interactions were recognized as complementary bumps (blue) and hollows (red), respectively, where two molecular surfaces touch one another. Moreover, the distances involved in the intermolecular interactions were quantitatively summarized on a 2D plot named *fingerprint* [29], which explores the nature and kind of intermolecular interaction experienced by the title compound. The Crystal Explorer program was utilized to perform this HS analysis, with the fingerprint plots using the standard 0.6–2.8 Å view of d_e versus d_i.

Molecular target prediction

The biological activities for the title compound were predicted by MolTarPred server [30]. The query structure (M-CNP) was drawn in 2D using MarvinSketch software (ChemAxon, Budapest, Hungary, http://www.chemaxon.com) and submitted to the MolTarPred webserver. This server predicts biological activities based on Tanimoto similarities between the Morgan fingerprints of the query compound and 607,659 ChEMBL compounds with experimental biological data. The predicted targets were those retrieved from the top 10 most similar compounds to the query. The confidence in a predicted target was estimated as the number of top hits known to possess activity against the predicted target [30].

Ligand and protein preparation

The 3D structure of the M-CNP, co-crystallized ligands, and known inhibitors and non-inhibitors were imported to Maestro workspace v.9.3 (Schrödinger, LCC, New York, 2012) and prepared using LigPrep v.2.5 at a neutral pH (7.4 ± 1.0). Subsequently, up to 2000 conformers were generated for the two enantiomers (*R* and *S*) using OMEGA v.2.5. [31, 32], and the AM1-BCC charges [33] were added using QUACPAC v.1.6.3 [34]. The 3D structures of the predicted protein targets were preprocessed using Protein Preparation Wizard available on Maestro workspace v.9.3 (Schrödinger, LCC, New York, 2012). Next, hydrogen atoms were included in the proteins, while bond orders and formal charges were adjusted.

The protonation states of polar amino acids were predicted by PROPKA v.3.1 [35] at neutral pH (7.4 ± 1.0). Subsequently, the full-atom protein structure minimization using the OPLS-2005 force field was carried out [36].

Molecular docking

The prepared proteins were then subjected to the grid-generation protocol using two binding site detection strategies. In the first strategy, grids were generated with a molecular probe available on OEDocking suite v.3.2.0 [37–39] for the detection of binding pockets around the protein. In the second strategy, co-crystallized ligands were considered geometric centers of the grids. Details of protein crystals and grids details are available on Supplementary Table S2. Finally, molecular docking calculations were performed implementing the high-resolution protocol of FRED program with the ChemGauss4 score function [37, 38, 40], both available on the OEDocking suite [37–39].

ADMET properties calculation and PAINS analysis

Some pharmacokinetics and toxicity (ADMET) properties were calculated for M-CNP using two different web services. The potential for hERG blockage was assessed using the PredhERG server [41, 42]. In parallel, the admetSAR server [43] was used to assess blood–brain barrier penetration [44], human intestinal absorption [44], inhibition and substrate metabolism prediction of five most important cytochrome P450 isoforms [45], AMES mutagenicity [46], and carcinogenicity [47]. Studies on pan-assay interference compounds (PAINS) [48, 49] were performed using False Positive Remover server [48].

Cytotoxicity assay

To evaluate the in vitro cell viability after treatment for 48 h with the M-CNP and to determine its IC50 concentrations, the MTT colorimetric method was used. These screenings were made for VERO (ATCC® CCL-81™—African green monkey kidney cell), B16F10 (ATCC® CRL-6475—murine melanoma cell), and MDA-MB-231 (ATCC® HTB-26™—human triple-negative metastatic breast adenocarcinoma). The selectivity indexes (SI) were determined using the formula: $SI = \frac{IC_{50\,\text{normal cell}}}{IC_{50\,\text{tumor cell}}}$. The values of $SI \geq 2.0$ were considered significant.

Results and discussion

X-ray crystal structure

The M-CNP possesses an *m*-nitro group in ring B and a *p*-chlorine atom in ring D. Their crystals belong to centrosymmetric monoclinic space group $P2_1/n$. The crystallographic parameters are presented in Table 1, followed by the Ortep representation in Fig. 1.

The main core of the M-CNP molecule is a dihydroquinolin-4(1H)-one motif (Scheme 1), where there is a sulfonylbenzene group attached to an N atom, a nitro-3-vinylbenzene attached to C8 atom, and a chlorobenzene group attached to C16 atom. The piperidinone cycle assumes chair conformation; thus, O1 and N1 atoms are not coplanar with the ring A. As M-CNP (a racemate) crystallizes in the $P2_1/n$ space group, the following discussions were made using a molecule with *S* configuration regarding chiral center C16. Additionally, this molecule has an *E* conformation at the double bond C7 = C8. Also, another aromatic substituent in dihydroquinolin-4(1H)-one is offset, related to ring A in 64.57° for ring B, 48.40° for ring C and 82.26° for ring D. The geometric parameters are presented in Table S1.

Table 1 Crystal data and structure refinement of M-CNP dihydroquinoline

Empirical formula	$C_{28}H_{19}ClN_2O_5S$
Formula weight	530.96
Temperature/K	293(2)
Crystal system	Monoclinic
Space group	$P2_1/n$
a/Å	11.5271 (6)
b/Å	15.4704 (6)
c/Å	13.7852 (6)
α/°	90
β/°	100.320 (5)
γ/°	90
Volume/Å^3	2418.53 (19)
Z	4
ρ_{calc}g/cm^3	1.458
μ/mm^{-1}	0.289
F(000)	1096.0
Crystal size/mm^3	0.45 × 0.40 × 0.30
Radiation	MoKα (λ = 0.71073)
Reflections collected	56262
Independent reflections	5510
Data/restraints/parameters	5510/0/334
Goodness-of-fit on F^2	1.044
Final R_1 index [$I \geq 2\sigma(I)$]	R_1 = 0.0380
Largest diff. peak/hole/e Å^{-3}	0.27/− 0.35

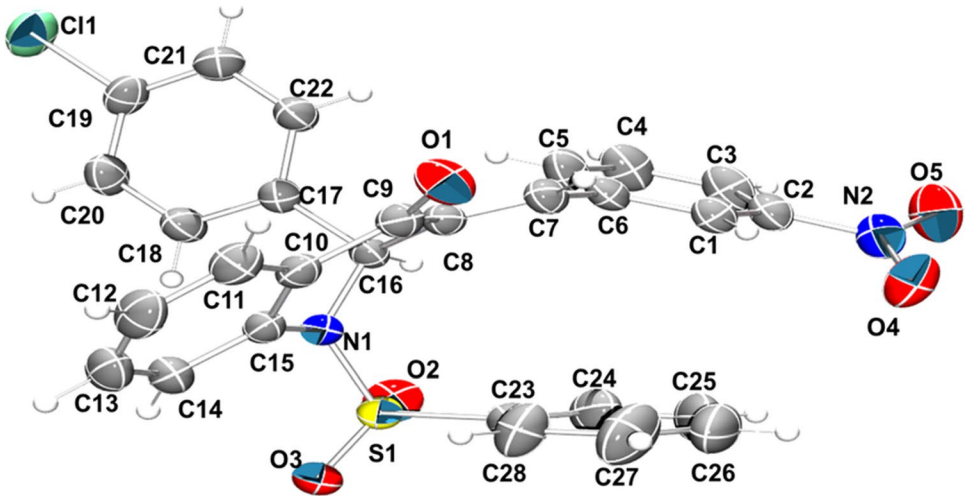


Fig. 1 Thermal ellipsoid drawing of M-CNP showing atom labeling and displacement ellipsoids at the 50% probability level. H atoms are shown as small spheres of arbitrary radii

The M-CNP molecule features the absence of strong H-bond donor; thus, their packing is stabilized by non-conventional H-bonds. There are four dimers in M-CNP structure: three with only C–H···O interactions (C5–H5···O2, C21–H21···O1 and C26–H26···O4), and one resulting of the combination between C–H···O (C12–H12···O4) and C–H···Cl (C28–H28···Cl1), resulting in an extremely stable packaging. In the first one, it is observed that the interaction between the nitrobenzene and sulfonyl groups generates a dimer with $R_2^2(18)$ motif (Fig. 2a). Furthermore, two intermolecular interactions generate a second dimer with $R_2^2(25)$ motif (Fig. 2b). One of these interactions is arranged between nitrobenzene group and ring A, while the other involves chlorobenzene group and ring C. The third dimer with $R_2^2(16)$ motif results from the interaction of carbonyl and chlorobenzene groups (Fig. 2c). Finally, the last dimer is stabilized by two interactions between nitrobenzene group and the ring C resulting in a $R_2^2(30)$ motif (Fig. 2d). The geometric parameters of intermolecular interactions are presented in Table 2.

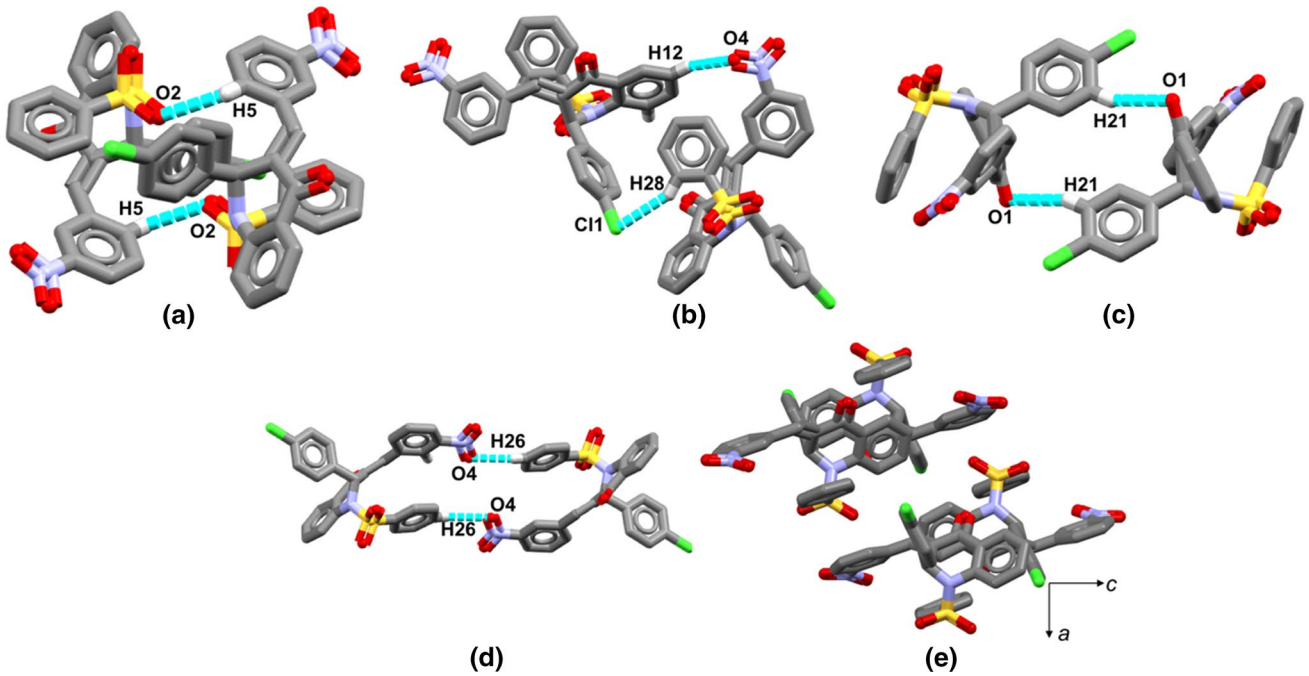


Fig. 2 Dimers observed in the crystal packing of M-CNP with motifs $R_2^2(18)$ (**a**), $R_2^2(25)$ (**b**), $R_2^2(16)$ (**c**) and $R_2^2(30)$ (**d**). The molecular packing of M-CNP is shown along the *b* axis (**e**)

Table 2 Non-classical hydrogen bonds of M-CNP

D–H···A	$d_{D\text{-}H}$ (Å)	$d_{H\cdots A}$ (Å)	$d_{D\cdots A}$ (Å)	$d_{H\cdots A}$ (°)	Symmetry code
C5–H5 ···O2	0.93	2.58	3.18	123	$1-x, 1-y, -z$
C12–H12 ···O4	0.93	2.43	3.33	162	$1/2-x, 1/2+y, 1/2-z$
C21–H21 ···O1	0.93	2.61	3.48	157	$-x, 1-y, -z$
C26–H26 ···O4	0.93	2.63	3.38	139	$1-x, -y, -z$
C28–H28 ···Cl1	0.93	2.88	3.72	151	$1/2-x, 1/2+y, 1/2-z$

In order to understand the supramolecular arrangement, the HS analysis was performed to M-CNP. The d_{norm} surface is illustrated in Fig. 3, where red dots represent the close interactions (I–V). The interaction C12–H12···O4 has a shorter character than others as shown in Fig. 3. The contacts (II) and (IV) are related to a dimeric interaction constituted by nitrobenzene group and ring A and chlorobenzene group and ring C (Fig. 2b).

The conformation of the structure is due to the stabilizing effect of intramolecular interactions like π···π, for example. The aromatic groups in M-CNP lead to two C–H···π and one π···π interaction (Fig. 4) in the molecular packing. These hydrophobic interactions connect the dimers and generate molecular layers in the crystal packing. The shape index map of HS helps to analyze molecular contacts that indicate C–H···π and π···π interactions. C–H···π interactions are recognized by a broad depression in the surface above the benzene ring. Red is used to represent the acceptor region, while blue represents the donor region of intermolecular contacts. Furthermore, π···π interactions are characterized by two triangular shapes like a "bowtie" above the aromatic rings indicating regions where there is overlap in the delocalized π system, often in centers of gravity (Cg) of the benzene rings [50].

The first C–H···π interactions involve C27 and H27 atoms, and chlorobenzene ring [H···π = 3.005 Å] (I) (Fig. 4a), while the second one involves C24 and H24 atoms, and chlorobenzene ring [H···π = 3.234 Å] (II) (Fig. 4b). The region marked with a circle represents the π···π interactions between rings B and C [3.790 Å, symmetry code = $1/2+x, 1/2-y, 1/2+z$] (Fig. 4c).

The combination of d_i and d_e on a 2D plot (fingerprint) provides more information about all contacts in the molecule. Moreover, it is highly sensitive to the immediate environment of the molecule and unique for a given molecule [29, 50]. In the fingerprints, the C – H···π interactions are represented by C···H contacts, while the π···π interactions are recognized by C···C contacts. Figure 5 presents the fingerprint of the M-CNP molecule.

The H···H contacts constitute a high percentage in the fingerprints (30.7%), as expected of an organic molecule. Hydrophobic interactions (C···C and C···H contacts) are also observed in M-CNP fingerprints. A narrow light blue spot at $d_i = d_e \approx 1.8$ Å is due to a single C···C contact typical from π···π interactions, representing 4.1% in the total fingerprint. Meanwhile, a couple of asymmetric lateral wings at minimum $d_i + d_e \approx 2.8$ Å is attributed to C–H···π interactions and it contributes with 18.6% in the fingerprint of M-CNP. Small peaks along 2.6 Å $< d_i + d_e <$ 3.6 Å (I) are

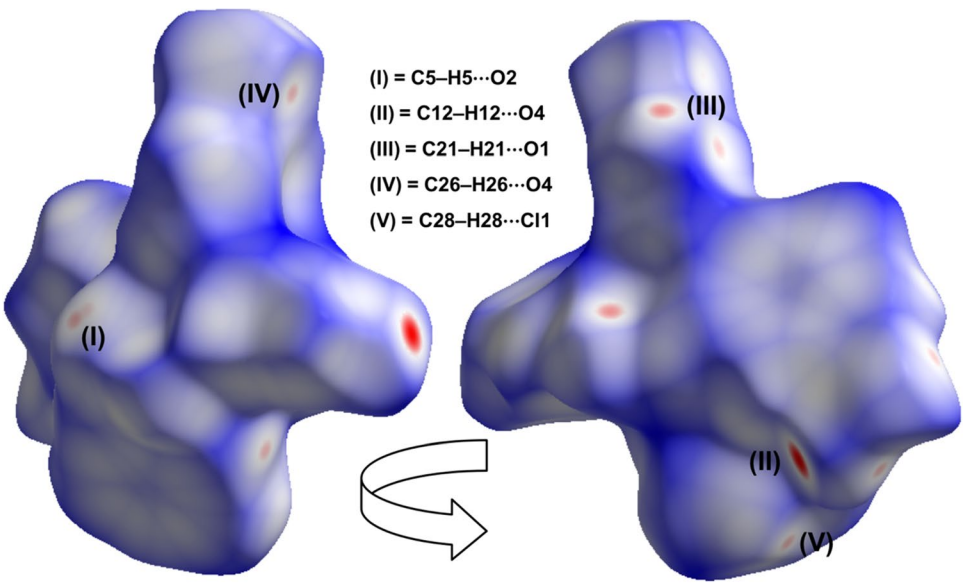


Fig. 3 The Hirshfeld surface d_{norm} map showing C–H···O and C–H···Cl interactions observed in the molecular packing of M-CNP. Blue indicates low intensity, and red indicates high-intensity contacts. The arrow is used to indicate that this is the backside of the molecule. (Color figure online)

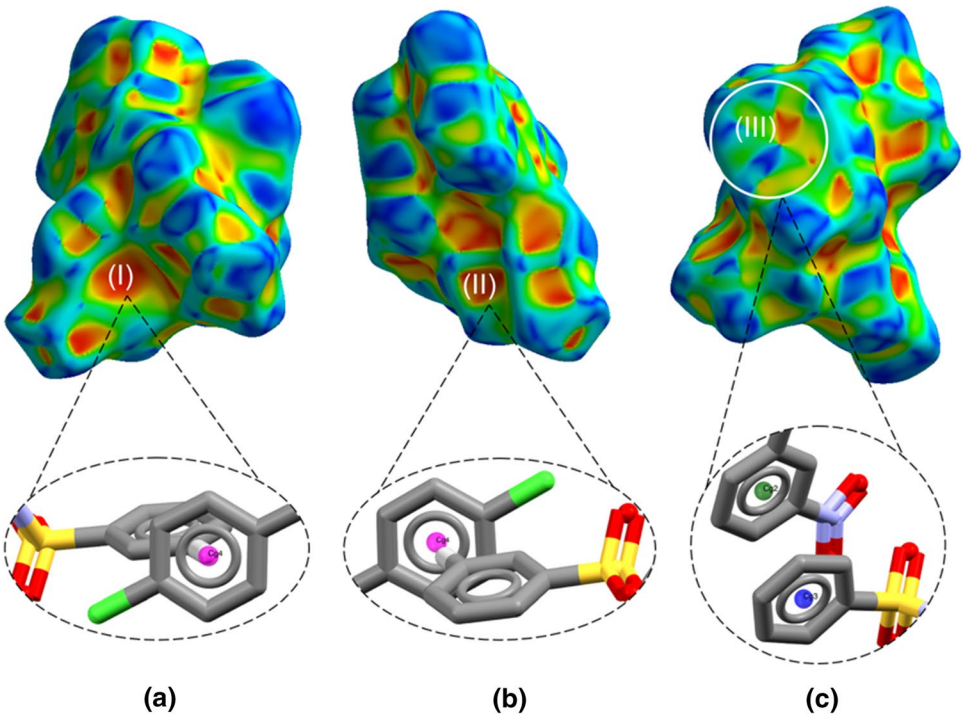


Fig. 4 The Hirshfeld surface shape index shows the C–H···π (**a**, **b**) and π···π (**c**) interactions establishing the crystal packing of M-CNP

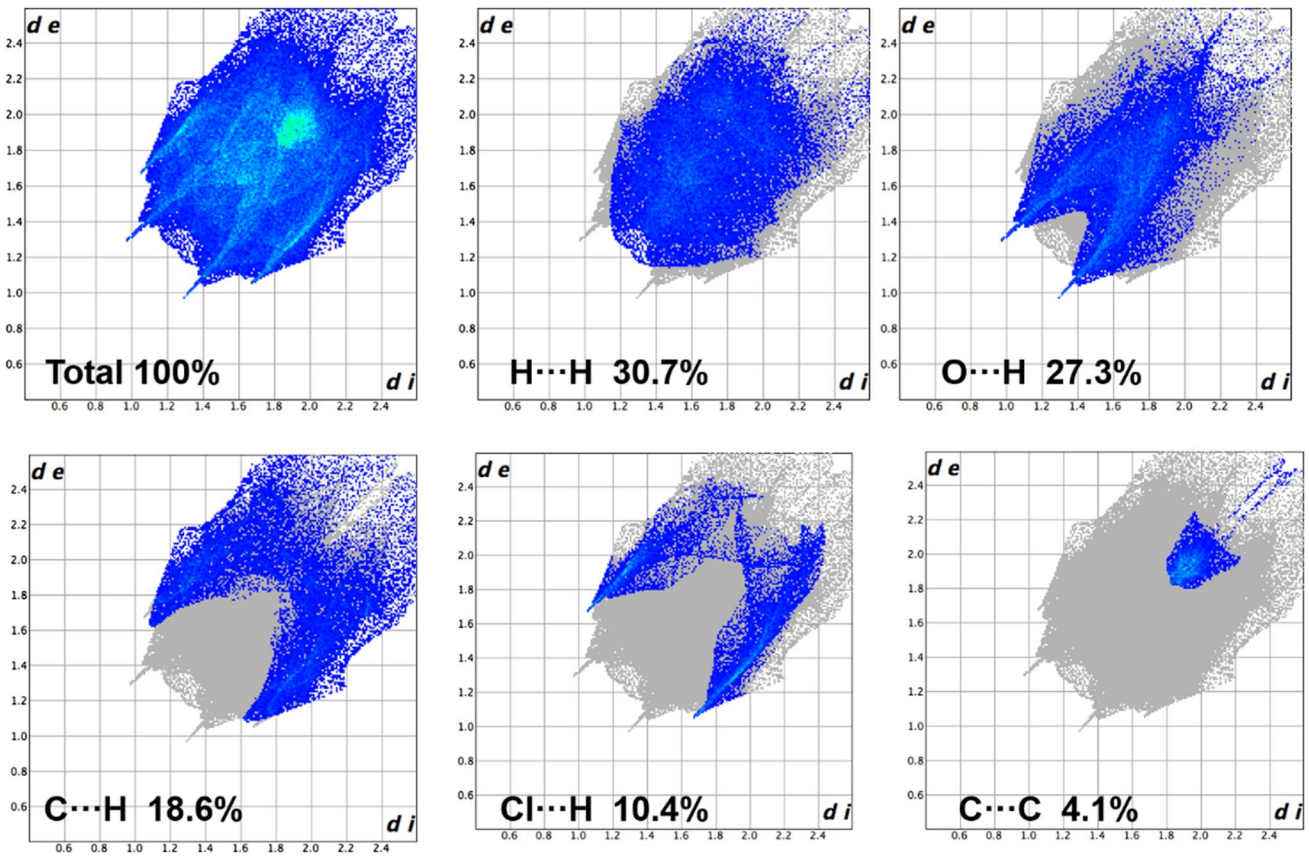


Fig. 5 Hirshfeld surface fingerprint plots of the nearest internal distance (d_i) *versus* the nearest external distance (d_e) for M-CNP. The colors represent the number of points that share the same d_i, d_e coordinate (light blue: many; dark blue: few). (Color figure online)

associated to C–H···O interactions corresponding to 27.3% of HS. Similarly, the spikes due to the C–H···Cl contact are pushed farther apart on the plot with the minimum d_i at 1.80 corresponding to 10.4% of all HS.

In silico identification of potential biological activities

The in silico studies have been integrated into all modern drug-discovery programs. Here, we performed a computational investigation aiming to identify potential biological activities for M-CNP using the MolTarPred server. The fundamental idea underlying this server is that compounds sharing enough similarity have enhanced the probability to share the same biological profile [51, 52]. The results of our preliminary screening indicated that M-CNP may interact with nuclear receptor ROR-gamma, survival motor neuron protein, Niemann-pick C1 protein, Ras-related protein Rab-9A, and aldehyde dehydrogenase 1A1 (see Table 3).

To improve the specificity in the predictions, our preliminary list of targets was used for further investigation in molecular docking studies. Initially, the FRED protocol was evaluated by re-docking, which removes the crystallographic ligand with subsequent docking. This re-docking showed the structural differences in the crystallographic ligand and the docked ligand. The root mean square deviation (RMSD) of the heavy atoms in the crystallographic ligand and re-docked ligand should be less than 2.0 Å [53]. In general, FRED could reproduce the crystallographic conformation for all molecular targets with RMSD values less than 2.0 Å (see Table S2 and Figure S1). Subsequently, M-CNP was docked against all predicted targets. As we can see in Table 3, M-CNP could act as an anticancer agent because of its Chemgauss4 of − 16.56 with human aldehyde dehydrogenase 1A1 (ALDH 1A1). The binding affinity of M-CNP is higher than that obtained for the co-crystallized ligand 3-benzyl-4-methyl-2-oxo-2H-chromen-7-yl methanesulfonate (Chemgauss4 of − 14.17, experimental IC_{50} of 2.8 µM [54]). In addition, we performed an additional docking validation study with ALDH 1A1 inhibitors and non-inhibitors (Table S3) deposited in the ChEMBL database [55] (Assay ID: CHEMBL1614458). All inhibitors (experimental IC_{50} ~ 0.02 µM) had scores below − 13.10, while non-inhibitors (experimental IC_{50} = 39.8 µM) had scores above − 11.85, demonstrating the reliability of docking prediction.

The ALDH 1A1 is a nicotinamide-adenine dinucleotide phosphate–positive ($NAD(P)^+$)-dependent enzyme that catalyzes the oxidation of retinal (retinaldehyde) to retinoic acid (RA) [56]. By serving as a ligand for nuclear RA receptors (RARs) and retinoid X receptors (RXRs), RA induces differentiation of cells and regulates gene transcription, signaling events, and post-translational modification of proteins. Therefore, the inhibition of ALDH 1A1 might play a key role in the regulation of growth and differentiation of cancer tissue types, such as breast, lung, esophagus, colon, and stomach epithelium [56].

Figure 6 illustrates the predicted binding mode of M-CNP into the active site of ALDH 1A1. The 4-chlorophenyl moiety of the ligand forms a π-stacking interaction with Phe171, while a positively charged nitrogen of the nitro group interacts with Tyr297 through a π-cation interaction. Apparently, the 4-chlorophenyl moiety also promotes a steric hindrance in the oxidized Cys303 (CsO303) pocket. In addition, the predicted binding mode allows several hydrophobic interactions between the phenyl rings of the ligand with residues Met175, Trp178, Cys302, Ile304, Tyr457, Phe466, and Val470.

Analysis of ADMET properties and PAINS

Poor ADMET properties are considered the main reasons for the failures of drug candidates in the preclinical and clinical phases [57]. Therefore, a preliminary assessment of ADMET properties can minimize the time and cost of research by identifying the strongest candidates for clinical development and rejecting those with a low probability of success [58]. M-CNP was subjected to in silico assessment of ADMET properties using Pred-hERG [41, 42] and admetSAR server [43].

Our analyses predicted M-CNP as a non-blocker of hERG, non-mutagenic, non-carcinogenic, and with acceptable human intestinal absorption and low permeability in the blood–brain barrier. In addition, M-CNP was expected

Table 3 Top five predicted targets for M-CNP

Predicted target	Organism	Confidence of prediction	Docking scores	
			M-CNP	Co-crystalized ligand
Nuclear receptor ROR-gamma	*Mus musculus*	3	− 11.02	− 17.43
Survival motor neuron protein	*Homo sapiens*	2	1.00	− 7.91
Niemann-pick C1 protein	*Homo sapiens*	2	1.00	–
Ras-related protein Rab-9A	*Homo sapiens*	2	− 4.00	− 13.23
Aldehyde dehydrogenase 1A1	*Homo sapiens*	2	− 16.56	− 14.17

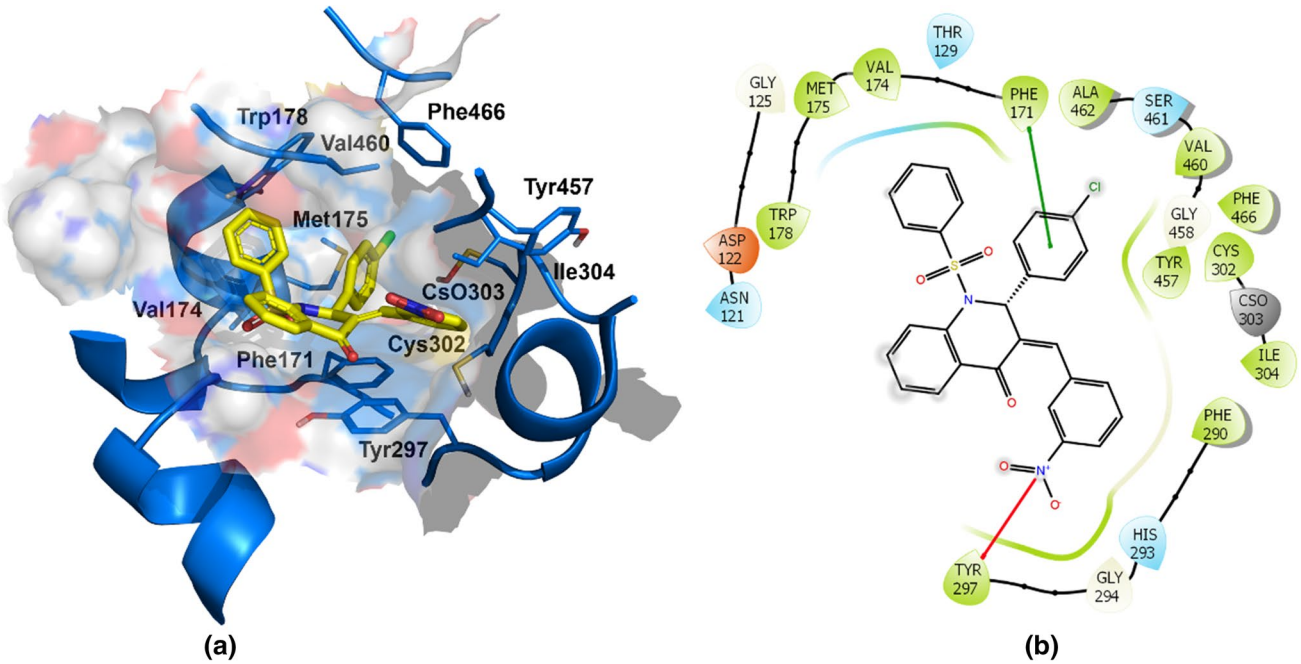


Fig. 6 **a** Binding mode of M-CNP with human aldehyde dehydrogenase 1A1, and **b** 2D ligand interactions with key residues in the active site. The π-stacking and π-cation interactions are presented as green and red lines, respectively. (Color figure online)

Table 4 Predicted ADMET properties and PAINS analysis of M-CNP

Property	M-CNP
hERG[a]	Non-blocker
Carcinogenicity[b]	Non-carcinogenic
Blood–brain barrier penetration[b]	Low permeability
AMES toxicity[b]	Non-mutagenic
CYP450 2C9[b]	Inhibitor
CYP450 2D6[b]	Inhibitor
CYP450 1A2[b]	Inhibitor
CYP450 2C19[b]	Inhibitor
CYP450 34A[b]	Inhibitor
Human intestinal absortion[b]	Moderate (> 30%)
PAINS	Not

[a]Pred-hERG [41, 42]

[b]admetSAR [44]

to have high inhibitory promiscuity against five major cytochrome P450 isoforms (1A2, 2C9, 2C19, 2D6, and 3A4). Subsequently, M-CNP was subjected to a computational sub-structure filter aiming to identify PAINS. This analysis suggests that M-CNP provides a reduced probability to nonspecifically interact or aggregate with proteins in prospective biological assays. These results are summarized in Table 4.

Cytotoxicity assay

The MTT results of M-CNP indicated a marked decrease in cell viability, with a high concentration. The values of IC_{50} for VERO, MDA-MB-231, and B16F10 were, respectively, 171.62, 54.59, and 76.08 µM, while the SI values for MDA-MB-231 and B16F10 were 3.14 and 2.25, respectively. In general, M-CNP was more active in cancer cell lines MDA-MB-231 and B16F10 than in normal cells (VERO), where the IC_{50} value was 171.62 µM. That indicates selectivity for cancer cells, pertinent in the search for more specific drugs. The selectivity indexes indicate the selectivity of the complexes against tumor cells and provide information about their potential for clinical and preclinical in vivo tests.

Conclusions

M-CNP was crystallized in centrosymmetric monoclinic space group $P2_1/n$. In addition, the dimeric supramolecular arrangement was formed by intermolecular interactions of type C–H···O and C–H···Cl, as confirmed by Hirshfeld surface analysis and molecular potential maps for M-CNP. Thereafter, molecular docking studies indicated that the compound could act as an anticancer agent because of its affinity with human aldehyde dehydrogenase (ALDH 1A1). In this model, the active sites 4-chlorophenyl moiety, nitrogen (NO_2 group), and several hydrophobic interactions between the phenyl rings of the ligand with

residues in ALDH 1A1 have a key role in the potential anticancer activity.

Acknowledgements The authors would like to acknowledge the Brazilian agencies Fundação de Amparo à Pesquisa do Estado de Goiás (FAPEG), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for financial support. We would like to thank the Universidade Federal de Juiz de Fora for the X-ray diffraction data collection, OpenEye Scientific Software Inc. (https://www.eyesopen.com/), and ChemAxon (https://chemaxon.com) for providing us with an academic license of their software. CHA and HBN are the researcher's fellows of CNPq.

Author Contributions WFV and HBN designed the project. JMFC, GDCO, and CNP performed the synthesis, crystallization, and spectroscopies studies of M-CNP. WFV, HBN, and PSCJ performed crystallography studies. JTMF, BJN, and CHA performed the in silico studies. EPSL performed the in vitro studies. All authors have written, critically reviewed, and approved the manuscript.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Müller-Schiffmann A, Sticht H, Korth C (2012) Hybrid compounds. BioDrugs 26:21–31. https://doi.org/10.2165/11597630-000000000-00000
2. Lahtchev KL, Batovska DI, Parushev SP et al (2008) Antifungal activity of chalcones: a mechanistic study using various yeast strains. Eur J Med Chem 43:2220–2228. https://doi.org/10.1016/j.ejmech.2007.12.027
3. Custodio J, Faria E, Sallum L et al (2017) The influence of methoxy and ethoxy groups on supramolecular arrangement of two methoxy-chalcones. J Braz Chem Soc 28:2180–2191. https://doi.org/10.21577/0103-5053.20170067
4. Carvalho PS, Custodio JMF, Vaz WF et al (2017) Conformation analysis of a novel fluorinated chalcone. J Mol Model 23:97. https://doi.org/10.1007/s00894-017-3245-8
5. Silva WA, Andrade CKZ, Napolitano HB et al (2013) Biological and structure-activity evaluation of chalcone derivatives against bacteria and fungi. J Braz Chem Soc 24:133–144. https://doi.org/10.1590/S0103-50532013000100018
6. Ávila HP, de Smânia FAE, Monache FD, Smânia A (2008) Structure–activity relationship of antibacterial chalcones. Bioorg Med Chem 16:9790–9794. https://doi.org/10.1016/j.bmc.2008.09.064
7. Nowakowska Z (2007) A review of anti-infective and anti-inflammatory chalcones. Eur J Med Chem 42:125–137. https://doi.org/10.1016/j.ejmech.2006.09.019
8. Rosa GP, Seca AML, do Barreto MC et al (2019) Chalcones and flavanones bearing hydroxyl and/or methoxyl groups: synthesis and biological assessments. Appl Sci 9:2846. https://doi.org/10.3390/app9142846
9. Al-Karawi AJM, Hammood AJ, Awad AA et al (2018) Synthesis and mesomorphism behaviour of chalcones and pyrazoles type compounds as photo-luminescent materials. Liq Cryst 45:1603–1619. https://doi.org/10.1080/02678292.2018.1446553
10. Özaslan MS, Demir Y, Aslan HE et al (2018) Evaluation of chalcones as inhibitors of glutathione S-transferase. J Biochem Mol Toxicol 32:e22047. https://doi.org/10.1002/jbt.22047
11. Padhye S, Ahmad A, Oswal N et al (2010) Bioorganic & medicinal chemistry letters fluorinated 2 0-hydroxychalcones as garcinol analogs with enhanced antioxidant and anticancer activities. Bioorg Med Chem Lett 20:5818–5821. https://doi.org/10.1016/j.bmcl.2010.07.128
12. Scozzafava A, Owa T, Mastrolorenzo A, Supuran C (2003) Anticancer and antiviral sulfonamides. Curr Med Chem 10:925–953. https://doi.org/10.2174/0929867033457647
13. Supuran CT (2008) Carbonic anhydrases: novel therapeutic applications for inhibitors and activators. Nat Rev Drug Discov 7:168–181. https://doi.org/10.1038/nrd2467
14. Abdelli A, Gaucher A, Efrit ML et al (2015) Arylation of allylphosphonates and application to the preparation of phosphonomethyl-coumarin, -quinolinone and -benzoxepinone skeletons. Tetrahedron Lett 56:1679–1681. https://doi.org/10.1016/j.tetlet.2015.02.038
15. Chung HJ, Kamli MR, Lee HJ et al (2014) Discovery of quinolinone derivatives as potent FLT3 inhibitors. Biochem Biophys Res Commun 445:561–565. https://doi.org/10.1016/j.bbrc.2014.02.029
16. Ghorab MM, Ragab FA, Heiba HI et al (2015) Synthesis, anticancer and radiosensitizing evaluation of some novel sulfonamide derivatives. Eur J Med Chem 92:682–692. https://doi.org/10.1016/j.ejmech.2015.01.036
17. De Castro MRC, Aragão ÂQ, da Silva CC et al (2015) Conformational variability in sulfonamide chalcone hybrids: crystal structure and cytotoxicity. J Braz Chem Soc 27:884–898. https://doi.org/10.5935/0103-5053.20150341
18. Snejko N, Cascales C, Gomez-Lor B et al (2002) From rational octahedron design to reticulation serendipity. A thermally stable rare earth polymeric disulfonate family with CdI2-like structure, bifunctional catalysis and optical properties. Chem Commun. https://doi.org/10.1039/b202639b
19. Dolomanov OV, Bourhis LJ, Gildea RJ et al (2009) OLEX2: a complete structure solution, refinement and analysis program. J Appl Crystallogr 42:339–341. https://doi.org/10.1107/S0021889808042726
20. Sheldrick GM (2015) SHELXT: integrated space-group and crystal-structure determination. Acta Crystallogr A 71:3–8. https://doi.org/10.1107/S2053273314026370
21. Sheldrick GM (2015) Crystal structure refinement with SHELXL. Acta Crystallogr Sect C 71:3–8. https://doi.org/10.1107/S2053229614024218
22. Farrugia LJ (2012) WinGX and ORTEP for Windows: an update. J Appl Crystallogr 45:849–854. https://doi.org/10.1107/S0021889812029111
23. Macrae CF, Bruno IJ, Chisholm JA et al (2008) Mercury CSD 2.0—new features for the visualization and investigation of crystal structures. J Appl Crystallogr 41:466–470. https://doi.org/10.1107/S0021889807067908
24. Spek AL (2003) Single-crystal structure validation with the program PLATON. J Appl Crystallogr 36:7–13. https://doi.org/10.1107/S0021889802022112
25. McKinnon JJ, Spackman MA, Mitchell AS (2004) Novel tools for visualizing and exploring intermolecular interactions in molecular crystals. Acta Crystallogr Sect B Struct Sci 60:627–668. https://doi.org/10.1107/S0108768104020300
26. Allen FH (2002) The Cambridge structural database: a quarter of a million crystal structures and rising. Acta Crystallogr Sect B Struct Sci 58:380–388. https://doi.org/10.1107/S0108768102003890

27. Groom CR, Allen FH (2014) The Cambridge structural database in retrospect and prospect. Angew Chemie 53:662–671. https://doi.org/10.1002/anie.201306438
28. McKinnon JJ, Jayatilaka D, Spackman MA (2007) Towards quantitative analysis of intermolecular interactions with Hirshfeld surfaces. Chem Commun. https://doi.org/10.1039/b704980c
29. Spackman MA, McKinnon JJ (2002) Fingerprinting intermolecular interactions in molecular crystals. CrystEngComm 4:378–392. https://doi.org/10.1039/B203191B
30. Peón A, Naulaerts S, Ballester PJ (2017) Predicting the reliability of drug-target interaction predictions with maximum coverage of target space. Sci Rep 7:3820. https://doi.org/10.1038/s41598-017-04264-w
31. (2017) OMEGA v.2.5.1: OpenEye Scientific Software, Santa Fe, NM, USA. http://www.eyesopen.com
32. Hawkins PCD, Skillman GA, Warren GL et al (2010) Conformer generation with OMEGA: algorithm and validation using high quality structures from the protein databank and cambridge structural database. J Chem Inf Model 50:572–584. https://doi.org/10.1021/ci100031x
33. Jakalian A, Jack DB, Bayly CI (2002) Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II, parameterization and validation. J Comput Chem 23:1623–1641. https://doi.org/10.1002/jcc.10128
34. OpenEye Scientific Software Inc. (2017) QUACPAC 1.6.3
35. Søndergaard CR, Olsson MHM, Rostkowski M, Jensen JH (2011) Improved treatment of ligands and coupling effects in empirical calculation and rationalization of p K a values. J Chem Theory Comput 7:2284–2295. https://doi.org/10.1021/ct200133y
36. Banks JL, Beard HS, Cao Y et al (2005) Integrated modeling program, applied chemical theory (IMPACT). J Comput Chem 26:1752–1780. https://doi.org/10.1002/jcc.20292
37. McGann M (2011) FRED pose prediction and virtual screening accuracy. J Chem Inf Model 51:578–596. https://doi.org/10.1021/ci100436p
38. McGann M (2012) FRED and HYBRID docking performance on standardized datasets. J Comput Aided Mol Des 26:897–906. https://doi.org/10.1007/s10822-012-9584-8
39. (2017) OEDocking v.3.2.0: OpenEye Scientific Software, Santa Fe, NM, USA. http://www.eyesopen.com
40. McGann MR, Almond HR, Nicholls A et al (2003) Gaussian docking functions. Biopolymers 68:76–90
41. Braga RC, Alves VM, Silva MFB et al (2014) Tuning HERG out: antitarget QSAR models for drug development. Curr Top Med Chem 14:1399–1415. https://doi.org/10.2174/1568026614666140506124442
42. Braga RC, Alves VM, Silva MFB et al (2015) Pred-hERG: a novel web-accessible computational tool for predicting cardiac toxicity. Mol Inform 34:698–701. https://doi.org/10.1002/minf.201500040
43. Cheng F, Li W, Zhou Y et al (2012) admetSAR: a comprehensive source and free tool for assessment of chemical ADMET properties. J Chem Inf Model 52:3099–3105. https://doi.org/10.1021/ci300367a
44. Shen J, Cheng F, Xu Y et al (2010) Estimation of ADME properties with substructure pattern recognition. J Chem Inf Model 50:1034–1041. https://doi.org/10.1021/ci100104j
45. Cheng F, Yu Y, Shen J et al (2011) Classification of cytochrome P450 inhibitors and noninhibitors using combined classifiers. J Chem Inf Model 51:996–1011. https://doi.org/10.1021/ci200028n
46. Hansen K, Mika S, Schroeter T et al (2009) Benchmark data set for in silico prediction of ames mutagenicity. J Chem Inf Model 49:2077–2081. https://doi.org/10.1021/ci900161g
47. Lagunin A, Filimonov D, Zakharov A et al (2009) Computer-aided prediction of rodent carcinogenicity by PASS and CISOC-PSCT. QSAR Comb Sci 28:806–810. https://doi.org/10.1002/qsar.200860192
48. Baell JB, Holloway GA (2010) New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. J Med Chem 53:2719–2740. https://doi.org/10.1021/jm901137j
49. Baell J, Walters MA (2014) Chemistry: chemical con artists foil drug discovery. Nature 513:481–483
50. Spackman MA, Jayatilaka D (2009) Hirshfeld surface analysis. CrystEngComm 11:19–32. https://doi.org/10.1039/B818330A
51. Rognan D (2007) Chemogenomic approaches to rational drug design. Br J Pharmacol 152:38–52
52. Klabunde T (2007) Chemogenomic approaches to drug discovery: similar receptors bind similar ligands. Br J Pharmacol 152:5–7
53. Westermaier Y, Barril X, Scapozza L (2015) Virtual screening: an in silico tool for interlacing the chemical universe with the proteome. Methods 71:44–57. https://doi.org/10.1016/j.ymeth.2014.08.001
54. Buchman CD, Hurley TD (2017) Inhibition of the aldehyde dehydrogenase 1/2 family by Psoralen and Coumarin derivatives. J Med Chem 60:2439–2455. https://doi.org/10.1021/acs.jmedchem.6b01825
55. Gaulton A, Bellis LJ, Bento AP et al (2012) ChEMBL: a large-scale bioactivity database for drug discovery. Nucl Acids Res 40:D1100–D1107. https://doi.org/10.1093/nar/gkr777
56. Tomita H, Tanaka K, Tanaka T, Hara A (2016) Aldehyde dehydrogenase 1A1 in stem cells and cancer. Oncotarget 7:11018–11032. https://doi.org/10.18632/oncotarget.6920
57. van de Waterbeemd H, Gifford E (2003) ADMET in silico modelling: towards prediction paradise? Nat Rev Drug Discov 2:192–204. https://doi.org/10.1038/nrd1032
58. Alqahtani S (2017) In silico ADME-Tox modeling: progress and prospects. Expert Opin Drug Metab Toxicol 13:1147–1158. https://doi.org/10.1080/17425255.2017.1389897

Affiliations

W. F. Vaz[1,2] · J. M. F. Custodio[3] · G. D. C. D'Oliveira[3] · B. J. Neves[4] · P. S. C. Junior[5] · J. T. M. Filho[3] · C. H. Andrade[4] · C. N. Perez[3] · E. P. Silveira-Lacerda[6] · H. B. Napolitano[1]

1 Universidade Estadual de Goiás, Anápolis, GO 75132-400, Brazil

2 Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso, Lucas do Rio Verde, MT 78455-000, Brazil

3 Universidade Federal de Goiás, Goiânia, GO 74690-900, Brazil

4 LabMol, Faculdade de Farmácia, Universidade Federal de Goiás, Goiânia, GO 74605-170, Brazil

5 Universidade Federal de Mato Grosso do Sul, Nova Andradina, MS 79750-000, Brazil

6 Laboratório de Genética Molecular e Citogenética, Instituto de Ciências Biológicas, Universidade Federal de Goiás, Goiânia, GO 74605-170, Brazil

Journal of Molecular Structure 1206 (2020) 127732

Contents lists available at ScienceDirect

Journal of Molecular Structure

journal homepage: http://www.elsevier.com/locate/molstruc

Machine learning prediction of the potential pesticide applicability of three dihydroquinoline derivatives: Syntheses, crystal structures and physical properties

Wesley F. Vaz [a, b], Giulio D.C. D'Oliveira [c], Caridad N. Perez [c], Bruno J. Neves [c], Hamilton B. Napolitano [a, d, *]

[a] Universidade Estadual de Goiás, 75132-400, Anápolis, GO, Brazil
[b] Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso, 78455-000, Lucas do Rio Verde, MT, Brazil
[c] Universidade Federal de Goiás, 74690-900, Goiânia, GO, Brazil
[d] Centro Universitário de Anápolis, 75075-010, Anápolis, GO, Brazil

ARTICLE INFO

Article history:
Received 15 October 2019
Received in revised form
7 January 2020
Accepted 13 January 2020
Available online 22 January 2020

Keywords:
Machine learning model
Tobacco mosaic virus
Fusarium oxysporum
dihydroquinoline
Crystal structure

ABSTRACT

Increasingly machine learning processes have been applied in the search and development of compounds that may have specific physicochemical properties to the desired application. This article describes how a machine learning model led us to the synthesis of three dihydroquinoline derivatives with potential application as a pesticide. The synthesized compounds were predicted to be active against the *Tobacco mosaic virus* (≅ 90%) and *Fusarium oxysporum* (≅ 78%). Regarding a correlation between the pesticide activity and the molecular structure, the new dihydroquinoline derivatives were structurally characterized using spectroscopic techniques and single crystal X-ray diffraction. They crystallized into orthorhombic (I) and monoclinic (II and III) crystal systems with supramolecular arrangements maintained primarily by non-classical C–H···O hydrogen bonds, which form dimers and chains in their molecular packaging. Frontier molecular orbitals and molecular electrostatic potential maps were undertaken using density functional theory in order to study the electronic properties of the observed molecular conformations. Finally, the developed approach is a useful tool on new pesticide investigation when experimental toxicity data are not available.

1. Introduction

Pesticides include herbicides, insecticides, fungicides, fumigants, and rodenticides and offer significant benefits in public health and food production. Unlike most other significant chemicals, pesticides are designed to impact living systems [1]. As a result, there is a worry regarding the environmental and human consequences of extensive pesticide use [1,2]. Several harmful effects on the overuse of pesticides include (1) potential destruction of biodiversity; (2) problems on environmental sustainability and global stability [2]; and (3) potential contamination on food, water and the environment [3]. Therefore, it is important to develop new compounds less harmful to living systems and machine learning is a powerful methodology regarding its wide use of molecular engineering applied to design compounds with desired properties [4]. Machine learning models have long been employed for drug discovery utilizing docking studies, virtual screening, molecule synthesis, small molecule physicochemical properties, solubility and beyond [5]. Moreover, machine learning methods have also been used to discover solutions to environmental issues, such as municipal solid waste management [6], soil/compost properties with bioavailability and risk assessment [7], to model the bioavailability of contaminants [8], and to obtain new antifungal and antiviral agents in agriculture [9–12].

To propose new compounds with biological potential, we developed a machine learning model from a training set of molecules with known activity. In this case, was assumed that the potential pesticide is associated with the chemical structure. The obtained model shows that alpha-beta unsaturated systems, sulfonamide group, and the nitro group are frequently founded in

* Corresponding author. Universidade Estadual de Goiás, 75132-400, Anápolis, GO, Brazil.
E-mail address: hamilton@ueg.br (H.B. Napolitano).

https://doi.org/10.1016/j.molstruc.2020.127732
0022-2860/

commercials pesticides. All those characteristics could be combined in a quinolinone moiety to produce a molecular structure favorable to pesticide action (Fig. 1). Furthermore, it is recognized that the application of quinoline compounds has been broadened high efficient pesticide agents [13–16].

Structurally, quinolones are compounds containing a 4-oxo-1,4-dihydroquinoline skeleton and they differ from a dihydroquinoline by the absence of a double bond between carbons 2 and 3 in the dihydroquinoline skeleton [17]. There are few crystalline structures, possessing the 2,3-dihydroquinolin-4(1H)-one nuclei, reported in the Cambridge Structural Database (CSD) [20]. Performing a CSD search (Version 5.40, Feb 2019), only 126 hits were found. Only for comparison purposes, when this same search is done for molecules with the chalcone scaffold, compounds from which quinolines can be generated, 3314 occurrences are found. Therefore, there is much to be studied in terms of the structural characteristics and crystalline state of dihydroquinolinones. This paper details the machine learning model applied to predict new compounds with potential biological activity. Based on the results of that model we synthesized three dihydroquinolinones: (*E*)-3-(4-chlorobenzylidene)-2-(4-ethoxyphenyl)-1-(phenylsulfonyl)-2,3-dihydroquinolin 4(*1H*)-one (**I**), (*E*)-2-(4-chlorophenyl)-3-(3-nitrobenzylidene)-1-(phenylsulfonyl)-2,3-dihydroquinolin-4(*1H*)-one (**II**) and (*E*)-3-(4-chlorobenzylidene)-2-(2-chlorophenyl)-1-(phenylsulfonyl)-2,3-dihydroquinolin-4(*1H*)-one (**III**), and presented a comprehensive study on crystal and electronic structure of these potential pesticide agents. Moreover, the importance of this is study is to show one of the great advantages of using in silico models to predict the bioactive potential of compounds that are diminishing the time and the requirement for experimental testing and, consequently, the cost related to it [46].

2. Experimental and computational approaches

2.1. Synthesis and spectroscopic analysis

The three compounds were synthesized using a previously reported methodology [21]. Basically, a chalcone precursor (1.0 mmol) and a substituted benzaldehyde (2.0 mmol) were dissolved in 15 mL of basic ethanol (KOH) and reacted at room temperature for 48 h (Scheme 1). The solution was filtered, and the precipitate was rinsed with 15 mL of ethanol. The precipitate was dissolved in dichloromethane (10 mL) and this solution was extracted with water. The organic phase could evaporate slowly, yielding the desired product. (*E*)-3-(4-chlorobenzylidene)-2-(4-ethoxyphenyl)-1-(phenylsulfonyl)-2,3-dihydroquinolin-4(1*H*)-one (**I**). Pale yellow crystalline solid, yield 66.2%, purity of 98.0%, MP 190–192 °C. ^{1}H NMR ($CDCl_3$) δ 1.36 (*t*, *J* 7.00 Hz, 3H), δ 3.95 (*q*, *J* 7.00 Hz, 2H), δ 6.56 (*d*, *J* 0.60 Hz, 1H), δ 6.77–6.80 (m, 2H), δ 7.10–7.13 (m, 4H), δ 7.19–7.23 (m, 2H), δ 7.27–7.32 (m, 3H), δ 7.37–7.40 (m, 2H), δ 7.49 (*tt*, *J* 1.25 Hz, 7.45 Hz, 1H), δ 7.52 (s, 1H), δ 7.54 (*ddd*, *J* 1.66 Hz, 7.36, *Hz*, 8.19 Hz, 1H), δ 7.71 (*dd*, *J* 1.05 Hz, 8.20 Hz, 1H), δ 7.89 (*dd*, *J* 1.65 Hz, 7.80 Hz, 1H); ^{13}C NMR ($CDCl_3$) δ 14.8, 59.6, 63.4, 114.7, 127.2, 127.9, 128.3, 128.5, 128.5, 128.8, 128.9, 129.4, 131.3, 131.3, 132.2, 133.2, 134.6, 136.4, 137.6, 138.6, 138.9, 159.0, 182.7; IR 1677 (m), 1609 (m), 1474 (w), 1354 (s), 1306 (w), 1247 (s); HRMS calculated for $C_{30}H_{24}ClNO_4S$ 530.1193, found 530.1083. (*E*)-2-(4-chlorophenyl)-3-(3-nitrobenzylidene)-1-(phenylsulfonyl)-2,3-dihydroquinolin-4(1*H*)-one (**II**). Pale yellow to white crystalline solid, yield 93.3%, purity of 98.4%, MP 206–208 °C.

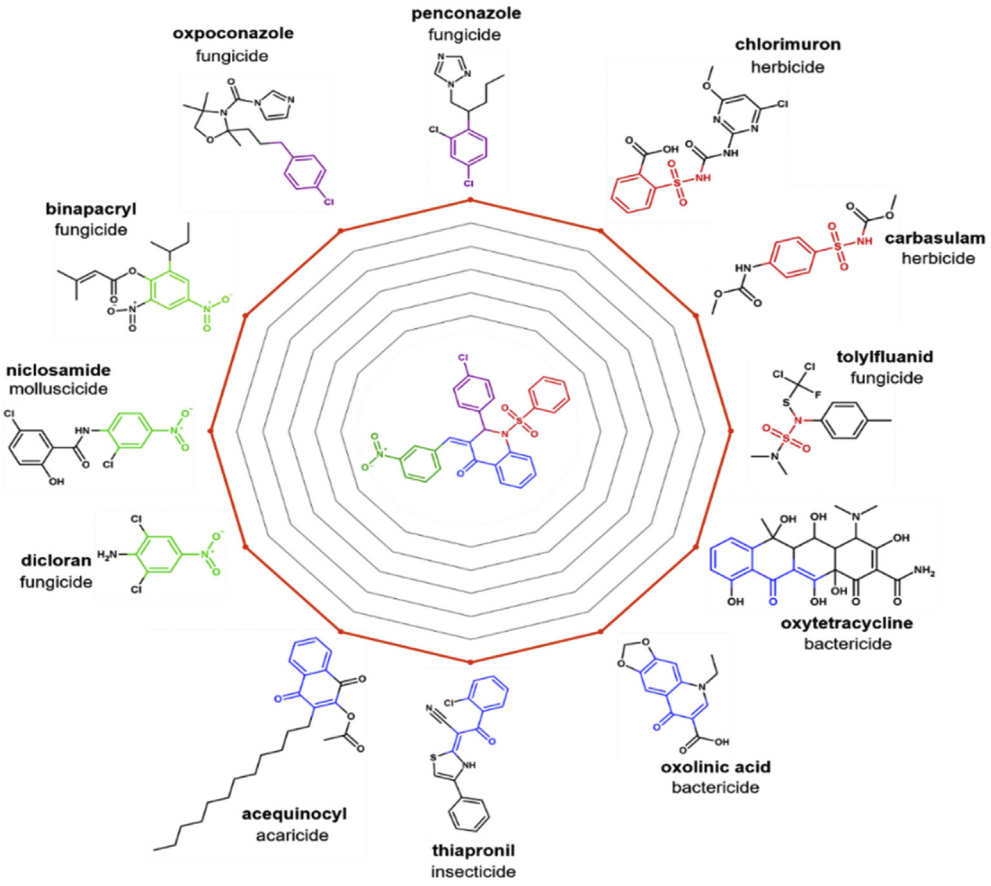
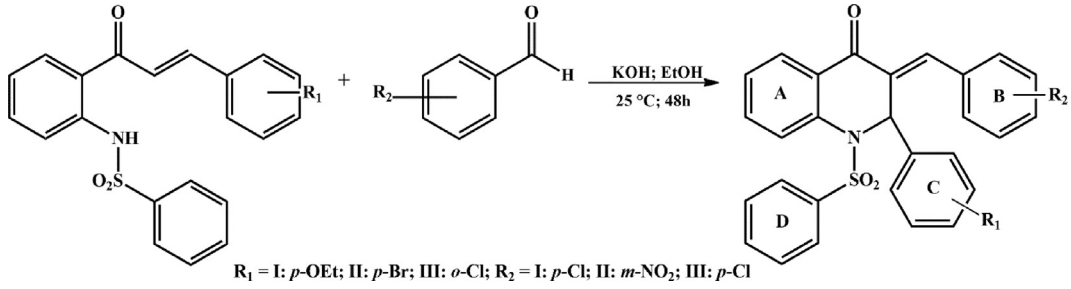


Fig. 1. Privileged substructures for the design of pesticides.



Scheme 1. Conditions for the synthesis of compounds **I**, **II** and **III**.

^{1}H NMR ($CDCl_3$) δ 6.47 (s, 1H), δ 7.08–7.11 (m, 2H), δ 7.24–7.28 (m, 2H), δ 7.28–7.30 (m, 2H), δ 7.34 (*ddd*, *J* 1.11 Hz, 7.41 Hz, 7.74 Hz, 1H), δ 7.36–7.38 (m, 1H), δ 7.43–7.45 (m, 2H), δ 7.55–7.61 (m, 3H), δ 7.62 (s, 1H), δ 7.70 (*dd*, *J* 1.03 Hz, 8.18 Hz, 1H), δ 7.91 (*dd*, *J* 1.62 Hz, 7.83 Hz, 1H), δ 8.03 (*t*, *J* 1.32, 1H), δ 8.30 (*ddd*, *J* 0.84 Hz, 2.21 Hz, 8.24 Hz, 1H); ^{13}C NMR ($CDCl_3$) δ 59.4, 123.1, 124.6, 125.1, 127.2, 127.7, 127.9, 128.2, 128.4, 129.1, 129.2, 130.4, 132.4, 132.8, 133.7, 134.2, 135.1, 135.2, 135.9, 137.3, 137.4, 138.8, 148.5, 182.0; IR 1675 (m), 1599 (m), 1478 (w), 1353 (s), 1305 (w), 1236 (m); HRMS calculated for $C_{28}H_{19}BrN_2O_5S$ 575.0276, found 575.0129. (*E*)-3-(4-chlorobenzylidene)-2-(2-chlorophenyl)-1-(phenylsulfonyl)-2,3-dihydroquinolin-4(1*H*)-one (**III**). Pale yellow crystalline solid, yield 57.6%, purity of 99.2%, MP 226–228 °C. ^{1}H NMR ($CDCl_3$) δ 6.76 (*dd*, *J* 1.35 Hz, 7.75 Hz, 1H), δ 6.95 (*dt*, *J* 1.30 Hz, 7.58 Hz, 1H), δ 6.96 (s, 1H), δ 7.08–7.11 (m, 2H), δ 7.11–7.13 (m, 2H), δ 7.20 (*ddd*, *J* 1.51 Hz, 7.44 Hz, 7.96 Hz, 1H), δ 7.23–7.26 (m, 2H), δ 7.35 (*ddd*, *J* 2.46 Hz, 6.09 Hz, 7.76 Hz, 1H), δ 7.40–7.43 (m, 2H), δ 7.49–7.54 (m, 4H), δ 7.57 (s, 1H), δ 7.89–7.91 (m, 1H); ^{13}C NMR ($CDCl_3$) δ 58.7, 126.7, 127.7, 128.0, 128.6, 129.2, 129.6, 129.9, 130.0, 130.2, 131.2, 131.2, 131.3, 132.1, 133.4, 134.7, 134.9, 135, 136.6, 136.7, 139.0, 139.0, 182.6; IR 1674 (m), 1604 (m), 1474 (w), 1359 (m), 1294 (m), 1239 (m); HRMS calculated for $C_{28}H_{19}Cl_2NO_3S$ 520.0541, found 520.0458. The NMR spectrums are available in supplementary data (Figs. S1–S10).

2.2. Crystal structure analysis

Suitable crystals of compounds **I**, **II** and **III** were carefully selected. The crystal of compound **I** was mounted in a Bruker APEX II CCD diffractometer with, graphite-monochromated MoKα radiation (λ = 0.71073 Å) and, data were measured at 120 K. Whereas the data collection of crystals of compounds **II** and **III** were carried out in the Agilent SuperNova diffractometer with MoKα radiation at room temperature (298 K). Throughout Olex2 system [22], the structure solutions were determined by direct methods with SHELXS [23] and refined by full-matrix least-squares on F^2 with SHELXL [24]. All the hydrogen atoms were placed in calculated positions and refined with fixed individual displacement parameters [U_{iso}(H) = 1.2U_{eq} (C) or 1.5U_{eq} (C)] according to the riding model (C–H bonds equal 0.93 Å for aromatic). Finally, the validation of chemical parameters was made using PARST [25], CRYSALISPRO [26] and PLATON [27]. Crystal, data collection, and structure refinement details are summarized in Table 1. The crystallographic information files of **I**, **II** and **III** were deposited in the Cambridge Structural Database under the codes 1953384, 1953387 and 1953385, respectively. The crystallographic structures were used to optimize the geometric parameters, and to calculate frontier molecular orbital (FMO) energies and molecular electrostatic potential maps (MEP). All calculations were undertaken throughout Gaussian09 package [28] through density functional theory (DFT) using B3LYP/6-311 + G(d) as functional and basis set, respectively.

2.3. Computational approach

Data Integration and Curation. All compounds were retrieved with biological and toxicological properties for a pool of plant pathogens and organism models (Table 3) from ChEMBL [29], PubChem Bioassay [30], and ECOTOX [31] databases. Before, activity thresholds were used to set active ($\leq$50 μM) and inactive compounds (>50 μM). Finally, all chemical structures and correspondent biological information (EC_{50}, MIC, or LC_{50}) were carefully standardized using Standardizer v.16.9.5.0 (ChemAxon, Budapest, Hungary, http://www.chemaxon.com) according to the protocols proposed by Fourches and colleagues [32–34]. Briefly, explicit hydrogens were added, whereas polymers, salts, metals, organometallic compounds, and mixtures were removed. Also, specific chemotypes such as aromatic rings and nitro groups were normalized. Also, we performed the analysis and exclusion of duplicates, as follows: (i) if duplicates presented discordance in biological activity, both entries would be excluded; and (ii) if the reported outcomes of the duplicates were the same, one entry would be retained in the dataset and the other excluded. Moreover, considering the different sizes of classes, the curated datasets were balanced using a linear under-sampling approach [35].

Machine learning modeling. Models were developed using pairwise combinations of FeatMorgan fingerprints (FCFP-like, radius 2 and bit vector of 2048 bits) calculated in RDKit (http://www.rdkit.org) [36] and Random Forest algorithm implemented in Scikit-learn v.0.19.2 (http://scikit-learn.org/) package available on Python v.3.6 (https://www.python.org) [37]. The five-fold external cross-validation (5FECV) approach was chosen for the estimation of the predictivity of developed models. The predictive performance of categorical models was evaluated using sensitivity (SE), specificity (SP), correct classification rate (CCR), positive predictive value (PPV), and negative predictive value (NPV). These metrics were calculated as follows:

$$SE = \frac{TP}{TP + FN} \quad (1)$$

$$SP = \frac{TN}{TN + FP} \quad (2)$$

$$CCR = \frac{SE + SP}{2} \quad (3)$$

Table 1
Crystal data and structure refinement for **I**, **II** and **III**.

	I	II	III
Crystal data			
Chemical formula	$C_{30}H_{24}ClNO_4S$	$C_{28}H_{19}BrN_2O_5S$	$C_{28}H_{19}Cl_2NO_3S$
M_r	530.01	575.42	520.40
Crystal system, space group	Orthorhombic, *Pbca*	Monoclinic, $P2_1/n$	Monoclinic, $P2_1/c$
Temperature (K)	296	294	293
a, *b*, *c* (Å)	15.2301 (7), 9.3995 (4), 36.7938 (16)	11.5059 (9), 15.6336 (9), 13.8456 (10)	12.6841 (7), 11.5547 (5), 17.4956 (9)
α, β, γ (°)	90, 90, 90	90, 100.280 (8), 90	90, 104.740 (5), 90
V (Å^3)	5267.2 (4)	2450.5 (3)	2479.8 (2)
Z	8	4	4
Radiation type	Mo *K*α	Mo *K*α	Mo *K*α
μ (mm^{-1})	0.26	1.81	0.38
Data collection			
Diffractometer	Bruker *APEX*-II CCD	SuperNova, Dual, Cu at zero, AtlasS2	SuperNova, Dual, Cu at zero, AtlasS2
Absorption correction	Multi-scan *SADABS2014*/5 was used for absorption correction. w*R*2(int) was 0.0755 before and 0.0502 after correction. The Ratio of minimum to maximum transmission is 0.9163. The λ/2 correction factor is 0.00150.	Multi-scan *CrysAlis PRO* 1.171.38.41 Empirical absorption correction using spherical harmonics, implemented in SCALE3 ABSPACK scaling algorithm.	Multi-scan *CrysAlis PRO* 1.171.38.43 Empirical absorption correction using spherical harmonics, implemented in SCALE3 ABSPACK scaling algorithm.
T_{min}, T_{max}	0.683, 0.745	0.695, 1.000	0.616, 1.000
No. of measured, independent and observed [$I > 2\sigma(I)$] reflections	35977, 5382, 4057	26684, 6226, 4699	58317, 6628, 5029
R_{int}	0.036	0.032	0.045
(sin θ/λ)$_{max}$ (Å^{-1})	0.625	0.698	0.699
Refinement			
$R[F^2 > 2\sigma(F^2)]$, w$R(F^2)$, *S*	0.045, 0.119, 1.07	0.040, 0.107, 1.02	0.047, 0.120, 1.03
No. of reflections	5382	6226	6628
No. of parameters	335	334	316
H-atom treatment	H-atom parameters constrained	H-atom parameters constrained	H-atom parameters constrained
$\Delta\rho_{max}$, $\Delta\rho_{min}$ (e Å^{-3})	0.19, −0.36	0.55, −0.76	0.25, −0.39

$$PPV = \frac{TP}{TP + FP} \quad (4)$$

$$NPV = \frac{TN}{TN + FN} \quad (5)$$

Here, TP and TN represent the number of true positives and true negatives, while FP and FN represent the number of false positives and false negatives, respectively.

Applicability domain (AD). The AD was estimated based on the Euclidean distances among the training set of each QSAR model generated in the 5FECV procedure. The distance of a test set compound to its nearest neighbor in the training set was compared to the predefined AD threshold level. The prediction was less reliable if the distance was greater than the threshold level. In our study, the AD was defined as a distance threshold (D_T) between a compound under prediction and the closest neighbors in the training set. The following equation was used for the calculation of the distance threshold:

$$D_T = \bar{y} + Z\sigma \quad (6)$$

In which ȳ is the average Euclidean distance of the k nearest neighbors within the modeling set, σ is the standard deviation of these Euclidean distances, and Z is an arbitrary parameter to control the significance level. We set the default value of this parameter Z at 0.5. If the compound distance exceeded the threshold, the prediction was considered to be less trustworthy [38,39].

3. Results and discussion

3.1. Molecular structure of bioactive dihydroquinolines

The three dihydroquinolin-4(1*H*)-one reported in this study crystallized in the following space groups: *Pbca* (**I**), $P2_1/n$ (**II**) and $P2_1/c$ (**III**). All of them have one molecule, with *R* conformation, in the asymmetric unity. The angle formed by the planes between the rings A and B is 54.34° (**I**), 64.60° (**II**), and 56.12° (**III**) showing that these compounds do not show planarity in their molecular structure. This feature also happens in previously reported dihydroquinolin-4(1*H*)-one crystal structures [40–42].

Mogul software was used to check how appropriate are the obtained 3D geometries. In **I** there are two bond angles (C9–C1–N1 and C17–C1–N1) and one torsion angle (C2–C1–C17–C18) with unusual values. In general, these angles are slightly larger (2° and 4°, respectively) than the mean value, while similar values for the torsion (−177.53°) are founded just only in 1.44% of the analyzed structures. In **II** there are three unusual values related to the torsions N1–C1–C17–C18 (26.82°), N1–C1–C17–C22 (−157.11°) and C24–C23–S1–O2 (−130.62°). Similar values for these torsions are founded approximately just in 1% of the structures of the analyzed molecular set. Finally, in **III** only the torsion C2–C1–C17–C22 (156.50) is unusual. The overlay of the three structures shows significant conformational differences related to their aromatic rings. The rings C and D in **III** in the opposite way when compared with **I** and **II** even the three structures having *R* conformation (Fig. 2). The other differences could be characterized by the torsion angles: C2–C10–C11–C16 = 153.5(2)°, 142.8(2)° and −145.71(19)°; C2–C1–C17–C18 = −177.5(2)°, 148.38(19)° and −23.4(2)° and

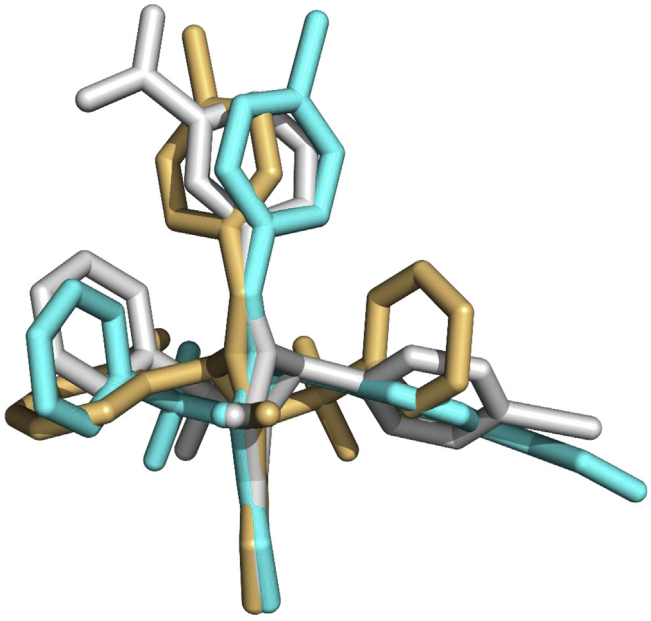


Fig. 2. The overlap of the three dihydroquinolin-4(1*H*)-one molecules under study, showing **I** (cyan), **II** (grey), and **III** (orange). H atoms have been omitted for clarity.

N1—S1—C23—C24 = 95.4(2)°, 115.0(2)° and 80.32(15)° for **I**, **II** and **III**, respectively.

The molecular packing of **I** is dominated by non-classical C—H···O hydrogen bonds (Table 2). The O3 atom of SO_2 group and an H atom of the ring A are involved in a $C_1^1(8)$ chain along to the *b* axis, while the O atom of carbonyl group an H atom of the ring B are involved in a $C_1^1(7)$ chain along to the *a* axis. The combination of these two chains leads to the formation of molecular pairs with $R_3^3(22)$ arrangement (Fig. 3a and b). Moreover, O2 atom of SO_2 group and an H atom of the ring D forms a $C_1^1(7)$ zigzag chain along to the *a* axis (Fig. 3c).

The molecular packing of **II** is also dominated by non-classical C—H···O hydrogen bonds (Table 2), but unlike what occurs in **I,** here dimer formation is predominant. The O3 atom of SO_2 group and the H12 atom of the ring B, the O1 atom of the carbonyl group and the H21 atom of the ring C and the O5 atom of NO_2 group and the H26 atom of the ring D are involved in a formation of a $R_2^2(18)$, $R_2^2(16)$ and $R_2^2(30)$ dimers, respectively (Fig. 4a, b, 4c). The only chain $C_1^1(12)$, formed in these packing grows along to the *b* axis in a zigzag way and it is a consequence of the interaction between the O5 atom of NO_2 group and the H6 atom of the ring A (Fig. 4d).

The crystal packing of compound **III,** just like in **I** and **II,** has non-classical C—H···O hydrogen bonds, listed in Table 2. A $R_2^2(16)$ dimer is formed by the interaction between the O1 atom of the carbonyl group and H19 atom of the ring D (Fig. 5a). In this packaging, there are also two chains: one $C_1^1(7)$ involving O3 atom from SO_2 group and the H5 atom of the ring A (Fig. 5b) and the other, $C_1^1(6)$, involving O2 atom of group SO_2 and H27 atom of the ring D (Fig. 5c). Both chains grow in a zigzag way along to the *b* axis.

The optimized structures were superimposed with those obtained experimentally (Fig. S13). It is noticed that in the gas phase the rings B, C, and D are more movable, and this was expected since the solid-state interactions are absent in these models. Another variation observed is the elongation, on average 0.67 Å, in the bond lengths involving the sulfur atom. The calculated FMO of **I**, **II** and **III** are represented in Fig. 6. In **I**, the highest occupied molecular orbital (HOMO) is concentrated in ring C while the lowest unoccupied molecular orbital (LUMO) extends over all dihydroquinoline-4(1H)-one moiety. In **II** HOMO extends over the rings A, B and D while LUMO is concentrated in ring C. In **III,** HOMO is localized over the vinyl group and the ring B while LUMO extends over all dihydroquinoline-4(1H)-one moiety. From the values of energy [43], the following global reactivity descriptors were calculated: E_{gap}, hardness and softness. For **I** these values are 3.60 eV, 1.80 eV and 7.20 eV, respectively. For **II** these values are 3.77 eV, 1.88 eV and 7.54 eV, respectively. Finally, for **III** these values are 3.60 eV, 1.80 eV and 7.20 eV, respectively. These values indicate that **III** is the highest electronic kinetic stable compound than **I** and **II**. Also, **III** has a higher resistance to transfer charge and is more susceptible to receive electrons than **I** and **II**.

The MEP representation of **I**, **II** and **III** are shown in Fig. 7. The high electron density sites where located around O atoms of carbonyl, sulfonyl and nitro groups while low electron density sites are located at aromatic rings. These characteristics were also observed in other dihydroquinolines reported in the literature [40—42]

Table 2
Hydrogen-bond geometry (Å, °) for **I**, **II** and **III**.

Compound	*D*—H···*A*	*D*—H	H···*A*	*D*···*A*	*D*—H···*A*
I	C6—H6···O3^{i}	0.93	2.48	3.289 (3)	145
	C16—H16···O1^{ii}	0.93	2.52	3.382 (3)	155
	C26—H26···O2^{iii}	0.93	2.55	3.248 (4)	132
II	C6—H6···O5^{iv}	0.93	2.42	3.331 (3)	165
	C12—H12···O3^{v}	0.93	2.57	3.187 (3)	124
	C21—H21···O1^{vi}	0.93	2.64	3.514	157
	C26—H26···O5^{vii}	0.93	2.62	3.417	144
III	C19—H19···O1^{viii}	0.93	2.54	3.303 (3)	139
	C27—H27···O2^{ix}	0.93	2.48	3.314 (3)	150
	C5—H5···O3	0.93	2.62	3.313	132

Symmetry codes: (i) $x, y+1, z$; (ii) $-x-1/2, y-1/2, z$; (iii) $x-1/2, y, -z+3/2$; (iv) $-x+1/2, y+1/2, -z+3/2$; (v) $-x+1, -y+1, -z+1$; (vi) $-x, -y+1, -z+1$; (vii) $-x+1, -y, -z+1$; (viii) $-x+2, -y+1, -z+1$; (ix) $-x+1, y+1/2, -z+1/2$.

Table 3
Summarized statistical characteristics of categorical models.

Model (class)	CCR	SE	SP	PPV	NPV	Coverage
Predictive pesticide activity against several organisms						
Tobacco mosaic (virus)	0.83	0.85	0.81	0.82	0.85	0.69
Bemisia tabaci (insect)	0.72	0.62	0.81	0.77	0.68	0.97
Helicoverpa armigera (caterpillar)	0.80	0.80	0.80	0.80	0.80	0.55
Agrobacterium tumefaciens (bacteria)	0.82	0.81	0.84	0.84	0.82	0.78
Botrytis fabae (fungus)	0.74	0.84	0.65	0.80	0.70	0.63
Aspergillus niger (fungus)	0.73	0.74	0.73	0.73	0.74	0.74
Fusarium oxysporum (fungus)	0.89	0.85	0.93	0.92	0.86	0.74
Fusarium solani (fungus)	0.76	0.84	0.70	0.73	0.81	0.70
Rhizoctonia solani (fungus)	0.71	0.73	0.70	0.72	0.71	0.71
Sclerotinia sclerotiorum (fungus)	0.76	0.79	0.72	0.74	0.78	0.72
Predictive ecotoxicology against organism models						
Artemia salina (crustacean)	0.83	0.81	0.85	0.84	0.82	0.88
Apis mellifera (honeybee)	0.79	0.76	0.81	0.80	0.77	0.88
Oncorhynchus mykiss (fish)	0.70	0.69	0.71	0.71	0.69	0.81

CCR: correct classification rate; SE: sensitivity; SP: specificity; PPV: positive predictive value; and NPV: negative predictive value; Coverage: percentage of test set compounds within the applicability domain.

3.2. Machine learning modeling

The combination of FeatMorgan fingerprints with the Random Forest algorithm led to externally predictive categorical models, with correct classification rate (CCR) values ranging between 0.70 and 0.89, sensitivity (SE) of 0.62—0.85, specificity (SP) of 0.65—0.93, and coverage of 0.55—0.97. The details of the statistical

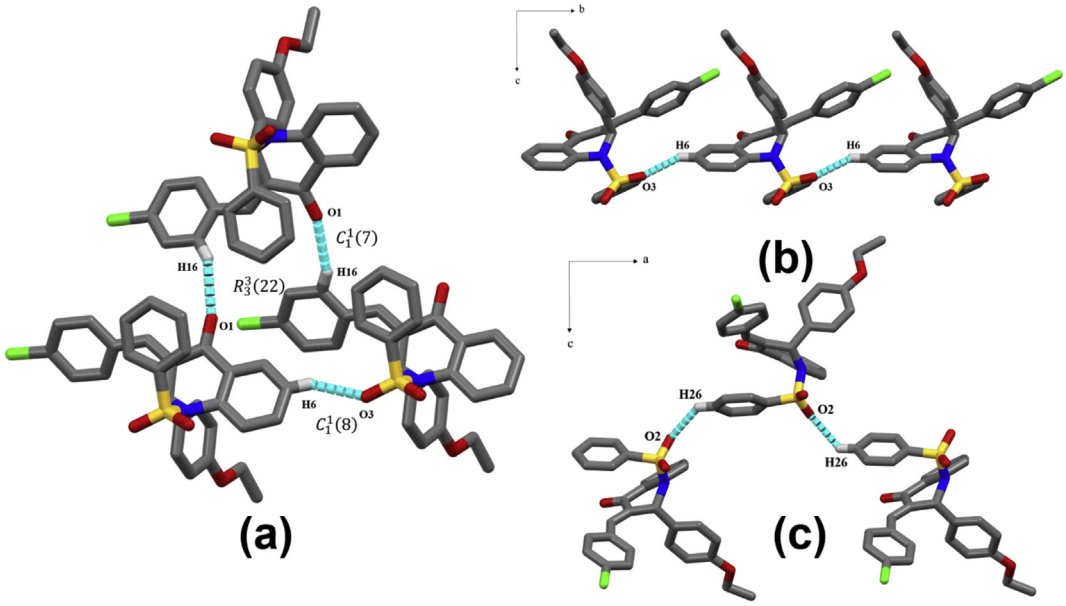


Fig. 3. Main C–H···O interactions in the molecular packing of compound **I**.

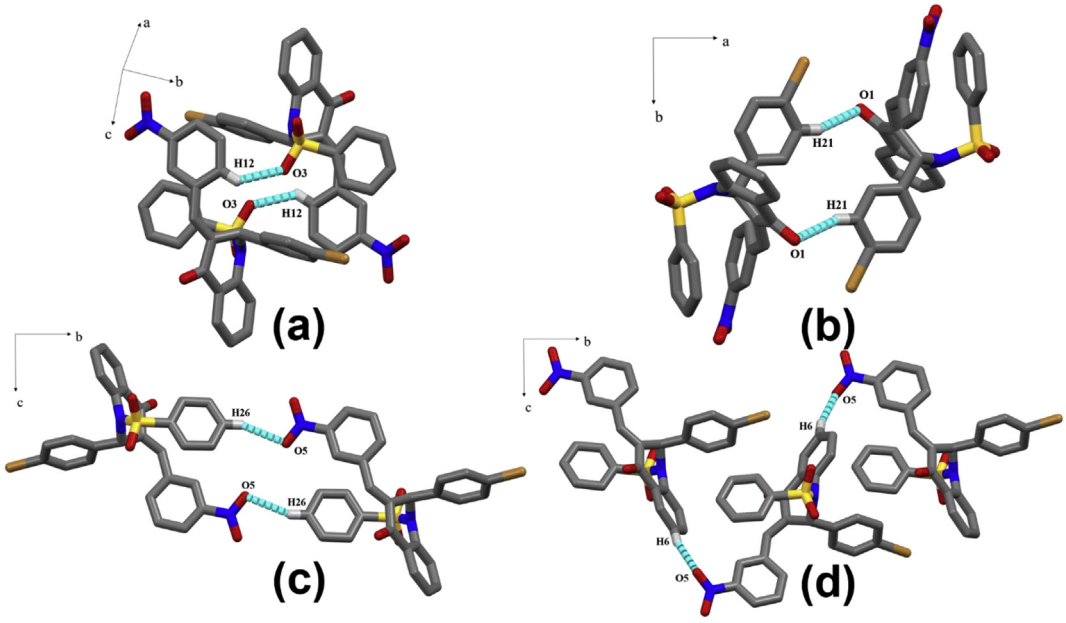


Fig. 4. Main C–H···O interactions in the molecular packing of compound **II**.

characteristics of categorical models are shown in Table 3. The model for *Fusarium oxysporum* (fungus) demonstrated the best performance among all other explored activity endpoints (CCR = 0.89, SE = 0.85, and SP = 0.93). On the other hand, the model developed for *Artemia salina* (CCR = 0.83, SE = 0.81, and SP = 0.85), a species of brine shrimp, showed the best performance between all ecotoxicity models.

The developed machine learning models were used as filters in a virtual screening aiming at prioritizing the putative biological applications of **I**, **II** and **III**. According to Table 4, all compounds were predicted as antiviral agents against *Tobacco mosaic virus* (TMV) (92%, 84% and 96% of probability for **I**, **II** and **III**, respectively), firstly discovered in tobacco, is one of the most devastating viruses to crops, which can infect over 350 different species of plants, including crops of tomato, pepper, cucumber, etc. As one of the most stable viruses, it can survive outside the plant and remain in a dormant state to infect growing crops. Once the plant is infected, no chemical cure is effectively available, and usually, all the infected crops should be removed [44]. Furthermore, all compounds were also predicted as an antifungal agent against *Fusarium oxysporum* (80%, 46% and 80% of probability for **I**, **II** and **III**, respectively). It includes many strains that cause vascular wilt diseases of economically significant crops all over the world [45]. On the other hand, predictive analysis of ecotoxicity potential indicates that **I**, **II**

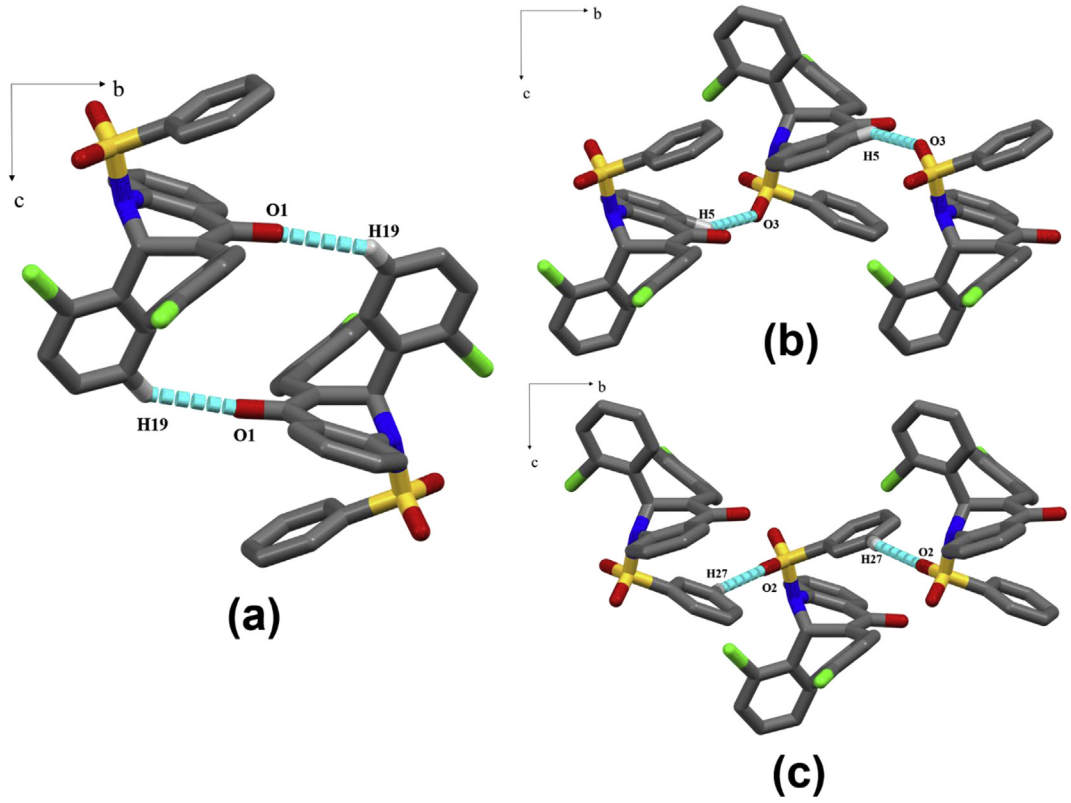


Fig. 5. Main C–H···O interactions in the molecular packing of compound **III**.

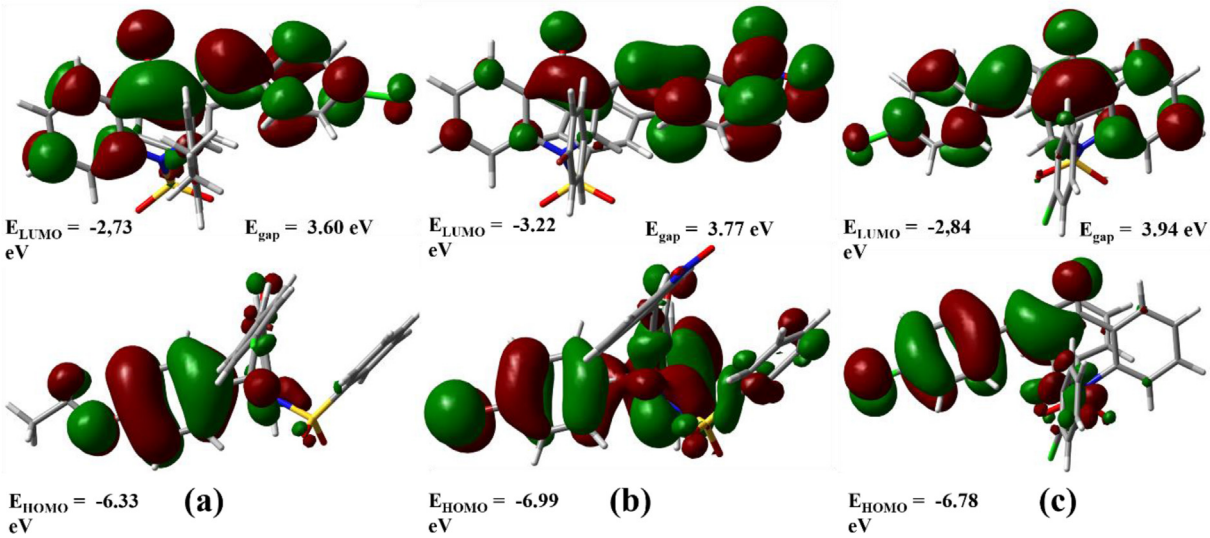


Fig. 6. Frontier molecular orbitals of compounds **I**, **II** and **III**.

and **III** may be toxic to honeybees (*Apis mellifera*, see Table 4).

4. Conclusions

This paper reports the use of a machine learning model along with public databases such as ChEMBL and PubChem Bioassay to guide the synthesis of new compounds with pesticide properties. In addition to the synthesis, spectroscopic characterizations, supramolecular studies and theoretical calculations of physical properties were realized to have a dense description of the proposed molecules. The results showed that dihydroquinoline molecules predicted antiviral (against *Tobacco mosaic*) and antifungal activity (*Fusarium oxysporum*) testifying the pesticide potential of these compounds. **I**, **II** and **III** have their molecular packing stabilized, mainly, by C–H···O interactions, that form dimers and chains. Moreover, the calculated electronic properties, showed **III** as the compound more electronic kinetic stable. Finally, the machine learning method was a powerful drive tool to design new dihydroquinolines derivatives as a potential pesticide.

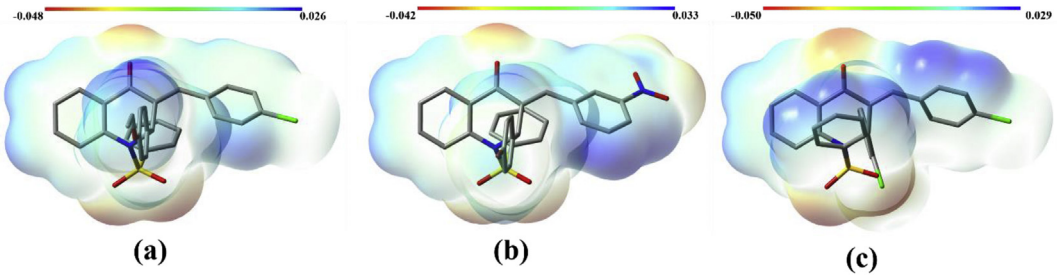


Fig. 7. Molecular electrostatic potential maps of compounds **I**, **II** and III.

Table 4
Predicted biological and toxicological properties of compounds **I**, **II** and **III**.

Endpoint (class)	I		II		III	
	Outcome (probability)	AD	Outcome (probability)	AD	Outcome (probability)	AD
Tobacco mosaic (virus)	**Active (92%)**	**reliable**	**Active (84%)**	**reliable**	**Active (96%)**	**reliable**
Bemisia tabaci (insect)	Inactive (76%)	unreliable	Inactive (61%)	unreliable	Inactive (66%)	unreliable
Helicoverpa armigera (caterpillar)	Inactive (98%)	unreliable	Inactive (96%)	unreliable	Inactive (88%)	unreliable
Agrobacterium tumefaciens (bacteria)	Inactive (76%)	unreliable	Inactive (64%)	unreliable	Inactive (82%)	unreliable
Botrytis fabae (fungus)	Active (80%)	reliable	Active (68%)	reliable	Active (68%)	reliable
Aspergillus niger (fungus)	Inactive (80%)	reliable	Inactive (72%)	reliable	Inactive (70%)	reliable
Fusarium oxysporum (fungus)	Active (80%)	reliable	Active (76%)	reliable	Active (80%)	reliable
Fusarium solani (fungus)	**Inactive (61%)**	**unreliable**	**Inactive (53%)**	**unreliable**	**Active (6%)**	**unreliable**
Rhizoctonia solani (fungus)	Active (61%)	reliable	Inactive (64%)	reliable	Inactive (65%)	reliable
Sclerotinia sclerotiorum (fungus)	Active (74%)	reliable	Inactive (57%)	reliable	Active (54%)	reliable
Predictive ecotoxicology against organism models						
Artemia salina (crustacean)	Nontoxic (30%)	unreliable	Nontoxic (16%)	unreliable	Nontoxic (25%)	unreliable
Apis mellifera (honeybee)	Toxic (74%)	reliable	Toxic (84%)	reliable	Toxic (82%)	reliable
Oncorhynchus mykiss (fish)	Nontoxic (28%)	reliable	Nontoxic (15%)	reliable	Nontoxic (29%)	reliable

CRediT authorship contribution statement

Wesley F. Vaz: Conceptualization, Formal analysis, Data curation, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing - original draft, Writing - review & editing. **Giulio D.C. D'Oliveira:** Data curation, Formal analysis, Investigation, Methodology, Writing - original draft. **Caridad N. Perez:** Data curation, Formal analysis, Investigation, Methodology, Writing - original draft. **Bruno J. Neves:** Conceptualization, Formal analysis, Data curation, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing - original draft. **Hamilton B. Napolitano:** Conceptualization, Formal analysis, Data curation, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing - original draft.

Acknowledgments

This paper was supported by Fundação de Amparo à Pesquisa do Estado de Goiás (FAPEG) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). We would like to thank the Universidade Federal de Juiz de Fora for the X-ray diffraction data collection of samples of compounds **II** and **III**. This research was developed with the support of the High-Performance Computation Nucleus of Universidade Estadual de Goiás.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.molstruc.2020.127732.

References

[1] A. Blair, B. Ritz, C. Wesseling, L. Beane Freeman, Pesticides and human health, Occup. Environ. Med. 72 (2015) 81–82, https://doi.org/10.1136/oemed-2014-102454.
[2] I. Mahmood, S.R. Imadi, K. Shazadi, A. Gul, K.R. Hakeem, Effects of pesticides on environment, in: Plant, Soil Microbes, Springer International Publishing, Cham, 2016, pp. 253–269, https://doi.org/10.1007/978-3-319-27455-3_13.
[3] F.P. Carvalho, Pesticides, environment, and food safety, Food Energy Secur. 6 (2017) 48–60, https://doi.org/10.1002/fes3.108.
[4] Y.-C. Lo, S.E. Rensi, W. Torng, R.B. Altman, Machine learning in chemoinformatics and drug discovery, Drug Discov. Today 23 (2018) 1538–1546, https://doi.org/10.1016/j.drudis.2018.05.010.
[5] S. Ekins, A.C. Puhl, K.M. Zorn, T.R. Lane, D.P. Russo, J.J. Klein, A.J. Hickey, A.M. Clark, Exploiting machine learning for end-to-end drug discovery and development, Nat. Mater. 18 (2019) 435–441, https://doi.org/10.1038/s41563-019-0338-z.
[6] M. Abbasi, A. El Hanandeh, Forecasting municipal solid waste generation using artificial intelligence modelling approaches, Waste Manag. 56 (2016) 13–22, https://doi.org/10.1016/J.WASMAN.2016.05.018.
[7] G. Wu, C. Kechavarzi, X. Li, S. Wu, S.J.T. Pollard, H. Sui, F. Coulon, Machine learning models for predicting PAHs bioavailability in compost amended soils, Chem. Eng. J. 223 (2013) 747–754, https://doi.org/10.1016/J.CEJ.2013.02.122.
[8] S. Cipullo, B. Snapir, G. Prpich, P. Campo, F. Coulon, Prediction of bioavailability and toxicity of complex chemical mixtures through machine learning models, Chemosphere 215 (2019) 388–395, https://doi.org/10.1016/J.CHEMOSPHERE.2018.10.056.
[9] S. Kar, K. Roy, J. Leszczynski, On Applications of QSARs in Food and Agricultural Sciences: History and Critical Review of Recent Developments, 2017, pp. 203–302, https://doi.org/10.1007/978-3-319-56850-8_7.
[10] V.S. Brauer, C.P. Rezende, A.M. Pessoni, R.G. De Paula, K.S. Rangappa, S.C. Nayaka, V.K. Gupta, F. Almeida, Antifungal agents in agriculture: friends and foes of public health, Biomolecules 9 (2019) 521, https://doi.org/10.3390/biom9100521.
[11] X. Yu, Y. Wen, C.-G. Liang, J. Liu, Y.-B. Ding, W.-H. Zhang, Design, synthesis and antifungal activity of psoralen derivatives, Molecules 22 (2017) 1672, https://doi.org/10.3390/molecules22101672.
[12] M. Chen, P. Li, D. Hu, S. Zeng, T. Li, L. Jin, W. Xue, B. Song, Synthesis, antiviral activity, 3D-QSAR, and interaction mechanisms study of novel malonate derivatives containing quinazolin-4(3H)-one moiety, Bioorg. Med. Chem. Lett 26 (2016) 168–173, https://doi.org/10.1016/J.BMCL.2015.11.006.
[13] Y.-J. Liu, J.-A. Zhang, L.-J. Zhang, X.-Z. Zou, N. Li, Y. Li, Study on pesticide activities of four ligands and their transition metal complexes with 8-

mercaptoquinoline and pyridine terminal groups, Inorg. Chem. Commun. 57 (2015) 40–43, https://doi.org/10.1016/J.INOCHE.2015.04.016.
[14] X.-Z. Zou, J.-A. Zhang, L.-J. Zhang, Y.-J. Liu, N. Li, Y. Li, S.-C. Wei, M. Pan, Crystal structures and biological activities of a symmetrical quinoline thioether ligand and its transition metal complexes, Inorg. Chem. Commun. 54 (2015) 21–24, https://doi.org/10.1016/J.INOCHE.2015.01.029.
[15] M. Xu, T. Wagerle, J.K. Long, G.P. Lahm, J.D. Barry, R.M. Smith, Insecticidal quinoline and isoquinoline isoxazolines, Bioorg. Med. Chem. Lett 24 (2014) 4026–4030, https://doi.org/10.1016/J.BMCL.2014.06.004.
[16] X.-H. Liu, Y.-M. Fang, F. Xie, R.-R. Zhang, Z.-H. Shen, C.-X. Tan, J.-Q. Weng, T.-M. Xu, H.-Y. Huang, Synthesis and in vivo fungicidal activity of some new quinoline derivatives against rice blast, Pest Manag. Sci. 73 (2017) 1900–1907, https://doi.org/10.1002/ps.4556.
[17] G. Psomas, D.P. Kessissoglou, Quinolones and non-steroidal anti-inflammatory drugs interacting with copper(ii), nickel(ii), cobalt(ii) and zinc(ii): structural features, biological evaluation and perspectives, Dalton Trans. 42 (2013) 6252, https://doi.org/10.1039/c3dt50268f.
[20] C.R. Groom, I.J. Bruno, M.P. Lightfoot, S.C. Ward, The Cambridge structural database, Acta Crystallogr. Sect. B Struct. Sci. Cryst. Eng. Mater. 72 (2016) 171–179, https://doi.org/10.1107/S2052520616003954.
[21] G. d'Oliveira, A. Moura, M. de Moraes, C. Perez, L. Lião, Synthesis, characterization and evaluation of in vitro antitumor activities of novel chalcone-quinolinone hybrid compounds, J. Braz. Chem. Soc. (2018), https://doi.org/10.21577/0103-5053.20180108.
[22] O.V. Dolomanov, L.J. Bourhis, R.J. Gildea, J.A.K. Howard, H. Puschmann, OLEX2: a complete structure solution, refinement and analysis program, J. Appl. Crystallogr. 42 (2009) 339–341, https://doi.org/10.1107/S0021889808042726.
[23] G.M. Sheldrick, A short history of *SHELX*, Acta Crystallogr. Sect. A Found. Crystallogr. 64 (2008) 112–122, https://doi.org/10.1107/S0108767307043930.
[24] G.M. Sheldrick, Crystal structure refinement with SHELXL, Acta Crystallogr. C 71 (2015) 3–8, https://doi.org/10.1107/S2053229614024218.
[25] M. Nardelli, PARST 95 – an update to PARST : a system of Fortran routines for calculating molecular structure parameters from the results of crystal structure analyses, J. Appl. Crystallogr. 28 (1995), https://doi.org/10.1107/S0021889895007138, 659–659.
[26] N. Snejko, C. Cascales, B. Gomez-Lor, E. Gutiérrez-Puebla, M. Iglesias, C. Ruiz-Valero, M.A. Monge, From rational octahedron design to reticulation serendipity. A thermally stable rare earth polymeric disulfonate family with CdI2-like structure, bifunctional catalysis and optical properties, Chem. Commun. (2002) 1366–1367, https://doi.org/10.1039/b202639b.
[27] A.L. Spek, Single-crystal structure validation with the program PLATON, J. Appl. Crystallogr. 36 (2003) 7–13, https://doi.org/10.1107/S0021889802022112.
[28] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G.A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H.P. Hratchian, A.F. Izmaylov, J. Bloino, G. Zheng, J.L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J.A. Montgomery Jr., J.E. Peralta, F. Ogliaro, M. Bearpark, J.J. Heyd, E. Brothers, K.N. Kudin, V.N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J.C. Burant, S.S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J.M. Millam, M. Klene, J.E. Knox, J.B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, R.L. Martin, K. Morokuma, V.G. Zakrzewski, G.A. Voth, P. Salvador, J.J. Dannenberg, S. Dapprich, A.D. Daniels, Ö. Farkas, J.B. Foresman, J.V. Ortiz, J. Cioslowski, D.J. Fox, Gaussian 09, Revision A.02, Gaussian Inc Wallingford CT, Wallingford CT, 2009, p. 34, https://doi.org/10.1159/000348293.
[29] A. Gaulton, L.J. Bellis, A.P. Bento, J. Chambers, M. Davies, A. Hersey, Y. Light, S. McGlinchey, D. Michalovich, B. Al-Lazikani, J.P. Overington, ChEMBL: a large-scale bioactivity database for drug discovery, Nucleic Acids Res. 40 (2012) D1100–D1107, https://doi.org/10.1093/nar/gkr777.
[30] Y. Wang, J. Xiao, T.O. Suzek, J. Zhang, J. Wang, Z. Zhou, L. Han, K. Karapetyan, S. Dracheva, B.A. Shoemaker, E. Bolton, A. Gindulyte, S.H. Bryant, PubChem's BioAssay database, Nucleic Acids Res. 40 (2012) D400–D412, https://doi.org/10.1093/nar/gkr1132.
[31] U.S. Environmental Protection Agency, Pesticide Ecotoxicity Database (Formerly: Environmental Effects Database (EEDB)), 1993.
[32] D. Fourches, E. Muratov, A. Tropsha, Trust, but verify: on the importance of chemical structure curation in cheminformatics and QSAR modeling research, J. Chem. Inf. Model. 50 (2010) 1189–1204, https://doi.org/10.1021/ci100176x.
[33] D. Fourches, E. Muratov, A. Tropsha, Curation of chemogenomics data, Nat. Chem. Biol. 11 (2015), https://doi.org/10.1038/nchembio.1881, 535–535.
[34] D. Fourches, E. Muratov, A. Tropsha, Trust, but verify II: a practical guide to chemogenomics data curation, J. Chem. Inf. Model. 56 (2016) 1243–1252, https://doi.org/10.1021/acs.jcim.6b00129.
[35] B.J. Neves, R.F. Dantas, M.R. Senger, C.C. Melo-Filho, W.C.G. Valente, A.C.M. de Almeida, J.M. Rezende-Neto, E.F.C. Lima, R. Paveley, N. Furnham, E. Muratov, L. Kamentsky, A.E. Carpenter, R.C. Braga, F.P. Silva-Junior, C.H. Andrade, Discovery of new anti-schistosomal hits by integration of QSAR-based virtual screening and high content screening, J. Med. Chem. 59 (2016) 7075–7088, https://doi.org/10.1021/acs.jmedchem.5b02038.
[36] S. Riniker, G. a Landrum, Open-source platform to benchmark fingerprints for ligand-based virtual screening, J. Cheminf. 5 (2013) 26, https://doi.org/10.1186/1758-2946-5-26.
[37] L. Breiman, Random forests, Mach. Learn. 45 (2001) 5–32, https://doi.org/10.1023/A:1010933404324.
[38] S. Zhang, A. Golbraikh, S. Oloff, H. Kohn, A. Tropsha, A novel Automated Lazy Learning QSAR (ALL-QSAR) approach: method development, applications, and virtual screening of chemical databases using validated ALL-QSAR models, J. Chem. Inf. Model. 46 (2006) 1984–1995, https://doi.org/10.1021/ci060132x.
[39] A. Tropsha, Best practices for QSAR model development, validation, and exploitation, Mol. Inform. 29 (2010) 476–488, https://doi.org/10.1002/minf.201000061.
[40] C.A. Moreira, J.M.F. Custódio, W.F. Vaz, G.D.C. D'Oliveira, C. Noda Perez, H.B. Napolitano, A comprehensive study on crystal structure of a novel sulfonamide-dihydroquinolinone through experimental and theoretical approaches, J. Mol. Model. 25 (2019) 205, https://doi.org/10.1007/s00894-019-4091-7.
[41] L.J. Michelini, W.F. Vaz, G.D.C. D'Oliveira, C.N. Pérez, H.B. Napolitano, Analysis of two novel 1–4 quinolinone structures with bromine and nitrobenzyl ligands, J. Mol. Model. 25 (2019) 55, https://doi.org/10.1007/s00894-019-3937-3.
[42] L. Michelini, W. Vaz, L. Naves, C. Pérez, H. Napolitano, Synthesis, characterization and conformational analysis of two novel 4(1H)-Quinolinones, J. Braz. Chem. Soc. (2019), https://doi.org/10.21577/0103-5053.20190124.
[43] S.S. Oliveira, L.G. Santin, L.R. Almeida, L.A. Malaspina, C. Lariucci, J.F. Silva, W.B. Fernandes, G.L.B. Aquino, R. Gargano, A.J. Camargo, H.B. Napolitano, Synthesis, characterization, and computational study of the supramolecular arrangement of a novel cinnamic acid derivative, J. Mol. Model. (2017), https://doi.org/10.1007/s00894-016-3203-x.
[44] J. Peng, K. Song, H. Zhu, W. Kong, F. Liu, T. Shen, Y. He, Fast detection of tobacco mosaic virus infected tobacco using laser-induced breakdown spectroscopy, Sci. Rep. 7 (2017) 44551, https://doi.org/10.1038/srep44551.
[45] T.R. Gordon, Fusarium oxysporum and the Fusarium wilt syndrome, Annu. Rev. Phytopathol. 55 (2017) 23–39, https://doi.org/10.1146/annurev-phyto-080615-095919.
[46] P. Banerjee, A.O. Eckert, A.K. Schrey, R. Preissner, ProTox-II: a webserver for the prediction of toxicity of chemicals, Nucleic Acids Res. 46 (2018) W257–W263, https://doi.org/10.1093/nar/gky318.

Journal of Molecular Structure 1226 (2021) 129326

Contents lists available at ScienceDirect

Journal of Molecular Structure

journal homepage: www.elsevier.com/locate/molstr

In silico-driven identification and structural analysis of nitrodihydroquinolinone pesticide candidates with antifungal activity

Wesley F. Vaz[a,b,*], Bruno J. Neves[c], Jean M.F. Custodio[d], Lucas L. Silva[a], Giulio D.C. D'Oliveira[e], Josiel A. Lemes[f], Bruno F.C. Lacerda[f], Solange X. Santos[a], Caridad N. Perez[e], Hamilton B. Napolitano[a,f,*]

[a] *Campus de Ciências Exatas e Tecnológicas, Universidade Estadual de Goiás, 75132-400, Anápolis, GO, Brazil*
[b] *Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso, 78455-000, Lucas do Rio Verde, MT, Brazil*
[c] *Laboratory for Molecular Modeling and Drug Design, Faculdade de Farmácia, Universidade Federal de Goiás, 74690-900, Goiânia, GO, Brazil*
[d] *Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame, IN, 46556, USA*
[e] *Universidade Federal de Goiás, 74690-900, Goiânia, GO, Brazil*
[f] *Laboratório de Novos Materiais, Centro Universitário de Anápolis, 75083-515, Anápolis, GO, Brazil*

ARTICLE INFO

Article history:
Received 20 July 2020
Revised 29 August 2020
Accepted 22 September 2020
Available online 22 September 2020

Keywords:
Dihydroquinolinones
X-ray diffraction
Predictive modeling
Antifungal activity

ABSTRACT

In this paper, we performed an *in silico*-driven design model to synthesize compounds with biological activity. This rational design has the advantage of decreasing the time and the need for experimental tests and, consequently, the cost related to the search for different candidates. In this way, there is a necessity for more studies that look for new molecules or compounds that may be alternatives to replace the most harmful chemicals for safer options. To contribute to filling this gap, we started an investigation looking for molecules with bioactive potential using a previously developed machine learning model. Leading us to the synthesis, spectroscopic and structural characterization of (*E*)-2-(4-chlorophenyl)-3-(4-nitrobenzylidene)-1-(phenylsulfonyl)-2,3-dihydroquinolin-4(1*H*)-one. Furthermore, considering the predicted biological profile, one of its isomers was incorporated in this study and submitted to experimental validation. The *in vitro* results indicated that the compounds have antifungal activity against *Aspergillus niger* in the same range of positive controls. Moreover, both compounds crystallized in the $P2_1/n$ space group, and their packing is mainly ruled by C–H···O interactions. Lastly, we hope that findings can be used as a starting point for new studies where the structural and biological knowledge of dihydroquinolinones leads to the designing of less toxic or nontoxic analogs antifungal agents by changing undesirable fragments by desirable ones in the molecular skeleton.

1. Introduction

Machine learning is a growing field of artificial intelligence that uses different statistical techniques to enable computers to learn from chemical and biological data types without being explicitly programmed for this task [1]. In this sense, machine learning model techniques have been used to develop quantitative structure-activity relationship models because they (1) do not require bioassay at the first stage, (2) demand less cost and time in estimating potential bioactivities, (3) decrease drug development costs and (4) reduce animal tests [2–4]. These features are essential regarding the average time (10-15 years) and cost (US$ 4-22 billion) to introduce a new pesticide product into the market [5,6].

As an example, a developed quantitative structure–toxicity relationship models provided efficient tools in the discovery and design of low aquatic toxic pesticides, giving new insights about how different physical-chemical properties influence in the high and low toxicity of a compound [7]. Also, these methods have been used to predict acute contact toxicity for bees being suitable to reliably predict the toxicity of structurally diverse pesticides, showing their potential to be applied in the screening and prioritization of new pesticides in the future [8]. Other applications of computational methodology toward pesticide discovery include ligand-based computer-aided pesticide design [9], natural products-based pesticides [10], and multi-target site inhibitor design [11].

Although machine learning-based virtual screening (VS) can be applied in many fields, it has been most used for drug

* Corresponding author.
E-mail addresses: wesfonseca@gmail.com, wesley.vaz@lrv.ifmt.edu.br (W.F. Vaz), hbnapolitano@gmail.com (H.B. Napolitano).

https://doi.org/10.1016/j.molstruc.2020.129326
0022-2860/

discovery. A few examples of the computational study of potential pesticides are found in the literature [12–16], and the design of new molecules with environmental importance is still scarce. Considering the relatively scarce studies on the environmental field, we sought to apply a machine learning-based VS model to propose new compounds with pesticide potential. The *in silico*-driven design of a new quinolinone (*E*)-2-(4-chlorophenyl)-3-(4-nitrobenzylidene)-1-(phenylsulfonyl)-2,3-dihydroquinolin-4(1*H*)-one (P-CNP) was performed using previously developed machine learning models [17]. Guided by these results, a quinolinone derivative was synthesized, crystallized, and structurally characterized. Besides, these outcomes on structural characterization were compared to meta-substituted isomer previously studied by our group [18]. In this sense, electronic properties were theoretically calculated at the B3LYP/6-311+G(d) level, to get a better understand of the molecular structures and supramolecular arrangements of both molecules. Lastly, both compounds were experimentally validated *in vitro* against *Aspergillus niger*, confirming their pesticide potential as new antifungals.

2. Materials and computational procedures

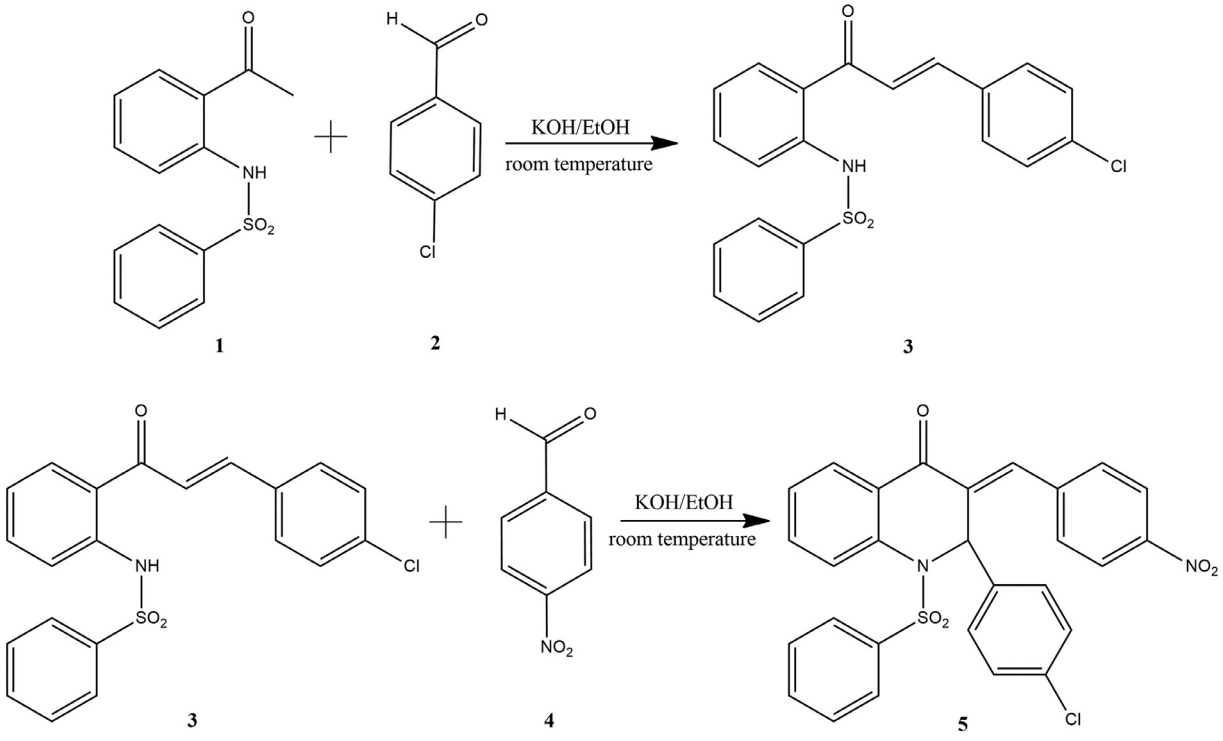
2.1. Synthesis and crystallization

The compound *N*-(2-acetylphenyl)benzenesulfonamide (1) (1.0 mmol) was reacted with 4-chlorobenzaldehyde (2) (3.0 mmol) in basic medium and produced the benzenesulfonamide chalcone (3) [19]. Then, P-CNP (5) was obtained employing Claisen-Schmidt condensation, reacting the benzenesulfonamide chalcone (3) (1.0 mmol) with *p*-nitro-benzaldehyde (4) (2.0 mmol) dissolved in 15 mL of basic ethanol (56.1 mg of potassium hydroxide dissolved) for 48h at 25°C, Scheme 1. So, the solution was filtered, and the precipitate was rinsed with 15 mL of ethanol. The precipitate was dissolved in dichloromethane (10 mL) and this solution was extracted with water. The organic phase evaporates slowly, yielding the product. This methodology is already known, and further details can be obtained in the literature [19,20]. Yellow crystalline solid, yield 81.9%, purity of 98.5%, mp 224 – 226 °C. ^{1}H NMR (500 MHz, $CDCl_3$) δ 6.55 (s, 1H), δ 7.09 – 7.12 (m, 2H), δ 7.22 – 7.26 (m, 2H), δ 7.27 – 7.29 (m, 2H), δ 7.29 – 7.32 (m, 2H), δ 7.33 – 7.36 (m, 3H), δ 7.54 (tt, J 1.25 Hz, 7.50 Hz, 1H), δ 7.59 (ddd, J 1.68 Hz, 7.35 Hz, 8.18 Hz, 1H), δ 7.63 (s, 1H), δ 7.71 (dd, J 1.00 Hz, 8.20 Hz, 1H), δ 7.90 (dd, J 1.60 Hz, 7.75 Hz, 1H), δ 8.27 – 8.29 (m, 2H); ^{13}C NMR (126 MHz, $CDCl_3$) δ 59.4, 124.3, 127.2, 127.7, 127.9, 128.2, 128.4, 128.9, 129.1, 129.4, 130.5, 133.3, 133.6, 134.9, 135.3, 137.3, 138.8, 139.6, 148.3, 181.9; IR 1673 (m), 1602 (m), 1480 (w), 1354 (s), 1298 (m), 1255 (w); HRMS calculated for $C_{28}H_{29}ClN_2O_5S$ 553.0601, found 553.0601. As mentioned before, the M-CNP molecule has already been studied individually and additional structural information about it can be obtained consulting the previous paper [18]. All spectra are available in supplementary materials; Figures S1 to S8.

2.2. X-ray diffraction analysis

Single crystal X-ray diffraction was carried out using Agilent SuperNova diffractometer with MoKα radiation at room temperature (298 K). The collected data were processed with CRYSALISPRO software [21]. P-CNP structure was solved with the ShelXS [22] structure solution program using direct methods and then refined by full-matrix least-squares refinement on F^2 using SHELXL [23] on Olex2 package [24]. All non-hydrogen atoms were refined with anisotropic parameters. All the hydrogen atoms were placed in calculated positions and refined with fixed individual displacement parameters [U_{iso}(H) = 1.2Ueq or 1.5Ueq] according to the riding model. Illustrations were made using Mercury [25] and CrystalExplorer [26]. The crystallographic information file (CIF) of P-CNP was deposited at the Cambridge Crystallographic Data Center (CCDC) [27,28] under code 1981002. Copies of the data can be obtained, free of charge, via www.ccdc.cam.ac.uk.

The physical-chemical properties of title compounds were theoretically computed at the B3LYP/6-311+G(d) level of theory imple-



Scheme 1. General scheme for the synthesis of P-CNP (**5**).

mented in the Gaussian09 package [29] using the crystallographic data from the CIF file. The initial geometric parameters were optimized without constraint, so the minimum energy structure obtained was used in harmonic vibrational calculations. The vibrational modes were assigned using Gaussview [30] animation tool and VEDA4 [31]. The same theory was used to calculate molecular electrostatic potential maps (MEP), frontier molecular orbitals energy (FMO), and electronic reactivity descriptors values.

2.3. Molecular modeling analysis

Machine learning models previously developed by Vaz and coauthors [17] were used to predict biological/toxicological properties of P-CNP and M-CNP against the following organisms: *Tobacco mosaic* (virus), *Bemisia tabaci* (insect), *Helicoverpa armigera* (caterpillar), *Agrobacterium tumefaciens* (bacteria), *Botrytis fabae* (fungus), *Aspergillus niger* (fungus), *Fusarium oxysporum* (fungus), *Fusarium solani* (fungus), *Rhizoctonia solani* (fungus), and *Sclerotinia sclerotiorum* (fungus). Moreover, ecotoxicity of compounds was accessed for *Artemia salina* (crustacean), *Apis mellifera* (honeybee), and *Oncorhynchus mykiss* (fish).

Briefly, the binary machine learning models used here were developed using pairwise combinations of FeatMorgan fingerprints (FCFP-like, radius 2 and bit vector of 2048 bits) calculated in RDKit (http://www.rdkit.org) [32] and Random Forest algorithm implemented in Scikit-learn v.0.19.2 (http://scikit-learn.org/) library of Python v.3.6 (https://www.python.org) [33]. The robustness and predictivity of models were accessed using a 5-fold external cross-validation procedure. All models showed correct classification rate (CCR) values ranging between 0.70 and 0.89, sensitivity (SE) of 0.62-0.85, specificity (SP) of 0.65-0.93, and coverage of 0.55-0.97. The applicability domain (AD) was estimated using the methodology developed by Golbraikh and coauthors [34]. If the compound distance exceeds the similarity threshold between training set compounds and a compound under investigation, the prediction may be considered less trustworthy [34]. Finally, the model predictions were mechanistically interpreted using predicted probability maps developed by Riniker and Landrum [35].

2.4. Antifungal assay

The antifungal activity of nitrodihydroquinolinones isomers was performed from the determination of the minimum inhibitory concentration (MIC) on *A. niger*, according to Espinel-Ingroff and coauthors. [36]. The fungus was reactivated from the collection of fungi culture of the Laboratory of Basic, Applied, and Scientific Dissemination of the State University of Goiás, by sowing in potato dextrose agar (PDA) medium. The inoculum was prepared from a conidial suspension in physiological solution with 0.02% tween 80. The tested compounds were individually added in Eppendorf's with potato-dextrose broth to obtain different concentrations between 2.4 and 5×10^4 μg mL^{-1}. The fungal inoculum was added to the medium to obtain a final concentration of 10^6 conidia mL^{-1}. The cultures were incubated at 25°C for 3 days and fungal growth was verified from the mycelial mass. As a positive control, the antifungal agents' fluconazole and thiophanate-methyl were used in the same concentrations evaluated and as a negative control only the potato-dextrose broth, both added with the fungal inoculum. The MIC was considered the lowest concentration of the tested compounds capable of preventing fungal growth.

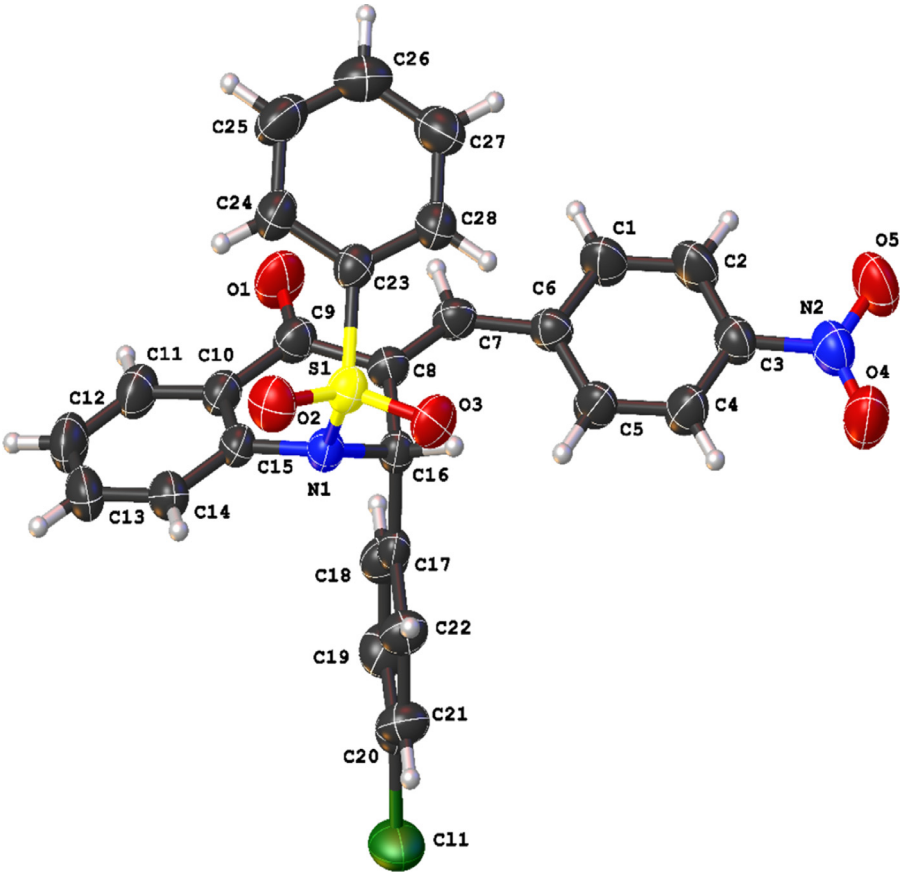


Fig. 1. Thermal ellipsoid drawing of P-CNP showing atom labeling and displacement ellipsoids at the 50% probability level. H atoms are shown as small spheres of arbitrary radii.

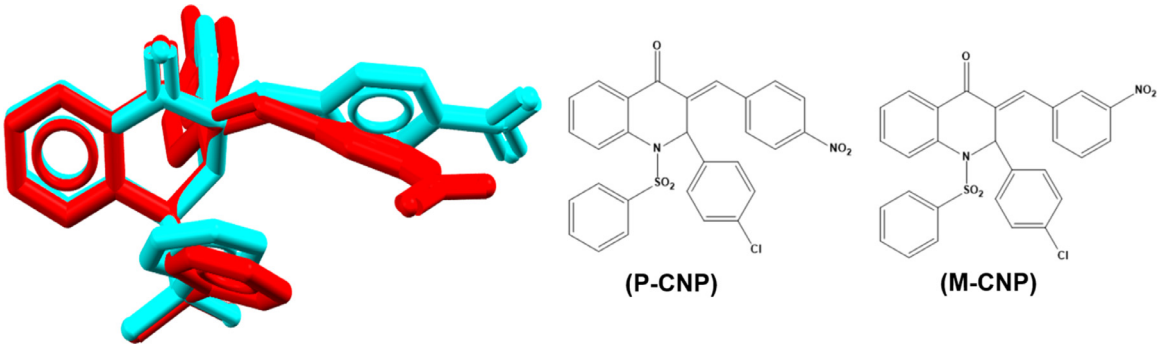


Fig. 2. The overlap of P-CNP (blue) and M-CNP (red) structures obtained from x-ray diffraction data.

Table 1
Crystal data and structure refinement of P-CNP.

	P-CNP
Empirical formula	$C_{28}H_{19}ClN_2O_5S$
Formula weight	530.96
Temperature/K	293(2)
Crystal system	monoclinic
Space group	$P2_1/n$
a/Å	11.1075(4)
b/Å	14.4637(5)
c/Å	15.2820(5)
α/°	90
β/°	96.658(4)
γ/°	90
V(Å^3)	2438.58(16)
Z	4
ρ_{calc}g/cm^3	1.446
μ/mm^{-1}	0.286
F(000)	1096
Crystal size/mm^3	0.53 × 0.41 × 0.34
Radiation	MoKα (λ = 0.71073)
Reflections collected	51829
Independent reflections	6508
Data/restraints/parameters	6508/0/341
Goodness-of-fit on F^2	1.029
Final R_1 index [I>=2σ (I)]	R_1 = 0.0428
Largest diff. peak/hole/e Å^{-3}	0.27/-0.37

Table 2
Hydrogen-bond geometry (Å, °) for P-CNP.

D—H···A	D—H	H···A	D···A	D—H···A
C4—H4···O2^{i}	0.93	2.44	3.355 (2)	168
C5—H5···O3^{i}	0.93	2.53	3.284 (2)	135
C21—H21···O5^{ii}	0.93	2.57	3.166 (2)	122
C22—H22···O5^{ii}	0.93	2.65	3.212 (2)	119
C25—H25···O1^{iii}	0.93	2.66	3.559 (2)	164
C27—H27···O4^{iv}	0.93	2.67	3.538 (3)	155
C28—H28···O5^{iv}	0.93	2.59	3.330 (2)	137

Symmetry codes: (i) −x, −y+1, −z+2; (ii) x−1, y, z; (iii) −x, −y+1, −z+1; (iv) −x+1, −y+1, −z+2.

3. Results and discussion

3.1. X-ray diffraction analysis

The compound P-CNP consists of a benzenesulfonyl group attached to N1 atom, a *p*-nitro vinylbenzene group attached to the C8 atom, and a chlorobenzene group attached to C16 atom. The piperidone cycle appears in the chair form. Thus O1 and N1 atoms are not coplanar with the ring *a*. The crystallographic parameters and ORTEP representation were presented in Table 1 and Fig. 1, respectively.

The *S* configuration was used in the structural analysis of P-CNP. So, the nitrobenzene rings were not planarly oriented concerning the dihydroquinolin-4(1*H*)-one motif ring. In P-CNP there is a deviation of 26.43(6)°, while in M-CNP [18] it is 64.58(6)°. Moreover, the values for torsion angles show that the NO_2 group has no coplanarity to the ring *b* in both isomers. That torsion [C2-C3-N2-O4 = = -176.0(2)] is 4(2)° in P-CNP and more pronounced in M-CNP [C1-C2-N2-O5 = -158.2(2)] where the deviation is approximately 22 (2) °. Fig. 2 shows the overlay of both structures.

P-CNP features the absence of a strong H-bond donor, just like M-CNP, being it packing stabilized by C–H···O H-bonds being three centrosymmetric dimers and one bifurcated interaction. Also, it is noted that the change in the position of the NO_2 group (form meta to para) led to an increase in the number of C–H···O interactions at P-CNP when compared to M-CNP [18] . In the first interaction, it is observed that the interaction between O atoms of nitrobenzene group and H atoms of benzenesulfonyl group generating a dimer with $R_2^2(7)$ motif (Fig. 3a). Also, two intermolecular interactions, involving H atoms of nitrobenzene group and O atoms of the SO_2 group generates a second dimer with $R_2^2(7)$ motif (Fig 3b). One of these interactions is formed between the carbonyl group and H atoms of benzenesulfonyl group with $R_2^2(20)$ motif (Fig. 3c), while the other involves a bifurcated interaction where two H atoms of chlorobenzene ring interact with one O atom of nitrobenzene ring $R_2^1(5)$ (Fig. 3d). Finally, the molecular packing of P-CNP is represented in Fig. 3e, while the molecular packing of M-CNP is represented in Fig. 3f and the hydrogen-bond geometries are given in Table 2.

To further analysis about intermolecular interactions, present in P-CNP, Hirshfeld Surface (HS) analysis was used in this study. The d_{norm} HS of the title compound is illustrated in Fig. 4, and is represented both, acceptor and donor regions, of six interactions. Blue indicates outlying contacts while red indicates closer contacts. The interaction C25–H25···O1, represented by (I), is related to a dimeric interaction formed by carbonyl and benzenesulfonyl groups. The interactions C4–H4···O2 (II) and C5–H5···O3 (III) are related by contacts between SO_2 and nitrobenzene groups. Additionally, C27–H27···O4 (IV) and C28–H28···O5 (V) interactions involve nitrobenzene and benzenesulfonyl groups, respectively. Finally, the bifurcated interaction involving the nitro group and chlorobenzene ring is represented by (VI) and is formed by C21–H21···O5 and C22–H22···O5 contacts. Based on such contact proximities , highlighted in red on the surface, the interaction (II) is the closest in the packing of P-CNP. The closest contacts suggest that the nitrobenzene group has a leading role in the supramolecular arrangement.

In addition to the C–H···O interactions, there are three C–H···π interactions that contribute to the crystalline arrangement of P-CNP. These hydrophobic interactions connect the dimers and are responsible for the molecular growth in a three-dimensional way. The shape index surface is a tool to analyze hydrophobic contacts; moreover, on this surface, C–H···π interac-

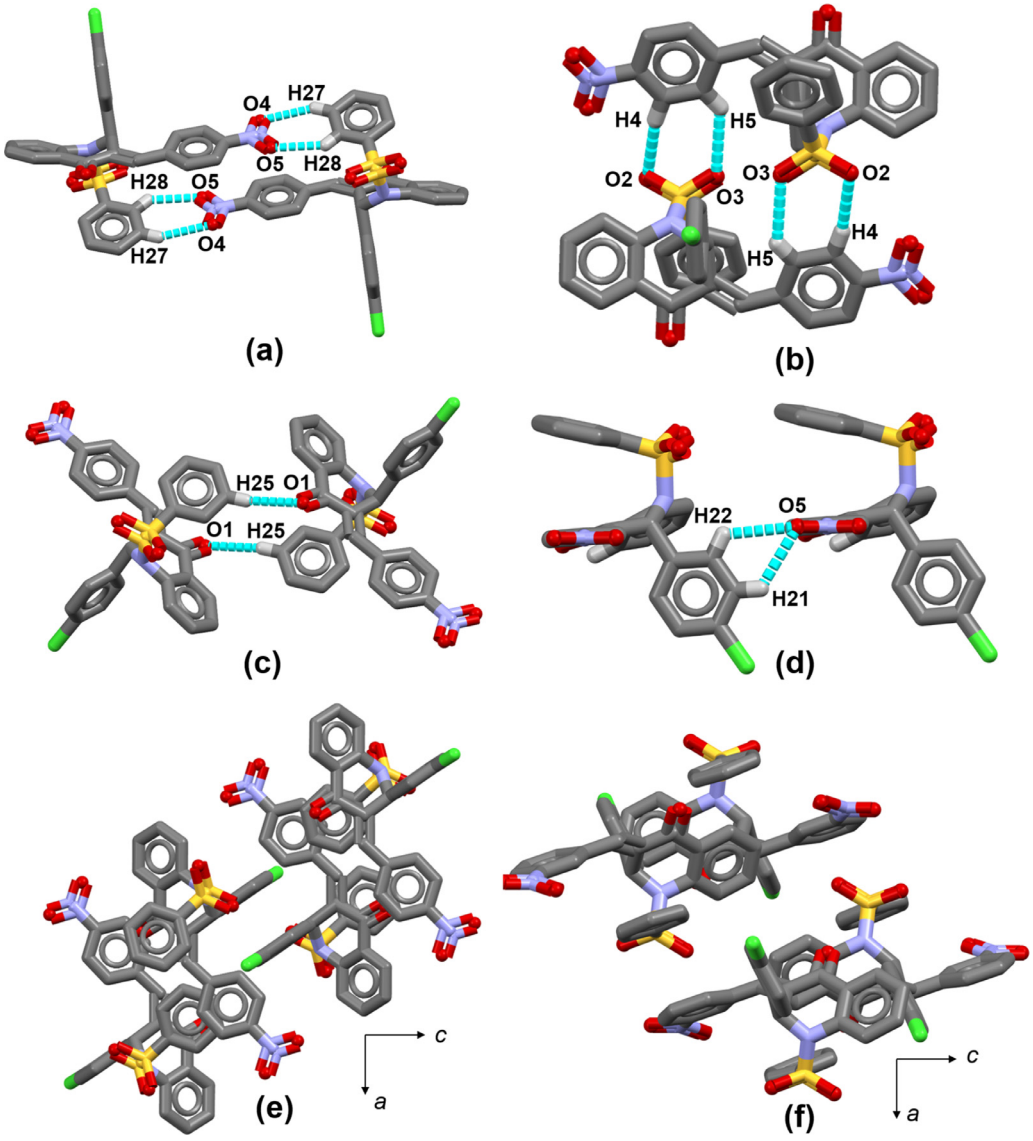


Fig. 3. Dimers observed in the crystal packing of P-CNP with motifs $R_2^2(7)$ (a) and (b), $R_2^2(20)$ (c), and bifurcated interaction $R_2^1(5)$ (d). The molecular packing of P-CNP (e) and M-CNP (f) are shown as viewed parallel to the b axis.

tions are recognized by a broad depression in the surface above the aromatic ring. Red is used to represent the acceptor region while blue represents the donor region of intermolecular contacts. Fig. 5 show the shape index surface for P-CNP, where the region marked with a circle represents C–H$\cdots\pi$ interactions. The first involves C12, H12 atoms, and benzenesulfonyl ring [H$\cdots$Cg$_{(D)}$ = 2.94 Å]. The second involves C1, H1 atoms, and benzenesulfonyl ring [H$\cdots$Cg$_{(C)}$ = 3.65 Å], and the last one involves C26, H26 atoms, and chlorobenzene ring [H$\cdots$Cg$_{(B)}$ = 3.64 Å].

The 2D fingerprint plot (d_e vs d_i) tool of HS was also used to study all the interatomic interactions in the isomers (Fig. 6). In both molecules, H$\cdots$H interactions (I) are the higher presence in the fingerprints, being 34% and 30.7%, in P-CNP and M-CNP, respectively. About C$\cdots$C interactions (II), they represent, respectively, 2% and 4.1% in P-CNP and M-CNP packing. Also, C$\cdots$H interactions (III) are 16.7% and 18.6% in the fingerprint of P-CNP and M-CNP, respectively. The H$\cdots$O interactions (IV) correspond to 25.2% and 27.3% of the interactions present in P-CNP and M-CNP, respectively. The occurrence of H$\cdots$Cl interactions (V) is almost the same in the isomers, being 10.1% for P-CNP and 10.4% for M-CNP. Lastly, the O$\cdots$C interactions (VI) are more frequent in P-CNP than in M-CNP being 6.6 % and 2.5%, respectively. Thus, it is possible to notice that the NO_2 group settled at para position leads to an increase in the occurrence only of H$\cdots$H and O$\cdots$C interactions indicating that in this position, the molecule experiences more steric effects than when it is in the meta position. On the other hand, the change in the position of the NO_2 group does not significantly affect the occurrence of H$\cdots$Cl interactions.

The theoretical and experimental vibrational spectra were plotted in Fig. 7. A scaling factor [37] equal to 0.968 was applied in theoretical wavenumbers to correct overestimated values obtained with DFT [38]. The IR intensities, the calculated vibrational frequencies, and the assigned vibrational modes for the main group of P-CNP and M-CNP are listed in Table 3.

The stretching vibrations for aromatic CH bonds usually arise in the range [39] of 3050 to 3010 cm^{-1}. This mode for P-CNP and M-CNP occurs at 3091 to 3057 cm^{-1} and 3065 to 3050 cm^{-1}, respectively. Regarding the calculated frequencies, this mode occurs practically in the same regions, at the range of 3127 to 3070 cm^{-1}. The vibrational stretching mode for the C=C group occurs as a weak band around 1680 to 1600 cm^{-1}. Meanwhile, the absorption

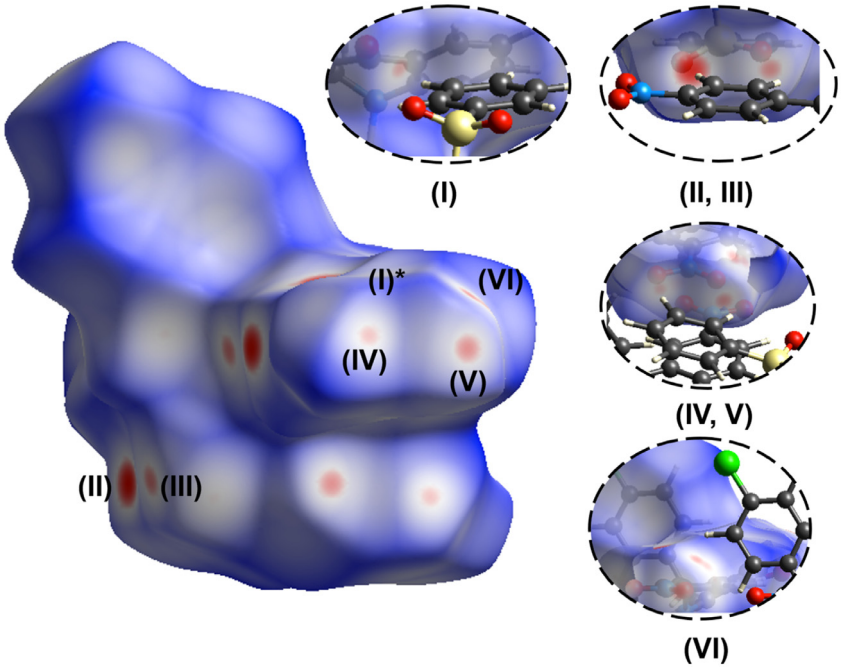


Fig. 4. The Hirshfeld surface d_{norm} map showing C–H···O interactions observed in the molecular packing of P-CNP.

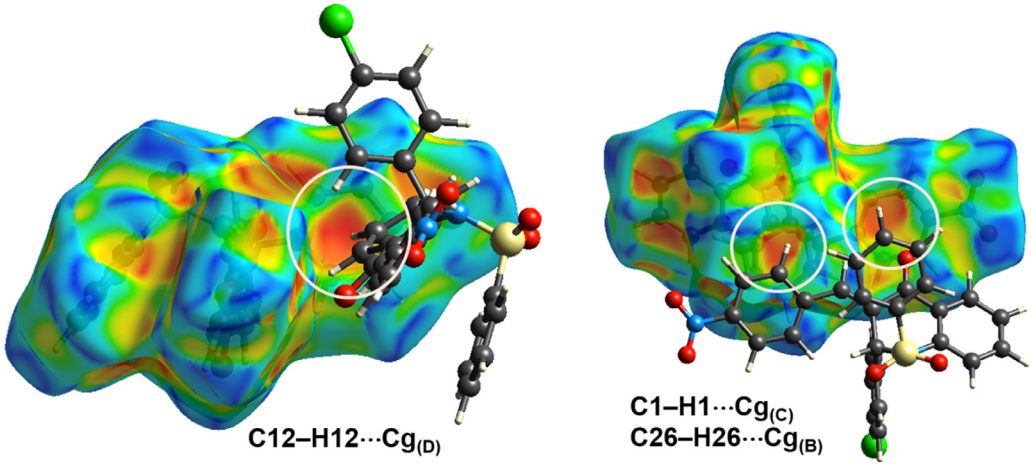


Fig. 5. Shape index surface showing the C–H···π interactions establishing the crystal packing of P-CNP. In the image on the right, the left circle represents C1−H1···$Cg_{(C)}$

Table 3
Theoretical and experimental vibrational assignments of P-CNP and M-CNP.

	P-CNP				M-CNP			
Vibrational mode	Unscaled[a]	Scaled[a]	Intensity[b]	IR[c]	Unscaled[a]	Scaled[a]	Intensity[b]	IR[c]
ν $(CH)_{Ar}$	3231.19 – 3172.28	3127.19-3070.28	-	3091 - 3057	3231.18 – 3172.76	3127.60-3070.49	-	3053 -3065
ν $(C\text{-}C)_{Ar}$	1650.37 – 1612.15	1597.37-1560.15	-	1573 -1476	1659.13 – 1612.03	1605.91– 1560.41	-	1562-1469
ν C=C	1650.37	1597.56	186.04	1489	1650.44	1597.62	222.00	1575
ν C=O	1725.75	1670.53	126.80	1673	1727.38	1672.10	131.82	1696
ν_{asym} SO_2	1304.45	1262.71	24.60	1309	1304.04	1262.31	35.94	1359
ν_{sym} SO_2	1125.59	1089.57	135.60	1143	1124.82	1088.82	143.71	1050
ν_{asym} NO_2	1574.90	1524.50	198.20	1529	1587.54	1536.73	229.44	1524
ν_{sym} NO_2	1367.08	1323.33	445.23	1354	1371.66	1327.76	340.04	1316
ν C_{Ar}–Cl	1103.00	1067.70	68.35	998	1103.10	1067.70	65.71	1012

[a] Scale factor = 0.968.
[b] IR intensities (K $mmol^{-1}$).
[c] cm^{-1}.

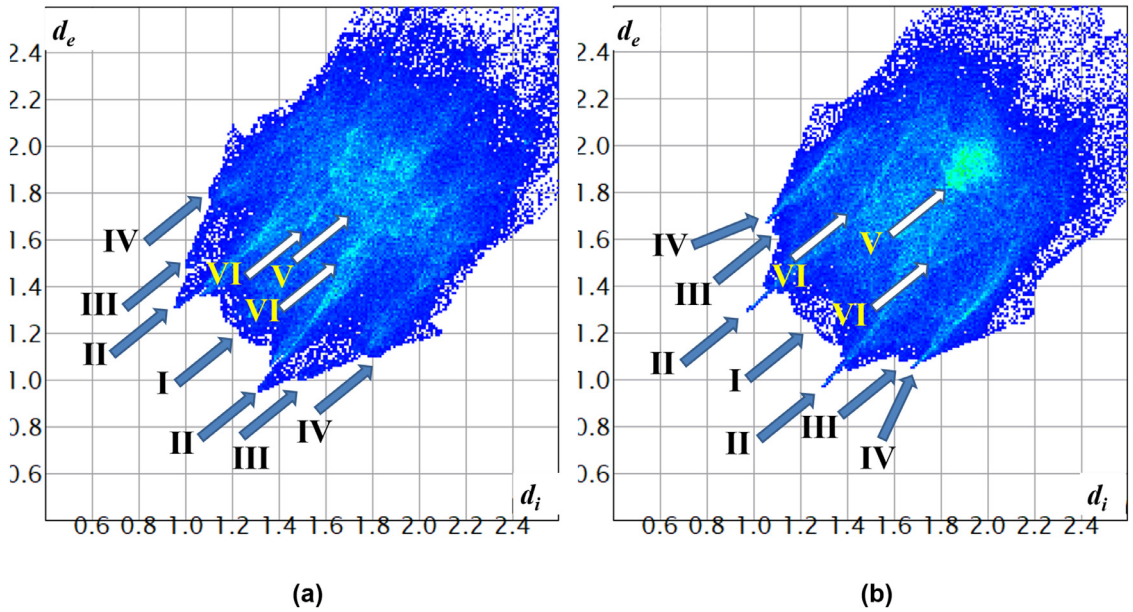


Fig. 6. 2D fingerprint plots of the nearest external distance (d*e*) versus the nearest internal distance (d*i*) for P-CNP (a) and M-CNP (b). I represent the regions corresponding to H···H interactions; II, the regions of C···C interactions; III, the regions of C···H interactions; IV, the regions of H···O interactions; V, the regions of H···Cl, and VI, the regions of O···C interactions. The colors represent the number of points that share the same *di, de* coordinate (light blue: many; dark blue: few).

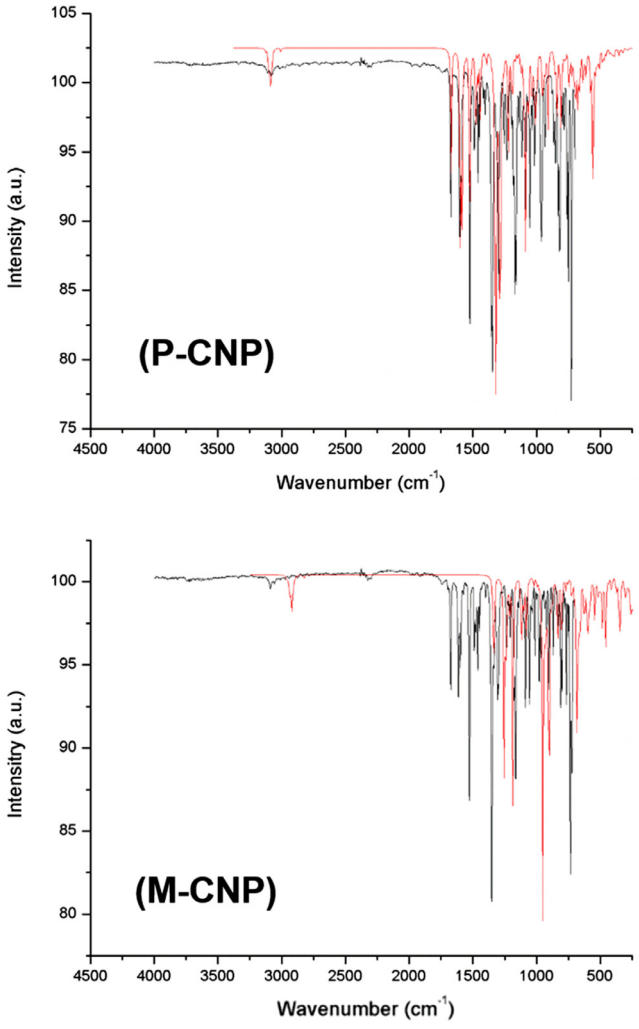


Fig. 7. The experimental (black) and theoretical (red) infrared spectra overlay of P-CNP and M-CNP.

of medium to the strong intensity from 1600 to 1450 is associated with this vibrational mode in aromatic rings [39]. νC=C mode was used for both P-CNP and M-CNP, and in both cases, the band arises at 1597 cm^{-1}, while in FT-IR, these bands appear at 1489 and 1575 cm^{-1}, respectively. Additionally, this stretching mode occurring in aromatic rings were theoretically assigned in a region from 1612 to 1650 cm^{-1} for both compounds. The values are in good agreement with the experimental results.

The carbonyl group, when conjugated with a double carbon bond, tends to show its stretch mode in the range of 1675 to 1680 cm^{-1} [39]. In P-CNP, the strong bands observed at 1673 cm^{-1} in FT-IR, and 1670 cm^{-1} observed in theoretical calculations are attributed to νC=O mode. On the other hand, M-CNP presents this mode at 1696 cm^{-1} (exp) and 1672 cm^{-1}(DFT). For the sulfone group, two stretching bands occur – one asymmetrical in the region of 1300 cm^{-1} and another, symmetrical in the region [39] of 1150 cm^{-1}. In the present study, the bands observed at 1309 cm^{-1}, 1359 cm^{-1} (experimental) and 1262 cm^{-1} (calculated) are assigned to SO_2 asymmetric mode for P-CNP and M-CNP, respectively, while symmetric vibrations occur at 1143 cm^{-1}, 1150 cm^{-1} (experimental) and 1089 cm^{-1} (calculated), respectively.

A nitro group, when conjugated to an aromatic ring, displaces its stretch bands around 1550 to 1490 cm^{-1} (asymmetric) and 1355 to 1315 cm^{-1} (symmetric) [39]. The asymmetric mode appears in experimental spectra at 1529 and 1524 cm^{-1}, while in calculated spectra, it occurs at 1524 and 1536 cm^{-1}, for P-CNP and M-CNP, respectively. Besides, for symmetric mode, experimental values can be found at 1354 and 1316 cm^{-1} and 1323 and 1327 in calculated spectra, for P-CNP and M-CNP, respectively. The νC–Cl vibration in aryl chlorides absorb between [39] 1100 and 1035 cm^{-1}. Here, experimental bands arise at 998 and 1012 cm^{-1}, and calculated bands occur at 1067 cm^{-1} for P-CNP and M-CNP, respectively.

The global electronic reactivity descriptors were obtained with FMO energies to calculate: ionization energy [$I = -E_{HOMO}$], electron affinity [$A = -E_{LUMO}$], electronic chemical potential [$\mu = 1/2(E_{HOMO} + E_{LUMO})$], hardness [$\eta = 1/2(E_{LUMO} - E_{HOMO})$], softness [$s = 1/\eta$], gap energy [$E_{gap} = E_{LUMO} - E_{HOMO}$] and electrophilicity [$\omega = \mu^2/2\eta$] [40]. These values for P-CNP and M-CNP were listed

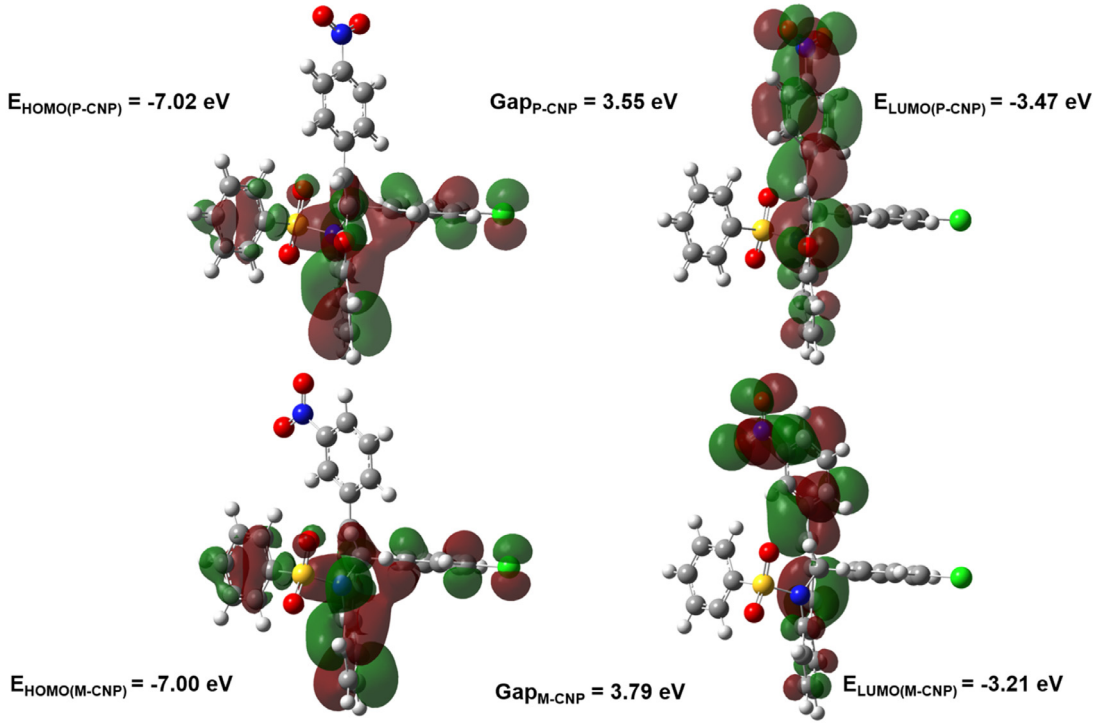


Fig. 8. Frontier molecular orbitals of P-CNP and M-CNP and their energies

Table 4
The global reactivity descriptors of P-CNP and M-CNP

Molecular properties	P-CNP	M-CNP
Ionization potential (I)	7.02	7.00
Electron affinity (A)	3.47	3.21
Chemical potential (μ)	-5.25	-5.11
Global hardness (η)	1.78	1.90
Softness (s)	0.56	0.53
Energy gap (eV)	3.55	3.79
Electrophilicity index (ω)	7.75	6.88

in Table 4, and the graphical representation of FMO is shown in Fig. 8.

The electron donor behavior in both compounds is similar, while P-CNP (lowest LUMO energy) can be the best electron acceptor (Table 4). The values of potential ionization and electron affinity, table 4, lead to the same conclusions. Also, P-CNP, when compared to M-CNP, is the softest molecule (lowest ΔE_{gap}), the most electronic reactive compound (less hardness, greater softness), and has low electronic kinetic stability. Finally, by electrophilicity values, P-CNP can be considered a stronger electrophile. To investigate the reactive sites for the electrophilic and nucleophilic attack, the MEP surfaces were calculated for the title compounds and are shown in Fig. 9. The negative potential, red regions, are mainly localized over the oxygen atoms, SO_2, and NO_2 groups, and these regions are most susceptible to electrophilic attack. The positive potential, blue regions, are localized around hydrogen atoms of the aromatic rings, indicating possible sites for nucleophilic attack. To conclude, the green regions, around chlorine atoms, represent the zero potential of the title molecule.

3.2. Molecular Modeling analysis

The developed machine learning models were then used as filters in a virtual screening to prioritize the putative pesticide applications of M-CNP and P-CNP. According to Table 5, both compounds were predicted as antifungal agents against *A. niger* (76% and 75% of probability for M-CNP and P-CNP, respectively). The *A. niger* is ubiquitous in soil and commonly reported in dead leaves, stored grain, compost piles, and other decaying vegetation. It causes a disease called black mold on certain fruits and vegetables such as grapes, onions, peanuts, and is a common contaminant of food [41–43]. Both compounds were predicted as active against *S. sclerotiorum* (70% and 71% of probability for M-CNP and P-CNP, respectively), a fungus, that causes a severe disease called white mold in a broad host range of more than 400 plant species (herbaceous, succulent plants, particularly flowers and vegetables), thereby becoming one of the most successful plant pathogens. On the other hand, predictive analysis of ecotoxicity potential indicates that M-CNP and P-CNP may be toxic to honeybees (*A. mellifera)*, see Table 5).

Subsequently, predicted probability maps (Fig. 10) from machine learning predictions were generated for both compounds aiming to visualize the atomic and fragment contributions for antifungal activity and acute contact toxicity for the honeybee. Here, the "weight" of an atom was considered as a predicted-probability difference obtained when the bits in the fingerprint corresponding to the atom are removed. Then, the normalized weights were used to color the atoms in topography-like maps in which green (active or nontoxic) and purple regions (inactive or toxic) indicate a decrease or increase in probabilities when the bits are removed. By analyzing the probability maps generated for the anti-*A. niger* activity, we identified that all atoms, except the nitro group, contribute favorably for the antifungal activity. Furthermore, for anti-*S. sclerotiorum*, nitro group, and acetophenone moiety also contributed to the antifungal property. On the other hand, mechanistic interpretation of potential acute contact toxicity for honeybee suggests that the propan-2-amine fragment (highlighted in purple) contributed positively to the toxicity. This structural and biological information provided by the QSAR models could be useful for designing or optimizing potent and nontoxic antifungal analogs by replacing unfavorable fragments with favorable fragments.

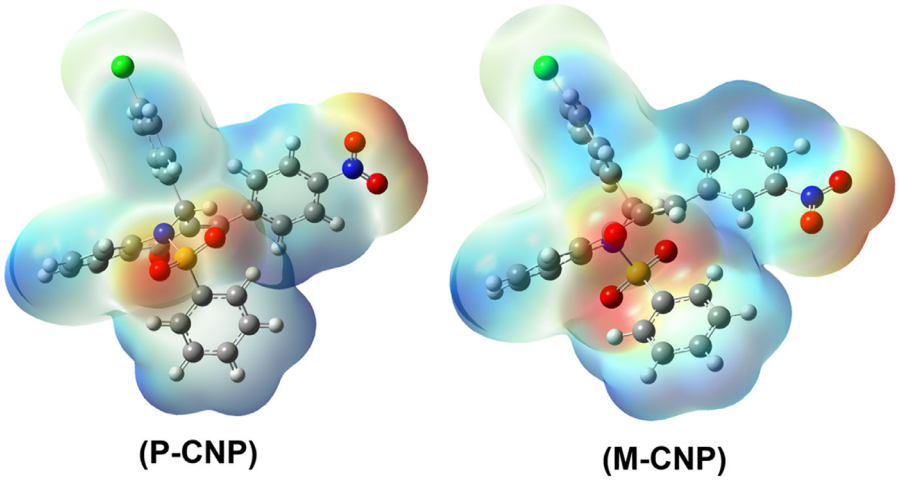


Fig. 9. Molecular electrostatic potential map of P-CNP and M-CNP. Blue represents areas with a positive charge, while red represents regions with a negative charge. Lastly, green represents the zero potential regions.

Table 5
Predicted biological and toxicological properties of synthesized compounds.

Endpoint (class)	M-CNP		P-CNP	
	Outcome (probability)	AD	Outcome (probability)	AD
Predictive pesticide activity against several organisms				
Tobacco mosaic (virus)	Inactive (8%)	reliable	Inactive (8%)	reliable
Bemisia tabaci (insect)	Active (61%)	unreliable	Active (63%)	unreliable
Helicoverpa armigera (caterpillar)	Inactive (86%)	unreliable	Inactive (85%)	unreliable
Agrobacterium tumefaciens (bacteria)	Inactive (76%)	unreliable	Inactive (76%)	unreliable
Botrytis fabae (fungus)	Inactive (67%)	reliable	Inactive (67%)	reliable
***Aspergillus niger* (fungus)**	**Active (76%)**	**reliable**	**Active (75%)**	**reliable**
Fusarium oxysporum (fungus)	Inactive (83%)	reliable	Inactive (83%)	reliable
Fusarium solani (fungus)	Inactive (59%)	unreliable	Inactive (59%)	unreliable
Rhizoctonia solani (fungus)	Inactive (71%)	reliable	Inactive (71%)	reliable
***Sclerotinia sclerotiorum* (fungus)**	**Active (70%)**	**reliable**	**Active (71%)**	**reliable**
Predictive ecotoxicology against organism models				
Artemia salina (crustacean)	Nontoxic (62%)	unreliable	Nontoxic (59%)	unreliable
Apis mellifera (honeybee)	Toxic (81%)	reliable	Toxic (81%)	reliable
Oncorhynchus mykiss (fish)	Nontoxic (82%)	reliable	Nontoxic (82%)	reliable

AD: applicability domain

Table 6
Effect of different concentrations of M-CNP and P-CNP on the mycelial growth of *Aspergillus niger* on potato-dextrose broth (PDB). The first appearance of the (-) symbol in each line represents the MIC value in which the compound shows activity.

	Minimum inhibitory concentration (mgL^{-1})								
	0*	24×10^{-3}	0.05	0.8	3	6	12	25	50
M-CNP	+	+	+	+	+	+	-	-	-
P-CNP	+	+	+	+	+	+	+	-	-
Positive controls									
Fluconazole	+	+	+	+	+	+	-	-	-
Thiophanate-methyl	+	+	+	+	+	+	+	+	-

* Negative control (only PDB and the fungal inoculum); - Absence of fungal growth; + Fungal growth.

3.3. Antifungal activity

Considering the *in-silico* profile of M-CNP and P-CNP, both evaluated compounds showed antifungal activity against *A. niger*. As we can see in Table 6, M-CNP and P-CNP showed MICs of 12 and 25 mgL^{-1}, respectively, therefore lower than thiophanate-methyl (50 mgL^{-1}). P-CNP showed slightly higher and M-CNP similar MIC to fluconazole (12 mgL^{-1}). These results open perspectives for future studies aiming to apply these molecules in drugs and/or agrochemicals with antifungal action. Also, future studies can be carried out to show if the use of these molecules can reduce a produce toxic environmental waste such as those generated by a standard antifungal drug like fluconazole and thiophanate-methyl. The properties of some natural and/or synthetic chalcones and their analogs against different organisms included fungi that have been previously described [44,45] and in some cases, the use of these molecules is more effective than commercial fungicides [46]. *A. niger* is also a fungus of great interest for microbiology, because its spores are present in abundance in the environment, mainly in the air, besides presenting a fast growth and tolerance to variations of

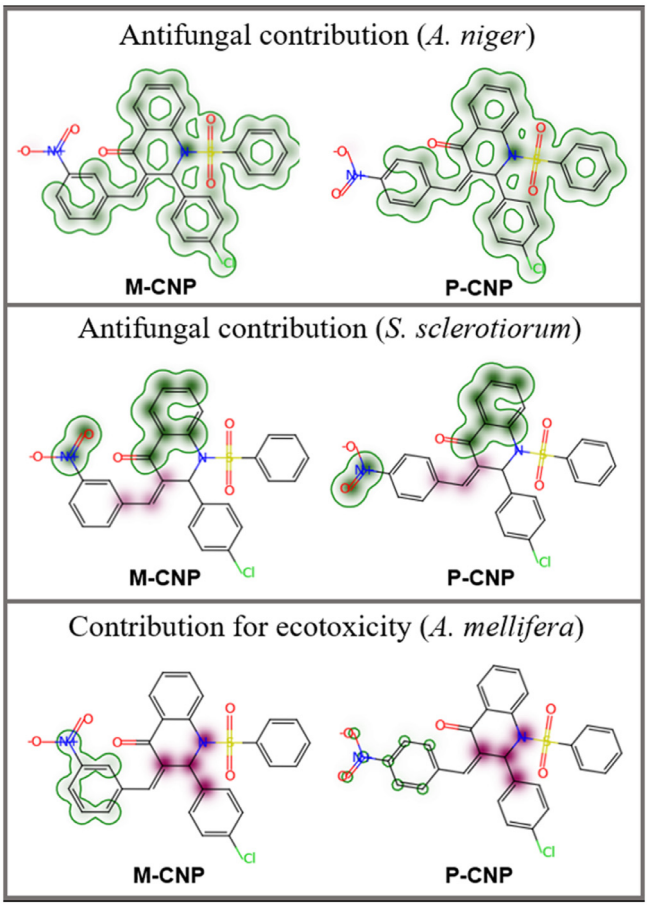


Fig. 10. Predicted influence of structural atoms and fragments on the antifungal activity and acute contact toxicity for the honeybee. Fragments increasing the toxicity are shown in green and fragments decreasing the toxicity are shown in purple.

pH, being able to contaminate a multitude of products [47,48]. For this reason, it is commonly used in laboratory tests to analyze the antifungal activity of different compounds, being part of the procedure of quality control of these products. Thus, the results obtained point to the potential of the tested compounds as microbial control agents.

4. Conclusions

P-CNP and M-CNP are not planar compounds, and with the relation of dihydroquinolin-4(1H)-one moiety the nitrobenzene ring is more twisted indicating that the molecule in para position is more rigid. Their molecular packing is stabilized by H-bonds, of the C–H···O kind, and they happen in greater number in P-CNP. Then, these interactions form dimers and bifurcated interactions being this pattern recurrent in the literature. The topological analysis showed that just H···H and O···C are more frequently in P-CNP than in M-CNP, being that a consequence of the changing of the NO_2 group position. Also, C–H···π interactions contribute to the crystalline state of the isomers. Moreover, in terms of electronic properties, both compounds have similar behavior, being P-CNP lightly more electronic kinetic stable. Finally, categorical machine learning models previously developed by our group were used to access the pesticide profile of P-CNP and M-CNP. Based on the machine learning model prediction on the bioactivity of nitrodihydroquinolinones, further testing against *A. niger* was undertaken. Consequently, we were able to identify the antifungal action of these two new chemical scaffolds.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Wesley F. Vaz: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Writing - original draft, Writing - review & editing. **Bruno J. Neves:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Writing - original draft, Writing - review & editing. **Jean M.F. Custodio:** Data curation, Formal analysis, Investigation, Methodology, Writing - original draft, Writing - review & editing. **Lucas L. Silva:** Data curation, Formal analysis, Investigation, Methodology, Writing - original draft, Writing - review & editing. **Giulio D.C. D'Oliveira:** Data curation, Formal analysis, Investigation, Methodology, Writing - original draft, Writing - review & editing. **Josiel A. Lemes:** Data curation, Formal analysis, Investigation, Methodology, Writing - original draft, Writing - review & editing. **Bruno F.C. Lacerda:** Data curation, Formal analysis, Investigation, Methodology, Writing - original draft, Writing - review & editing. **Solange X. Santos:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Writing - original draft, Writing - review & editing. **Caridad N. Perez:** Data curation, Formal analysis, Investigation, Methodology, Writing - original draft, Writing - review & editing. **Hamilton B. Napolitano:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Writing - original draft, Writing - review & editing.

Acknowledgments

The authors would like to acknowledge the Brazilian agencies' FAPEG, CNPq, and CAPES for financial and fellowship supports. We would like to thank Universidade Federal de Juiz de Fora for X-ray diffraction data collection and ChemAxon Inc. (http://www.chemaxon.com) for providing the academic license of their software. This research was developed with the support of the High-Performance Computing Center of the Universidade Estadual de Goiás.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.molstruc.2020.129326.

References

[1] S. Ekins, A.C. Puhl, K.M. Zorn, T.R. Lane, D.P. Russo, J.J. Klein, A.J. Hickey, A.M. Clark, Exploiting machine learning for end-to-end drug discovery and development, Nat. Mater. 18 (2019) 435–441, doi:10.1038/s41563-019-0338-z.
[2] W.M. Berhanu, G.G. Pillai, A.A. Oliferenko, A.R. Katritzky, Quantitative Structure-Activity/Property Relationships: The Ubiquitous Links between Cause and Effect, Chempluschem 77 (2012) 507–517, doi:10.1002/cplu.201200038.
[3] N. Bosc, F. Atkinson, E. Felix, A. Gaulton, A. Hersey, A.R. Leach, Large scale comparison of QSAR and conformal prediction methods and their applications in drug discovery, J. Cheminform 11 (2019) 4, doi:10.1186/s13321-018-0325-4.
[4] B.J. Neves, R.C. Braga, C.C. Melo-Filho, J.T. Moreira-Filho, E.N. Muratov, C.H. Andrade, QSAR-Based Virtual Screening: Advances and Applications in Drug Discovery, Front. Pharmacol. 9 (2018) 1275, doi:10.3389/fphar.2018.01275.
[5] V. Storck, D.G. Karpouzas, F. Martin-Laurent, Towards a better pesticide policy for the European Union, Sci. Total Environ. 575 (2017) 1027–1033, doi:10.1016/j.scitotenv.2016.09.167.
[6] D. Bourguet, T. Guillemaud, The Hidden and External Costs of Pesticide Use, in: L. E (Ed.), Sustain. Agric. Rev., Springer, 2016, pp. 35–120, doi:10.1007/978-3-319-26777-7_2.
[7] L. He, K. Xiao, C. Zhou, G. Li, H. Yang, Z. Li, J. Cheng, Insights into pesticide toxicity against aquatic organism: QSTR models on Daphnia Magna, Ecotoxicol. Environ. Saf. 173 (2019) 285–292, doi:10.1016/j.ecoenv.2019.02.014.

[8] F. Como, E. Carnesecchi, S. Volani, J.L. Dorne, J. Richardson, A. Bassan, M. Pavan, E. Benfenati, Predicting acute contact toxicity of pesticides in honeybees (Apis mellifera) through a k-nearest neighbor model, Chemosphere 166 (2017) 438–444, doi:10.1016/j.chemosphere.2016.09.092.
[9] B. Bordás, T. Kőmíves, A. Lopata, Ligand-based computer-aided pesticide design. A review of applications of the CoMFA and CoMSIA methodologies, Pest Manag. Sci. 59 (2003) 393–400, doi:10.1002/ps.614.
[10] R. Yang, T. Xu, J. Fan, Q. Zhang, M. Ding, M. Huang, L. Deng, Y. Lu, Y. Guo, Natural products-based pesticides: Design, synthesis and pesticidal activities of novel fraxinellone derivatives containing N -phenylpyrazole moiety, Ind. Crops Prod. 117 (2018) 50–57, doi:10.1016/j.indcrop.2018.02.088.
[11] J. Gressel, Perspective: present pesticide discovery paradigms promote the evolution of resistance – learn from nature and prioritize multi-target site inhibitor design, Pest Manag. Sci. 76 (2020) 421–425, doi:10.1002/ps.5649.
[12] C. Braeuning, A. Braeuning, H. Mielke, A. Holzwarth, M. Peiser, Evaluation and improvement of QSAR predictions of skin sensitization for pesticides, SAR QSAR Environ. Res. 29 (2018) 823–846, doi:10.1080/1062936X.2018.1518261.
[13] B. Souyei, A. Hadj Seyd, F. Zaiz, A. Rebiai, Application of Inverse QSAR/QSPR Analysis for Pesticides Structures Generation, Acta Chim. Slov. 66 (2019) 315–325, doi:10.17344/acsi.2018.4793.
[14] L.-T. Qin, Y.-H. Chen, X. Zhang, L.-Y. Mo, H.-H. Zeng, Y.-P. Liang, QSAR prediction of additive and non-additive mixture toxicities of antibiotics and pesticide, Chemosphere 198 (2018) 122–129, doi:10.1016/j.chemosphere.2018.01.142.
[15] X. Wang, M. Wang, J. Yan, M. Chen, A. Wang, Y. Mei, W. Si, C. Yang, Design, Synthesis and 3D-QSAR of New Quinazolin-4(3 H)-one Derivatives Containing a Hydrazide Moiety as Potential Fungicides, ChemistrySelect. 3 (2018) 10663–10669, doi:10.1002/slct.201801575.
[16] R. Zanni, M. Galvez-Llompart, I. Garcia-Pereira, J. Galvez, R. Garcia-Domenech, Molecular topology and QSAR multi-target analysis to boost the in silico research for fungicides in agricultural chemistry, Mol. Divers. 23 (2019) 371–379, doi:10.1007/s11030-018-9879-3.
[17] W.F. Vaz, G.D.C. D'Oliveira, C.N. Perez, B.J. Neves, H.B. Napolitano, Machine learning prediction of the potential pesticide applicability of three dihydroquinoline derivatives: Syntheses, crystal structures and physical properties, J. Mol. Struct. 1206 (2020) 127732, doi:10.1016/j.molstruc.2020.127732.
[18] W.F. Vaz, J.M.F. Custodio, G.D.C. D'Oliveira, B.J. Neves, P.S.C. Junior, J.T.M. Filho, C.H. Andrade, C.N. Perez, E.P. Silveira-Lacerda, H.B. Napolitano, Dihydroquinoline derivative as a potential anticancer agent: synthesis, crystal structure, and molecular modeling studies, Mol. Divers. (2020), doi:10.1007/s11030-019-10024-x.
[19] M.R.C. de Castro, R.F. Naves, A. Bernardes, C.C. da Silva, C.N. Perez, A.F. Moura, M.O. de Moraes, F.T Martins, Tandem chalcone-sulfonamide hybridization, cyclization and further Claisen–Schmidt condensation: Tuning molecular diversity through reaction time and order and catalyst, Arab. J. Chem. 13 (2020) 1345–1354, doi:10.1016/j.arabjc.2017.11.005.
[20] G. d'Oliveira, A. Moura, M. de Moraes, C. Perez, L. Lião, Synthesis, Characterization and Evaluation of in vitro Antitumor Activities of Novel Chalcone-Quinolinone Hybrid Compounds, J. Braz. Chem. Soc. 29 (2018) 2308–2325, doi:10.21577/0103-5053.20180108.
[21] N. Snejko, C. Cascales, B. Gomez-Lor, E. Gutiérrez-Puebla, M. Iglesias, C. Ruiz-Valero, M.A. Monge, From rational octahedron design to reticulation serendipity. A thermally stable rare earth polymeric disulfonate family with CdI2-like structure, bifunctional catalysis and optical properties, Chem. Commun. (2002) 1366–1367, doi:10.1039/b202639b.
[22] G.M. Sheldrick, SHELXT – Integrated space-group and crystal-structure determination, Acta Crystallogr. A. 71 (2015) 3–8, doi:10.1107/S2053273314026370.
[23] G.M. Sheldrick, Crystal structure refinement with SHELXL, Acta Crystallogr. Sect. C. 71 (2015) 3–8, doi:10.1107/S2053229614024218.
[24] O.V. Dolomanov, L.J. Bourhis, R.J. Gildea, J.A.K. Howard, H. Puschmann, OLEX2: A complete structure solution, refinement and analysis program, J. Appl. Crystallogr. 42 (2009) 339–341, doi:10.1107/S0021889808042726.
[25] C.F. Macrae, I.J. Bruno, J. a. Chisholm, P.R. Edgington, P. McCabe, E. Pidcock, L. Rodriguez-Monge, R. Taylor, J. van de Streek, P.A. Wood, Mercury CSD 2.0 – new features for the visualization and investigation of crystal structures, J. Appl. Crystallogr. 41 (2008) 466–470, doi:10.1107/S0021889807067908.
[26] M.J. Turner, J.J. McKinnon, S.K. Wolff, D.J. Grimwood, P.R. Spackman, D. Jayatilaka, M.A. Spackman, CrystalExplorer17, University of Western Australia, 2017 http://hirshfeldsurface.net.
[27] F.H. Allen, The Cambridge Structural Database: a quarter of a million crystal structures and rising, Acta Crystallogr. Sect. B Struct. Sci. 58 (2002) 380–388, doi:10.1107/S0108768102003890.
[28] C.R. Groom, F.H. Allen, The Cambridge Structural Database in retrospect and prospect, Angew. Chemie. 53 (2014) 662–671, doi:10.1002/anie.201306438.
[29] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G.A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H.P. Hratchian, A.F. Izmaylov, J. Bloino, G. Zheng, J.L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J.A. Montgomery Jr., J.E. Peralta, F. Ogliaro, M. Bearpark, J.J. Heyd, E. Brothers, K.N. Kudin, V.N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J.C. Burant, S.S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J.M. Millam, M. Klene, J.E. Knox, J.B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, R.L. Martin, K. Morokuma, V.G. Zakrzewski, G.A. Voth, P. Salvador, J.J. Dannenberg, S. Dapprich, A.D. Daniels, Ö. Farkas, J.B. Foresman, J.V Ortiz, J. Cioslowski, D.J. Fox, Gaussian 09, Revision A.02, 34, Gaussian Inc Wallingford CT, Wallingford CT, 2009, doi:10.1159/000348293.
[30] R. Dennington, T. Keith, J. Millam, GaussView, Version 5, (2009).
[31] M.H. Jamroz, Vibrational Energy Distribution Analysis VEDA 4, Warsaw, 2004.
[32] S. Riniker, G.a Landrum, Open-source platform to benchmark fingerprints for ligand-based virtual screening, J. Cheminform 5 (2013) 26, doi:10.1186/1758-2946-5-26.
[33] L. Breiman, Random forests, Mach. Learn. 45 (2001) 5–32, doi:10.1023/A:1010933404324.
[34] A. Golbraikh, M. Shen, Z. Xiao, Y.-D. Xiao, K.-H. Lee, A. Tropsha, Rational selection of training and test sets for the development of validated QSAR models., J. Comput. Aided. Mol. Des. 17 (2003) 241–253, doi:10.1023/a:1025386326946.
[35] S. Riniker, G.a. Landrum, Similarity maps - A visualization strategy for molecular fingerprints and machine-learning methods, J. Cheminform 5 (2013) 1–7, doi:10.1186/1758-2946-5-43.
[36] A. Espinel-Ingroff, K. Dawson, M. Pfaller, E. Anaissie, B. Breslin, D. Dixon, A. Fothergill, V. Paetznick, J. Peter, M. Rinaldi, T. Walsh, Comparative and collaborative evaluation of standardization of antifungal susceptibility testing for filamentous fungi, Antimicrob. Agents Chemother. 39 (1995) 314–319.
[37] J.P. Merrick, D. Moran, L. Radom, An evaluation of harmonic vibrational frequency scale factors, J. Phys. Chem. A. 111 (2007) 11683–11700, doi:10.1021/jp073974n.
[38] A.P. Scott, L. Radom, Harmonic Vibrational Frequencies: An Evaluation of Hartree−Fock, Møller−Plesset, Quadratic Configuration Interaction, Density Functional Theory, and Semiempirical Scale Factors, J. Phys. Chem. 100 (1996) 16502–16513, doi:10.1021/jp960976r.
[39] D.L. Pavia, G.M. Lampman, G.S. Kriz, J.A. Vyvyan, Introduction to Spectroscopy, 5th Editio, Cengage Learning, Mason, Ohio, 2015.
[40] A. Dhandapani, S. Manivarman, S. Subashchandrabose, Synthesis, single crystal structure, Hirshfeld surface and theoretical investigations on pyrimidine derivative, Chem. Phys. Lett. 655–656 (2016) 17–29, doi:10.1016/j.cplett.2016.04.009.
[41] H.A.H. Hasan, Patulin and aflatoxin in brown rot lesion of apple fruits and their regulation, World J. Microbiol. Biotechnol. 16 (2000) 607–612, doi:10.1023/A:1008982511653.
[42] X. Nie, S. Yu, M. Qiu, X. Wang, Y. Wang, Y. Bai, F. Zhang, S Wang, Aspergillus flavus SUMO Contributes to Fungal Virulence and Toxin Attributes, J. Agric. Food Chem 64 (2016) 6772–6782, doi:10.1021/acs.jafc.6b02199.
[43] M.-K. Zhang, J. Tang, Z.-Q. Huang, K.-D. Hu, Y.-H. Li, Z. Han, X.-Y. Chen, L.-Y. Hu, G.-F. Yao, H. Zhang, Reduction of Aspergillus niger Virulence in Apple Fruits by Deletion of the Catalase Gene cpeB, J. Agric. Food Chem. 66 (2018) 5401–5409, doi:10.1021/acs.jafc.8b01841.
[44] L. Svetaz, A. Tapia, S.N. López, R.L.E. Furlán, E. Petenatti, R. Pioli, G. Schmeda-Hirschmann, S.A. Zacchino, Antifungal chalcones and new caffeic acids esters from Zuccagnia punctata acting against soybean infecting fungi, J. Agric. Food Chem. 52 (2004) 3297–3300, doi:10.1021/jf035213x.
[45] C.M. Jimenez, D.A. Sampietro, M.A. Sgariglia, J.R. Soberón, M.A. Vattuone, Isolation, Identification and Usefulness of Antifungal Compounds from Zuccagnia punctata for Control of Toxigenic Ear Rot Pathogens, Nat. Prod. Commun. 9 (2019) 1934578X1400901, doi:10.1177/1934578x1400901014.
[46] G.P. Rao, P. Kumar, Mathuresh-Singh, H.N. Singh, O.P. Pandey, Efficacy of chalcone, hydrazide and oxadiazole derivatives against fungal pathogens of sugarcane, 1994.
[47] R.M. Baird, S.P. Denyer, Chapter 1 - Introduction to Microbiology, in: S.P. Denyer, R.M. Baird (Eds.), Guid. to Microbiol. Control Pharm. Med. Devices, 2nd ed., CRC press, London, 2007, pp. 1–21.
[48] K.F. Nielsen, J.M. Mogensen, M. Johansen, T.O. Larsen, J.C. Frisvad, Review of secondary metabolites and mycotoxins from the Aspergillus niger group, Anal. Bioanal. Chem. 395 (2009) 1225–1242, doi:10.1007/s00216-009-3081-5.

4.5 The hydroxyl-atrazine herbicides metabolites. The intriguing aspects of ionizing states and structural motifs diversity (manuscrito)

Wesley F. Vaz[ab], Luan F. Diniz[c], Valter H. C Silva[a], Nayara Coutinho[a], Flavio Olimpio[a], Bruno, Caue Ribeiro[e], Hamilton B. Napolitano[a], Paulo de Sousa Carvalho Jr[e,]*

[a]Ciências Exatas e Tecnológicas, Universidade Estadual de Goiás, 459, 75001-970, Anápolis, GO, Brazil.
[b]Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso, 78455-000, Lucas do Rio Verde , MT, Brazil

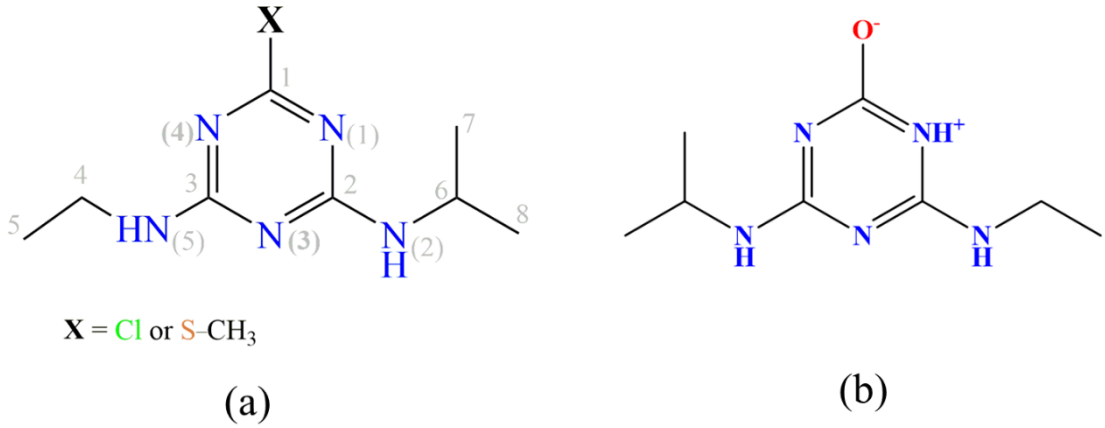
[c] Instituto de Física de São Carlos, Universidade de São Paulo, CP 369, 13560-970, São Carlos, SP, Brazil;
[d]USP/SP
[e]National Nanotechnology Laboratory for Agribusiness (LNNA), EMBRAPA Instrumentação, CP 13560-970, São Carlos, SP, Brazil
[f]Instituto de Física, Universidade Federal do Mato Grosso do Sul, 790xx-xxx, Campo Grande, Mato Grosso do Sul, Brazil.
*e-mail: paulo.sousa@ufms.br

ABSTRACT. (250 palavras).

Keywords: (4 palavras)

1. Introduction

Despite recognizing the importance of pesticides (*i.e* herbicides, insecticides, fungicides) in agriculture, their inappropriate and excessive application lead to the contamination of soil and groundwater[1–3]. When release in soil, as pollutants, they are commonly found in the environment as degraded products and metabolites resulted from abiotic and biotic processes. The degradation products, in general, more polar and more stable than the parent compound, hence, pose a greater potential risk for the environment. Ametryn (AMT) and Atrazine (ATZ), scheme 1(a) and (b), respectively, for example, are representative s-triazine herbicides and act by inhibiting photosynthetic electron transport. In the environment, ATZ and AMT are degraded at a low rate to several 2-triazine substituents[4] (Fig. S1, ESI). Among them, hydroxyatrazine (O=ATZ) is the main product of the first step degradation of s-triazines, being a nontoxic compound although persistent in the environment.



Scheme 1 - Molecular structure of ametryn (AMT) (a), and atrazine (ATZ) (a) and their metabolite (b), the hydroxyatrazine. The atoms labels are shown in grey.

The environmental pollution by triazines has raised an environmental concern due to its ecotoxicology and seeks biodegradation of these products in contaminated soils[5]. Toward the sustainable use of triazines, far beyond the optimization performance of herbicides providing dosage reduction, their metabolites, and degradation intermediates identification are demanded to predict the fate and persistence in the environment, toxicity and, establishment of degradation pathways[6].

Structurally, OH-ATZ has a hydroxyl group near to basic N-atom of triazine ring which allows the occurrence of hydroxy↔keto tautomer, a consequence of the

self-protonation of the molecule[7] leading to a diversity of solid forms. Beyond the protonation uncertainties associated with salt/cocrystal formation, compounds able to tautomerize may exhibit even more complex behavior in the solid-state, which impacts their administration and application. Herein, as part of the understating degradation process of s-triazine derivatives, we have reported, for the first time, the crystal structure of O=ATZ salts with HCl, HNO_3, methanesulfonic and oxalic acids. These products were structurally characterized by single-crystal (SCXRD) and powder X-ray diffraction (PXRD), thermal analysis (DSC and TGA), spectroscopy. Their solid-state characterization has related physicochemical properties aiming to support the understanding of its behaviour when it is formed in the environment.

2. Experimental

2.1. Chemicals

All chemicals were of analytical or chromatographic grade and were used as received without further purification. Ametryn (AMT) and atrazine (ATZ), PESTANAL® analytical standard (purity 98.5%) were purchased from Sigma-Aldrich, Brazil. The salt formers nitric (HNO_3), hydrochloric (HCl), methanosulfonic (HMes) and oxalic acids (HOxa), and the ethanol and methanol solvents were purchased from local sources.

2.2. preparation of O=ATZ salts

The O=ATZ-Cl and O=ATZ-NO_3 salts were obtained through the dissolution of ametrine in a 2:1 ethanol/water mixture, then the corresponding acid, at a concentration of 1 molar, was added to the system, while O=ATZ-OXA and O=ATZ-MES were prepared by an equimolar amount of neutral atrazine and the corresponding organic acid. Then, this mixture was then dissolved in a 2:1 ethanol/water system and the solutions were kept at room temperature for solvent evaporation. After a few hours it was possible to notice that the system acquired a yellowish color. Suitable crystals for single-crystal X-ray diffraction (SCXRD) were obtained over 2–3 days.

2.3. X-ray crystallography

The single-crystal X-ray diffraction data of the O=ATZ salts, except O=ATZ-OXA, were collected on a Bruker Apex-II CCD diffractometer with graphite-monochromated MoKα radiation (λ = 0.71073 Å), and data were measured at 120 K. Acquisition, indexing, data integration, cell determination, and final parameters were obtained using the software SAINT and SADABS[8]. The O=ATZ-OXA crystal was collected using an Agilent SuperNova diffractometer with MoKα radiation and kept at room temperature (298 K). Data collection, reduction, and cell refinement were made with the CRYSALISPRO software[9]. Using Olex2[10] the structures were solved by direct methods, and the models obtained were refined by full-matrix least-squares on F^2 (SHELXL)[11]. All non-hydrogen atoms were refined anisotropically. After that, the hydrogen atoms were located from electron-density difference maps and were positioned geometrically and refined using the riding model [Uiso(H) = 1.2Ueq or 1.5Ueq]. For the refinement of the O=AMT-NO_3 structure, which exhibits a disorder in the alkyl fragment, several restraints have been applied for splitting this group into two parts with 70:30% occupancies. Similar protocols have been applied to the refinement of the O=AMT-OXA, splitting the ethyl group, in both conformers, in two parts with 70:30% occupancies. Also, in the B conformer, restraints were applied in the isopropyl group that was split into two parts with 60:40% occupancies.

3. Results and Discussion

3.2 Characterization of ATZ metabolite

The four salts crystals of OH-ATZ (O=ATZ-Cl, O=ATZ-NO_3, O=ATZ-MES, and O=ATZ-OXA) have been obtained from the crystallization of solutions prepared by reacting the ATZ or AMT herbicide with acids (HCl, HNO_3, HMes, and HOxa) in a molar ratio of 1:1 (section 2). The OH-ATZ is the main metabolite of *s*-triazine herbicides and it is formed in situ as a result of the degradation of s-triazine herbicide at an acid medium with a low kinetic[12]. Until now, the solid-state of OH-ATZ still unknown and their features have demonstrated an intriguing aspect: the OH-ATZ

occurs as salts in the oxo-tautomer in the zwitterionic form. This characteristic is important because it determines their mobility in the soil and water ground. Towards this understanding, detailed structural and supramolecular descriptions of these salts are provided below. The asymmetric unit (ASU) view of each salt is shown in Fig. S1 ESI, while the packing diagrams are depicted in Figure S2 ESI. Crystal data, refinement, and geometric parameters of OH-ATZ salts are shown in Table S1. OH-ATZ, differently from the parent ATZ (pK_a = 1.7), is a weak base that below its pK_a value, *i.e.* 5.2, is protonated. Further reactions for the OH-ATZ are possible: the tautomerization of the compound. The OH-ATZ can exist in two different tautomeric forms (Scheme 1), *i.e.* the 2-hydroxyl tautomer (OH-ATZ) and the 2-oxo one (O=ATZ). Noteworthy, because the OH-ATZ reaches three protonation sites, the tautomerism of OH-ATZ may precede or follow/advance to protonation in the crystallization medium. From this observation, it is expected that when released into the soil, as it is being an environment of different ions, tautomerism and easy protonation/deprotonation of OH-ATZ provide high mobility in the environment. On the other hand, these features and their possible cooperative effect are related to the population and the stabilization of a particular tautomer in the medium.

3.2.1 O=ATZ-Cl crystal structure

When the O=ATZ zwitterion reacts with HCl acid, the protons transfer from the acid to the N-atom of the *s*-triazine ring to form O=ATZ HCl salt. The O=ATZ HCl crystallizes in the Triclinic space group $P\bar{1}$ with a positively charged zwitterion ($O{=}ATZ^+$), a chloride anion, and one water molecule in the ASU (Z' =1). In the $O{=}ATZ^+$ cation, the C1−O1 bond is 1.213(3) Å and corresponds to a typical double bonds C=O in ketone systems (1.222 Å). Moreover, the alkyl fragments are flexible; the $OH\text{-}ATZ^+$ adopts the antiperiplanar geometry between the protonated $N_{triazine}$ and the C7 atoms. Selected bond lengths are presented in Table 1. The $O{=}ATZ^+\cdots Cl^-$ ionic pair is stabilized by a bifurcated $NH^+\cdots Cl^-$ H-bonds (Table x). Further, the water molecule is associates to $O{=}ATZ^+$ cation by $N2H2\cdots O_w$ H-bond, and to Cl^- anion by $N5H5\cdots Cl^-$ one to form a $R_2^3(8)$ synthon. As results, adjacent $O{=}ATZ^+\cdots Cl^-$ ones are arranged into a 1D chain along the [001] direction linked by (Fig. 1a). Because the cation occurs as 2-oxo-tautomer, the inclusion of water

provides the formation of sheet structure bridging centrosymetrically chains through OH···O_{keto} H-bonds (Fig. 1a). The 3D layered packing is generated by linking these sheet frameworks through C6H6c···O1 interactions (Fig. 1b). Thus, hydrophobic zones along the structures are created by the contacts of alkyls fragments.

Although there are hydrophobic domains - the hydrophilic part that comes from hydration, facilitates the rupture of the structure by water. So, the OH-ATZ molecule can be leached, drained, and, consequently, contaminate the surface and underground of the water. It can be inferred that these structural factors are responsible for the high persistence of this pollutant in the field.

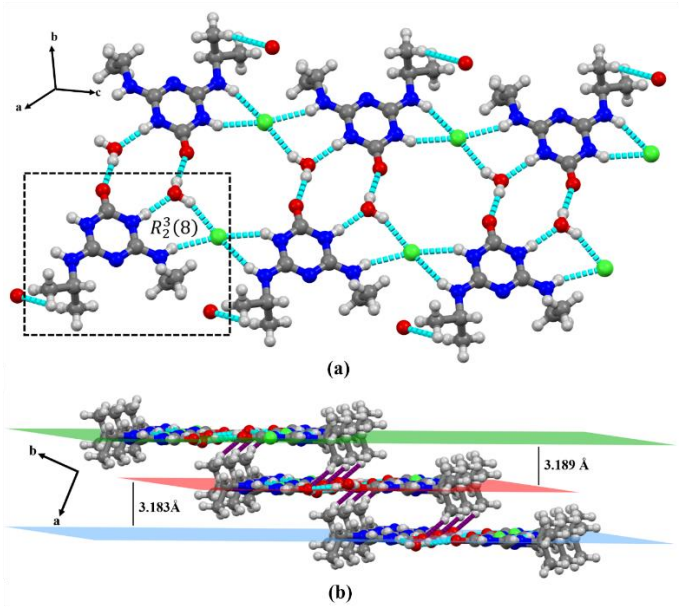


Figure 1 – The chain formed by a water molecule and ionic pair in the O=ATZ packing; the synthons are highlighted in dotted lines (a). In the crystal packing of O=ATZ-Cl, these chains are disposed into a ladder-type arrangement.

3.2.2 O=ATZ-NO_3 crystal structure

In a medium with a high concentration of HNO_3, the ATZ degrades to O=ATZ which is protonated by the acid to form O=AMT-NO_3 salt. It crystallizes in the orthorhombic space group $P2_12_12_1$ with a positively charged O=ATZ molecule and a NO_3^- counter-ion (Z' = 1) in the ASU. Since the O=ATZ^+ is H-bonded to NO_3^- anion on both sides through the $\boldsymbol{R_2^2(8)}$ synthons, a 1D chain of the ionic pairs is

formed along the [010] direction (Fig. 2a). Along to [010] direction, the chains are arranged in a square shape packing, by a twofold axis, due to the C6–H6C···O1 and C8A–H8AC···O4 interactions (Fig. 2b). Finally, these units are held together by C5—H5D···O2 forming a columnar arrangement along to [100] direction (Fig. 2b). Moreover, anion-π/triazine interactions[13] (Fig S3 ESI) were also found in the packing. It has been demonstrated that these relatively weak contacts contribute greatly to the stabilization, performing an important role in the formation of a specific crystal pattern on layered structure[14–17]. The columns are joined by dispersion forces, mainly represented by H···H contacts from the N-alkyl fragments of adjacent columns (Fig2c). This characteristic is also observed in the crystal packing of ametryn[18].

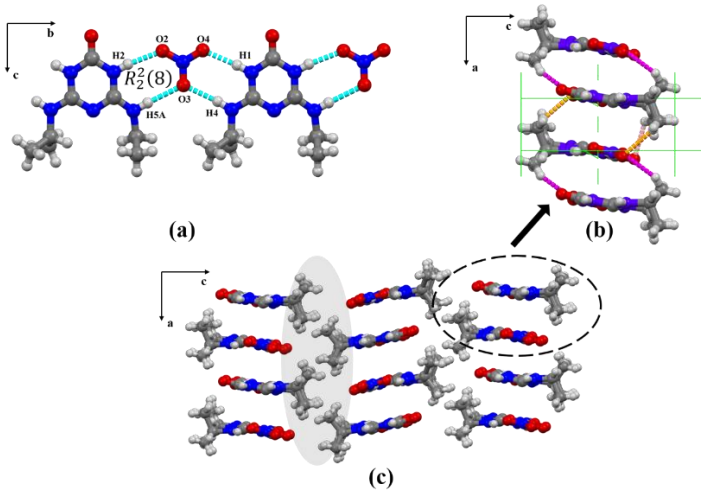


Figure 2 – The chain formed by the hydrogen bonds involved at the synthon formation (a). These chains are arranged in a square motif (b) by a 2-axis fold (the green lines) and the C5—H5D···O2 (orange), C6—H6C···O1 (pink), and C8A—H8AC···O4 (magenta) interactions. In the crystal packing (c) is possible to note hydrophobic regions (grey emphasis).

3.2.3 O=ATZ-MES crystal structure

The HMes acid is not a natural agent of soil or water ground, and it has been used in our screening to demonstrate the hydrolysis of ATZ and AMT by strong organic acids and the salt formation with O=ATZ by the s-*triazine*···SO_3^- synthon. The O=AMT-MES crystallizes in the monoclinic space group C2/c with a $O{=}ATZ^+$ cation and one Mes^- anion in the ASU. The s-triazine ring of $O{=}ATZ^+$ cation interacts on both sides with the $-SO_3^-$ portion of Mes^- anions by a $R_2^2(8)$ synthon that extends the ionic pairs in a chain along to [116] direction. In the O = AMT-Mes,

these chains cross each other almost orthogonally (Fig 3a) stabilized by π···π interactions (center-to-center: 3.668(2) Å) between the staked cations and CH···O and CH···O H-bonds between the Mes^- anion and $O{=}ATZ^+$ cations from neighboring chains. Noticeably, as an oxo-tautomer, the cation has a terminal H-bond acceptor group that has a key role in the packing of adjacent chains (Fig. 3a). From the packing of crossing chains, a like-square arrangement is observed (Fig. 3b). Like nitrate salt, along the [001] direction, these units stacked each other stabilized by CH···O interaction from the terminal alkyl groups of $O{=}ATZ^+$ cation to adjacent SO_3^- group. O=ATZ-Mes does not have a predominance of hydrophilic compounds, an overview of the packaging is shown in Fig.3c

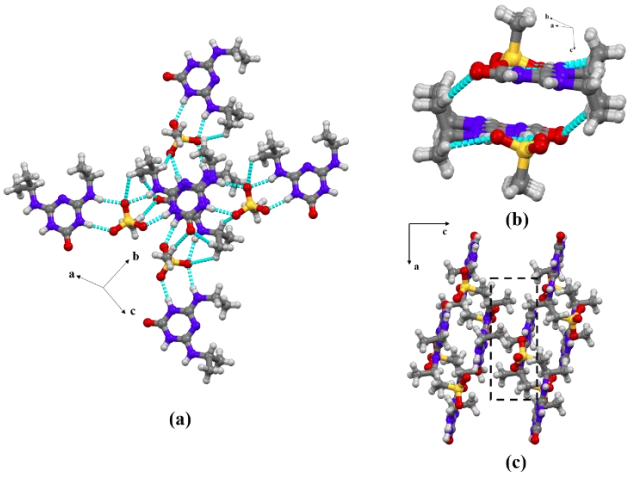
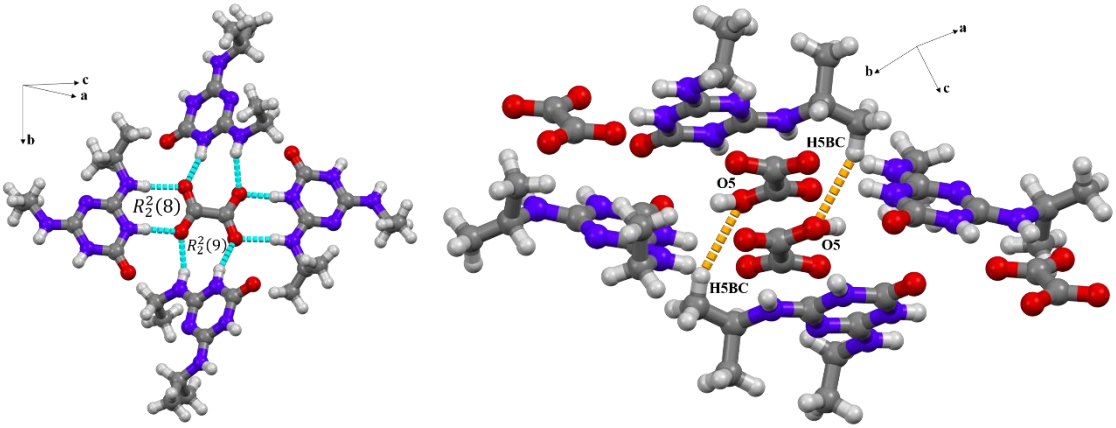


Figure 3 – The cross chains (a) formed in the molecular packing of O=AMT-MES (a); the square motif formed by C5—H5A···O1 and C6—H6C···O3 interactions (b), and a small hydrophobic zone in the structure formed by the contact of the alkyl groups of neighboring molecules (c).

3.2.4 O=ATZ-OXA crystal structure

Oxalic acid was also able to degrade AMT to the O=ATZ. Having pKa's = 1.25 and 4.25, both doubly charged oxalate (OXA^{2-}) and the hydrogen-oxalate ($HOXA^-$) anions are present in the medium, and with protonation of O=ATZ generates the O=AMT-OXA salt. The pKa of OH-ATZ[19] is equal to 5.2 this means that it can coexist protonated or non-protonated in equilibrium. The salt crystallizes in the triclinic space group $P\bar{1}$ with two independent $O{=}ATZ^+$ cations (*A*- and *B*-) and a OXA^{2-} (half molecule) and a $HOXA^-$ anions in their ASU (Z'= 2). In addition

to Δρ map analyses (see Fig. S4, ESI), the anions were differentiated considering their C–O bond length. Protonated carboxylic acids have unequal C-O bond lengths of 1.21 Å and 1.30 Å for a C-OH bond, while deprotonated carboxylates have identical C-O bond lengths of 1.26 Å due to resonance[20]. For $HOXA^-$ anion, the C–O bond length is 1.24 Å while for carboxylate is 1.28Å. When *A*- and *B*- are overlay (Fig. S5 ESI †) and geometries are checked with mogul[21,22] is possible to notice their structural differences. From the overlay is possible to notice that the ethyl group assumes opposite directions (C3B—N5B—C7B—C8B = -88º and C3A—N5A—C7A—C8A = 84º) and that the isopropyl group present 54º of difference in the conformers (C3B—N5B—C7B—C8B = -88º and C3A—N5A—C7A—C8A = 84º). In the salt, the OXA^{2-} anion is H-bonded to *A*-O=ATZ^+ cations along with its axial side forming $\boldsymbol{R_2^2(8)}$ synthons and to the B-ones by the normal direction of anion by a $\boldsymbol{R_2^2(9)}$ synthons, forming a tetrameric unit (Fig 4a). These units are further associated with the $HOXA^-$ anion resulting in a 3D network as shown in Fig. 4b.



(a) **(b)**

Figure 4 – The tetrameric unit (a) and the units forming the 3D arrangement of O=AMT-OXA. In (b) the interactions represented in (a) were not displayed to clarify.

4. Conclusions

Acknowledgments

We would like to thank the Brazilian funding agencies FAPESP and FAPEG (P.S.C.-Jr, V.H.CS, B. x grants: # 2020/xxxx), CNPq (C. R grant N. xxxx), and CAPES agencies for the financial support. The authors would also like to thank Dra. Charlane Cimini Correa (Federal University of Juiz de Fora) and Dr. Chris Herbert for allowing access to the X-ray diffraction facility.

Associated content

† Electronic supplementary information (ESI) available: X-ray Crystallographic information files (CIFs), Ortep-3 diagrams, PXRD patterns, and solubility experiments data. This material is available free of charge *via* the Internet at http://pubs.acs.org/

References

1 S. Stipičević, N. Galzina, N. Udiković-Kolić, T. Jurina, G. Mendaš, M. Dvoršćak, I. Petrić, K. Barić and V. Drevenkar, *Geoderma*, 2015, **259–260**, 300–309.

2 T. J. C. O'Riordan, *Curr. Opin. Green Sustain. Chem.*, 2018, **13**, 158–163.

3 J. M. Green, *Pest Manag. Sci.*, 2014, **70**, 1351–1357.

4 L. Chen, X. Hu, Y. Yang, C. Jiang, C. Bian, C. Liu, M. Zhang and T. Cai, *Chem. Eng. J.*, 2018, **351**, 523–531.

5 Y. Liu, L. Y. Ma, Y. C. Lu, S. S. Jiang, H. J. Wu and H. Yang, *Ecotoxicol. Environ. Saf.*, 2017, **140**, 264–270.

6 G. W. Stratton, *Arch. Environ. Contam. Toxicol.*, 1984, **13**, 35–42.

7 W. Qin, X. Yuan, L. Sun, Z. Qiang and D. Xia, *J. Hazard. Mater.*, 2019, **373**,

600–607.

8 Bruker, *APEX II*.

9 O. D. Rigaku, *CrysAlis PRO*, 2015, Rigaku Oxford Diffraction, Yarnton, England.

10 O. V. Dolomanov, L. J. Bourhis, R. J. Gildea, J. A. K. Howard and H. Puschmann, *J. Appl. Crystallogr.*, 2009, **42**, 339–341.

11 G. M. Sheldrick, *Acta Crystallogr. Sect. C Struct. Chem.*, 2015, **71**, 3–8.

12 B. Balci, N. Oturan, R. Cherrier and M. A. Oturan, *Water Res.*, 2009, **43**, 1924–1934.

13 M. Zaccheddu, C. Filippi and F. Buda, *J. Phys. Chem. A*, 2008, **112**, 1627–1632.

14 Y. V. Nelyubina, K. A. Lyssenko, D. G. Golovanov and M. Y. Antipin, *CrystEngComm*, 2007, **9**, 991.

15 L. Adriaenssens, C. Estarellas, A. Vargas Jentzsch, M. Martinez Belmonte, S. Matile and P. Ballester, *J. Am. Chem. Soc.*, 2013, **135**, 8324–8330.

16 P. U. Maheswari, B. Modec, A. Pevec, B. Kozlevčar, C. Massera, P. Gamez and J. Reedijk, *Inorg. Chem.*, 2006, **45**, 6637–6645.

17 R. E. Dawson, A. Hennig, D. P. Weimann, D. Emery, V. Ravikumar, J. Montenegro, T. Takeuchi, S. Gabutti, M. Mayor, J. Mareda, C. A. Schalley and S. Matile, *Nat. Chem.*, 2010, **2**, 533–538.

18 P. S. Carvalho, G. G. F. Guimarães, L. F. Diniz, J. Ellena and C. Ribeiro, *Green Chem.*, 2019, **21**, 6419–6429.

19 Z. Vryzas, E. Papadopoulou-Mourkidou, G. Soulios and K. Prodromou, *Eur. J. Soil Sci.*, 2007, **58**, 1186–1199.

20 J. Lin, E. Pozharski and M. A. Wilson, *Biochemistry*, 2017, **56**, 391–402.

21 I. J. Bruno, J. C. Cole, M. Kessler, J. Luo, W. D. S. Motherwell, L. H. Purkis, B. R. Smith, R. Taylor, R. I. Cooper, S. E. Harris and A. G. Orpen, *J. Chem. Inf. Comput. Sci.*, 2004, **44**, 2133–2144.

22 S. J. Cottrell, T. S. G. Olsson, R. Taylor, J. C. Cole and J. W. Liebeschuetz, *J. Chem. Inf. Model.*, 2012, **52**, 956–962.

The hydroxyl-atrazine herbicides metabolites. The intriguing aspects of ionizing states and structural motifs diversity.

Wesley F. Vaz[ab], Luan F. Diniz[c], Valter H. C Silva[a], Nayara Coutinho[a], Flavio Olimpio[a], Bruno, Caue Ribeiro[e], Hamilton B. Napolitano[a], Paulo de Sousa Carvalho Jr[e,]*

[a]Ciências Exatas e Tecnológicas, Universidade Estadual de Goiás, 459, 75001-970, Anápolis, GO, Brazil.
[b]Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso, 78455-000, Lucas do Rio Verde , MT, Brazil

[c] Instituto de Física de São Carlos, Universidade de São Paulo, CP 369, 13560-970, São Carlos, SP, Brazil;
[d]USP/SP
[e]National Nanotechnology Laboratory for Agribusiness (LNNA), EMBRAPA Instrumentação, CP 13560-970, São Carlos, SP, Brazil
[f]Instituto de Física, Universidade Federal do Mato Grosso do Sul, 790xx-xxx, Campo Grande, Mato Grosso do Sul, Brazil.
*e-mail: paulo.sousa@ufms.br

SUPPORT INFORMATION

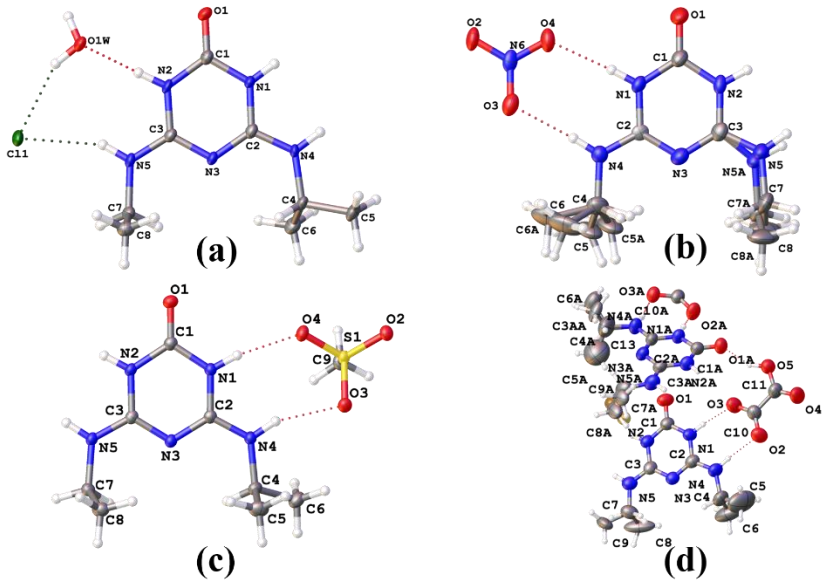


Figure S1 – Asymmetric unit of OH-ATZ-CL (a), O=ATZ-NO_3 (b), O=ATZ-MES (c), and O=ATZ-OXA (d). The ellipsoid plot is drawn at 50% of probability.

Table S1. Crystallographic parameters of the four metabolites

	O=AMT-Cl	**OH-AMT-MES**	**OH-AMT- NO_3**	**OH-AMT- OXA**
Empirical formula	$C_8H_{18}ClN_5O_2$	$C_9H_{19}N_5O_4S$	$NO_3 \cdot C_8H_{16}N_5O$	$C_{19}H_{32.7}N_{10}O_8$
Formula weight	251.72	293.35	260.27	529.25
Temperature/K	120.0	120.0	120.0	292.8
Crystal system	triclinic	Monoclinic	Orthorhombic	triclinic
Space group	P-1	C2/*c*	$P2_12_12_1$	P-1
a/Å	8.076(3)	12.715(3)	6.5104(8)	9.2094(8)
b/Å	8.137(3)	15.635(3)	9.7373(12)	11.8262(12)
c/Å	9.834(5)	14.649(3)	19.250(2)	12.1902(12)
α/°	102.360(4)	90	90	83.704(8)
β/°	95.585(7)	112.553(4)	90	85.728(8)
γ/°	90.975(5)	90	90	84.120(8)
Volume/Å3	627.8(4)	2689.6(10)	1220.3(3)	1310.0(2)
Z/ Z’	2/1	8/1	4/1	2/?
ρcalcg/cm3	1.332	1.449	1.417	1.342
μ/mm-1	0.301	0.260	0.115	0.106
F(000)	268.0	1248.0	552.0	561.0
Crystal size/mm3	0.209 × 0.175 × 0.106	0.214 × 0.126 × 0.125	0.392 × 0.369 × 0.18	? × ? × ?
Radiation	MoKα (λ = 0.71073)	MoKα (λ = 0.71073)	MoKα (λ = 0.71073)	Mo Kα (λ = 0.71073)
2Θ range for data collection/°	4.262 to 56.662	4.338 to 57.852	4.688 to 52.734	6.688 to 52.734
Index ranges	-10 ≤ h ≤ 10, -10 ≤ k ≤ 10, -13 ≤ l ≤ 13	-17 ≤ h ≤ 17, -21 ≤ k ≤ 21, -19 ≤ l ≤ 19	-8 ≤ h ≤ 8, -12 ≤ k ≤ 12, 0 ≤ l ≤ 24	-11 ≤ h ≤ 10, -14 ≤ k ≤ 14, -15 ≤ l ≤ 13
Reflections collected	9195	22695	4908	10906
Independent reflections	3077 [R_{int} = 0.0759, R_{sigma} = 0.0986]	3543 [R_{int} = 0.0528, R_{sigma} = 0.0330]	2476 [Rint = 0.0230, Rsigma = 0.0291]	5336 [R_{int} = 0.0393, R_{sigma} = 0.0796]
Data/restraints/parameters	3077/0/151	3543/0/176	2476/47/184	5336/46/382
Goodness-of-fit on F2	1.021	1.042	1.167	1.029
Final R indexes [I>=2σ (I)]	R_1 = 0.0630, wR_2 = 0.1139	R_1 = 0.0350, wR_2 = 0.0831	R1 = 0.0977, wR2 = 0.2628	R_1 = 0.1066, wR_2 = 0.2621

Table S2. Hydrogen-bond geometry (Å, º) for the four metabolites

D—H···*A*	*D*—H	H···*A*	*D*···*A*	*D*—H···*A*	**Symmetry code**
		O=ATZ-Cl			
N1—H1···Cl1	0.85	2.43	3.18	147	x,y,1+z
N2—H2···O1*W*	0.88	1.85	2.72	171	1555.03
N4—H4···Cl1	0.88	2.36	3.18	155	x,y,1+z
N5—H5···Cl1	0.88	2.42	3.21	150	x,y,1+z
O1*W*—H1*WB*···Cl1	0.87	2.33	3.14	155	x,y,1+z
O1*W*−H1*W*a···O1	0.87	1.93	2.79	168	-x,-y,1-z
C6−H6c···O1	0.98	2.68	3.64	167	-x,1-y,2-z
		O=ATZ-MES			
D—H···*A*	*D*—H	H···*A*	*D*···*A*	*D*—H···*A*	**Symmetry code**
N1—H1···O4	0.88	2.01	2.88 (2)	170	x,y,z

N2—H2···O2	0.88	2.03	2.89 (2)	166	x-1/2,+y+1/2,+z
N4—H4···O3	0.88	2.10	2.93 (2)	158	x,y,z
N5—H5···O3	0.88	2.06	2.90(2)	159	x-1/2,+y+1/2,+z
C5—H5*A*···O1	0.98	2.69	3.66 (2)	168	-x+1,-y+1,-z+1
C6—H6*A*···O1	0.98	2.49	3.44 (2)	165	-x+1,+y,-z+1/2+1
C6—H6*C*···O3	0.98	2.67	3.34 (2)	125	x,y,z
C9—H9*C*···O1	0.98	2.59	3.50 (2)	154	-x+1,+y,-z+1/2+1

O=ATZ-NO_3

D—H···*A*	*D*—H	H···*A*	*D*···*A*	*D*—H···*A*	**Symmetry code**
N1—H1···O4	0.88	2.01	2.89 (6)	175	x,y,z
N4—H4···O3	0.88	2.05	2.93 (6)	171	x,y,z
N5—H5A···O3	0.88	2.24	3.11 (4)	168	x,+y-1,+z
N2—H2···O2	0.88	2.03	2.90 (6)	170	x,+y-1,+z
C5—H5D···O2	0.98	2.69	3.60 (4)	154	-x,+y-1/2,-z+1/2
C6—H6C···O1	0.98	2.59	3.15 (4)	116	-x+1,+y+1/2,-z+1/2
C8A—H8AC···O4	0.98	2.60	3.41 (2)	140	-x+1,+y-1/2,-z+1/2

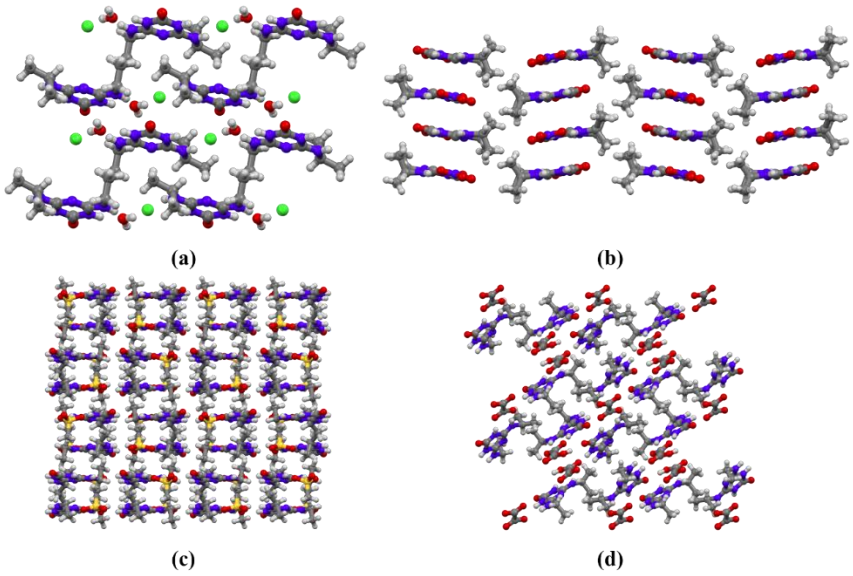


Figure S2 – The crystal packing of O=AMT-Cl (a), O=AMT-NO_3 (b) , O=AMT-MES (c), and O=AMT-OXA (d). (a) and (b) are represented in b view while (c) and (d) in a view

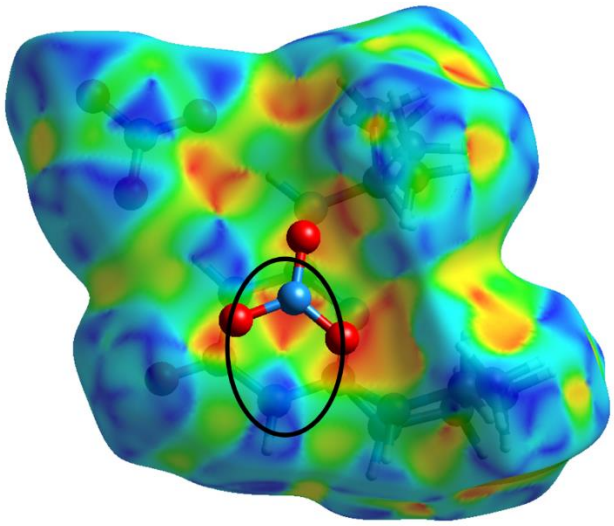


Figure S3 – The bowtie style shape, circled in black, formed by opposite blue and red triangles, in a shape index Hirshfeld surface, confirms the anion-π/triazine interactions in the crystal packing of O=AMT-NO_3.

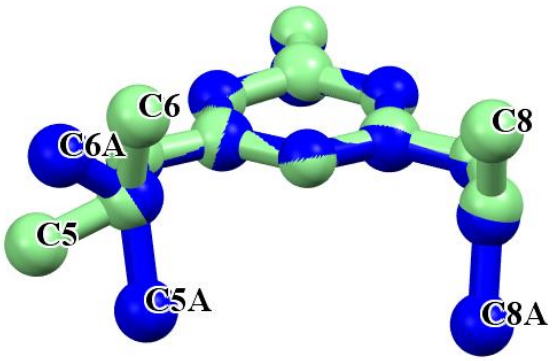


Figure S5 – The overlay of A (blue) and B (light green) conformers of O=AMT-OXA.

5 CONSIDERAÇÕES FINAIS

O estudo estrutural das dihidroquinolinas sugeriu o potencial fungicida dessa classe de compostos e a caracterização dos produtos de degradação de s-triazinas indicou seu impacto no meio ambiente. Os resultados obtidos corroboram os objetivos propostos. Primeiramente, os compostos foram estruturalmente elucidados por meios cristalográficos (difração dos raios X), gerando a entrada de mais dez novas moléculas dessa classe para o banco de dados de pequenas moléculas. Em seguida, com a aplicação dos modelos de *machine learning*, investigou-se o potencial biológico do conjunto em uma possível atividade anticâncer, que foi confirmada por ensaios citotóxicos. Percebeu-se que a anel clorofenílico forma uma interação com o resíduo Phe171 e promove um impedimento estérico no resíduo Cys303, enquanto o anel nitrobenzênico interage com o resíduo Tyr297. Com base nisso, os métodos computacionais foram utilizados com enfoque no potencial pesticida de três dihidroquinolinas.

Os resultados mostraram que as moléculas podem atuar como antiviral (vírus do mosaico do tabaco) e antifúngico (*Fusarium oxysporum*). Por último, dois isômeros foram apontados com potencial antifúngico (*A. niger* e *S. slerotiorum*) e posteriormente comprovado por ensaio antifúngico para o *A. niger*. Em relação ao *A. niger*, apenas o grupo substituinte ligado ao anel benzênico, que está ligado a dupla ligação, não contribui com a atividade biológica. Enquanto isso, para o *S. slerotiorum* o grupo ligado ao anel benzênico, que está ligado a dupla ligação, contribui significativamente para atividade, juntamente com o núcleo quinolínico.

Com relação ao metabolito dos herbicidas ametrina e atrazina, foram estruturalmente elucidados e caracterizados quatro de seus sais, esse conhecimento auxiliará no entendimento da cinética de degradação dos resíduos desses herbicidas. Dessa maneira, essa tese contribui para o campo de estudo nos seguintes aspectos: no desenvolvimento de algoritmos de *machine learning* que predizem corretamente, uma vez que há validação experimental, as dihidroquinolinas em estudo como compostos com ação fungicida; pioneirismo na detalhada caracterização estrutural de dez dihidroquinolinas e quatro produtos de degradação dos herbicidas ametrina e atrazina; e, por fim, mostra que o entendimento estrutural pode, e deve, ser utilizado

como instrumento de partida no planejamento racional de compostos com atividade pesticida.

Além disso, outras atividades acadêmicas também foram realizadas. Dentre elas destaca-se a participação nos seguintes eventos: IUCr-Unesco *Bruker OpenLab* 2, realizado no ano de 2016 em Montevidéu – URU (curso de verão); *67th Annual Meeting of the American Crystallographic Association* (ACA), realizado no ano de 2017 em Nova Orleans – EUA (apresentação de banner) e III *Reunión de la Asociación Latinoamericana de Cristalografía*, realizado no ano de 2018 em Valparaíso – CHI (apresentação oral). Destaca-se também a participação no Programa de Doutorado Sanduíche no Exterior (PDSE - CAPES) na *University of Notre Dame* (Indiana - EUA) durante os meses de novembro de 2018 a abril de 2019 e no programa *Co-financed ShortTerm Research Grant Brazil, 2019*, financiado pelo Serviço alemão de intercambio acadêmico (*Deutscher Akademischer Austauschdienst* - DAAD), na *Technische Universität Darmstadt* durante os meses de novembro de 2020 a janeiro de 2021.

Finalmente, em termos de produção científica, fui coautor em outros artigos em áreas como: materiais ópticos (CUSTODIO *et al.*, 2019a, 2020a, 2020b; DE PAULA *et al.*, 2019; RAVINDRACHARY *et al.*, 2005; VALVERDE *et al.*, 2017), biocombustíveis (FARIA *et al.*, 2021) , elucidação estrutural (AGUIAR *et al.*, 2020; CARVALHO *et al.*, 2017; CUSTODIO *et al.*, 2017; MICHELINI *et al.*, 2020, 2019; MOREIRA *et al.*, 2019), meio-ambiente (BOAVENTURA *et al.*, 2018, 2020) e saúde (CUSTODIO et al., 2018, 2019b; MICHELINI et al., 2018; SALLUM et al., 2019).

REFERÊNCIAS

ABBASI, M.; EL HANANDEH, A. Forecasting municipal solid waste generation using artificial intelligence modelling approaches. **Waste Management**, v. 56, p. 13–22, 2016.

AGUIAR, A. S. N. *et al*. Synthesis, characterization, and computational study of a new heteroaryl chalcone. **Journal of Molecular Modeling**, v. 26, n. 9, p. 243, 2020.

BANERJEE, P. *et al*. ProTox-II: a webserver for the prediction of toxicity of chemicals. **Nucleic Acids Research**, v. 46, n. W1, p. W257–W263, 2018.

BARTÓK, A. P. *et al*. Machine learning unifies the modeling of materials and molecules. **Science Advances**, v. 3, n. 12, p. e1701816, 2017.

BARONAS, R. L. Agrotóxico versus pesticida: notas de leitura sobre polêmica e amemória discursiva. Bakhtiniana: **Revista de Estudos do Discurso**, v. 14, n. 2, p. 62–87, 2019.

BENFENATI, E. *et al*. Chapter 2 - Databases for pesticide ecotoxicity. *In*: BENFENATI, E. (org.). **Quantitative Structure-Activity Relationships (QSAR) for Pesticide Regulatory Purposes**. Amsterdam: Elsevier, 2007. p. 59–81.

BERTHON, G. Where Do Agrochemicals Come From?. **CHIMIA International Journal for Chemistry**, v. 74, n. 10, p. 822–824, 2020.

BETHSASS, J.; COLANGELO, A. European Union Bans Atrazine, While the United States Negotiates Continued Use. **International Journal of Occupational and Environmental Health**, v. 12, n. 3, p. 260–267, 2006.

BLAKE, A. J.; CLEGG, W. Introduction to diffraction. *In*: **Crystal structure analysis**: principles and practice. International Union of Crystallography book series; 13. 2nd ed. ed. Oxford; Oxford University Press, 2009. p. 9–24.

BOAVENTURA, K. D. J. *et al*. A Agroecologia Ontem, Hoje e Amanhã em Terras Tupiniquins: Conceito, Contexto e Perspectivas. **História Ambiental, Latinoamericana y Caribeña**, v. 8, n. 2, p. 180–209, 2018.

BOAVENTURA, K. D. J. *et al*. Educação Ambiental e Percepção Acerca do Fogo e seus Impactos no Cerrado: Uma Pesquisa Qualitativa. **Fronteiras: Journal of Social, Technological and Environmental Science**, v. 9, n. 3, p. 355–379, 2020.

BOSC, N. *et al*. Large-scale comparison of QSAR and conformal prediction methods and their applications in drug discovery. **Journal of Cheminformatics**, v. 11, n. 1, p. 4, 2019.

CANTRELL, C. L.; DAYAN, F. E.; DUKE, S. O. Natural Products as Sources for New Pesticides. **Journal of Natural Products**, v. 75, n. 6, p. 1231–1242, 2012.

CARVALHO, P. S. *et al*. Conformation analysis of a novel fluorinated chalcone. **Journal of Molecular Modeling**, v. 23, n. 3, p. 97, 2017.

CARVALHO, P. S. *et al*. Highly water-soluble agrichemicals by using engineered organic salts for reducing adverse environmental impacts. **Green Chemistry**, v. 21, n. 23, p. 6419–6429, 2019.

CASTRO, M. R. C. *et al*. Conformational Variability in Sulfonamide Chalcone Hybrids: Crystal Structure and Cytotoxicity. **Journal of the Brazilian Chemical Society**, v. 27, n. 5, p. 884–898, 2015.

CHARLTON, R. L. *et al*. Repurposing as a strategy for the discovery of new anti-leishmanials: the-state-of-the-art. **Parasitology**, v. 145, n. 2, p. 219–236, 2018.

CIPULLO, S. *et al*. Prediction of bioavailability and toxicity of complex chemical mixtures through machine learning models. **Chemosphere**, v. 215, p. 388–395, 2019.

COJOCARU, O. A. *et al*. Ionic liquid forms of the herbicide dicamba with increased efficacy and reduced volatility. **Green Chemistry**, v. 15, n. 8, p. 2110, 2013.

CONSONNI, V.; TODESCHINI, R. Molecular Descriptors. *In*: T., P.; J., L.; M, C. (org.). **Recent Advances in QSAR Studies. Challenges and Advances in Computational Chemistry and Physics**. Dordrecth: Springer, 2010. p. 29–102.

CORREY, G. J. *et al*. Overcoming insecticide resistance through computational inhibitor design. **Proceedings of the National Academy of Sciences**, v. 116, n. 42, p. 21012–21021, 2019.

CULLITY, B. D. **Elements of Proteins X-ray Crystallography**. Massachusetts: Addison-Wesley, 1978.

CUSTODIO, J. M. F. *et al*. The Influence of Methoxy and Ethoxy Groups on Supramolecular Arrangement of Two Methoxy-chalcones. **Journal of the Brazilian Chemical Society**, v. 28, n. 11, p. 2180–2191, 2017.

CUSTODIO, J. M. F. *et al*. Structural insights into a novel anticancer sulfonamide chalcone. **New Journal of Chemistry**, v. 42, n. 5, p. 3426–3434, 2018.

CUSTODIO, J. M. F. *et al*. Sulphonamide chalcones: Conformationally diverse yet optically similar. **Journal of Molecular Structure**, v. 1198, p. 126896, 2019a.

CUSTODIO, J. M. F. *et al*. Benzenesulfonyl incorporated chalcones: Synthesis, structural and optical properties. **Journal of Molecular Structure**, v. 1208, p. 127845, 2020a.

CUSTODIO, J. M. F. *et al*. On the potential as nonlinear optical material of a new chalcone derivative and its crystal and topological analysis**. Journal of Molecular Structure**, v. 1201, p. 127131, 2020b.

CUSTODIO, J. M. F. F. *et al*. Solvent-driven structural adaptation in a novel anticancer sulfonamide chalcone. **Journal of Molecular Structure**, v. 1175, p. 389–397, 2019b.

D'OLIVEIRA, G. *et al*. Synthesis, Characterization and Evaluation of in vitro Antitumor Activities of Novel Chalcone-Quinolinone Hybrid Compounds. **Journal of the Brazilian Chemical Society**, v. 29, n. 11, p. 2308–2325, 2018.

DE PAULA, R. L. G. *et al*. A Comprehensive Topological Analysis on a New Bromine-Chalcone with Potential Nonlinear Optical Properties. **The Journal of Physical Chemistry A**, v. 123, n. 40, p. 8632–8643, 2019.

DENG, J.-H. *et al*. π-π stacking interactions: Non-negligible forces for stabilizing porous supramolecular frameworks. **Science Advances**, v. 6, n. 2, p. eaax9976, 2020.

DESIRAJU, G. R. Crystal engineering: structure, property and beyond. **IUCrJ**, v. 4, n. 6, p. 710–711, 2017.

DOLOMANOV, O. V. *et al*. OLEX2: a complete structure solution, refinement and analysis program. **Journal of Applied Crystallography**, v. 42, n. 2, p. 339–341, 2009.

DUDEK, A.; ARODZ, T.; GALVEZ, J. Computational Methods in Developing Quantitative Structure-Activity Relationships (QSAR): A Review. **Combinatorial Chemistry & High Throughput Screening**, v. 9, n. 3, p. 213–228, 2006.

EL-GAMAL, K. M. *et al*. Synthesis, docking, QSAR, ADMET and antimicrobial evaluation of new quinoline-3-carbonitrile derivatives as potential DNA-gyrase inhibitors. **Journal of Molecular Structure**, v. 1166, p. 15–33, 2018.

ERICKSON, L. E.; LEE, K. H.; SUMNER, D. D. Degradation of atrazine and related s-triazines. **Critical Reviews in Environmental Control**, v. 19, n. 1, p. 1–14, 1989.

FARIA, E. C. M. *et al*. Comparative Study of Chalcones and Their Potential as Additives for Biofuels. **Energy & Fuels**, v. 35, n. 1, p. 552–560, 2021.

FELCZAK, A. *et al*. Quinoline biodegradation by filamentous fungus Cunninghamella elegans and adaptive modifications of the fungal membrane composition. **Environmental Science and Pollution Research**, v. 23, n. 9, p. 8872–8880, 2016.

FILALI BABA, Y. *et al*. Synthesis, crystal structure, spectroscopic characterization, Hirshfeld surface analysis, molecular docking studies and DFT calculations, and antioxidant activity of 2-oxo-1,2-dihydroquinoline-4-carboxylate derivatives**.** **Journal of Molecular Structure**, v. 1188, p. 255–268, 2019.

FOUCOUT, L. *et al*. Synthesis, radiosynthesis and biological evaluation of 1,4-dihydroquinoline derivatives as new carriers for specific brain delivery. **Organic & Biomolecular Chemistry**, v. 7, n. 18, p. 3666, 2009.

FREEDMAN, B. PESTICIDES. *In*: FREEDMAN, B. (org.). **Environmental Ecology.** 2 ed. San Diego: Academic Press, 1995. p. 213–277.

GÁSPÁRI, Z.; PERCZEL, A. Protein Dynamics as Reported by NMR. In: **Protein Dynamics as Reported by NMR**. Elsevier, 2010. p. 35–75.

GIACOVAZZO, C. *et al*. **Fundamentals of Crystallography**. 3 ed. Oxford: Oxford University Press, 2011.

GLUSKER, J. P.; TRUEBLOOD, K. N. **Crystal Structure Analysis**: A Primer. OUP Oxford, 2010.

GRILLO, R.; ABHILASH, P. C.; FRACETO, L. F. Nanotechnology Applied to Bio-Encapsulation of Pesticides. **Journal of Nanoscience and Nanotechnology**, v. 16, n. 1, p. 1231–1234, 2016.

GROOM, C. R. *et al*. The Cambridge structural database. **Acta Crystallographica Section B: Structural Science, Crystal Engineering and Materials**, v. 72, n. 2, p. 171–179, 2016.

GU, Y.; LI, M. Molecular Modeling. In: ISHIDA, H.; AGAG, T. (org.). **Handbook of Benzoxazine Resins**. Elsevier, 2011. p. 712.

JESCHKE, P. Latest generation of halogen-containing pesticides. **Pest Management Science**, v. 73, n. 6, p. 1053–1066, 2017.

JESCHKE, P. Current status of chirality in agrochemicals. **Pest Management Science**, v. 74, n. 11, p. 2389–2404, 2018.

JITONNOM, J. Computer-aided pesticide design: A short review. *In*: CHANDRASEKAR, R.; TYAGI, B. K.; GUI, Z. Z. (org.). **Short Views on Insect Biochemistry and Molecular Biology**. 1 ed. South India: International Book Mission Academic, 2014. p. 685 – 707.

KLINK, C. A. *et al*. A conservação do Cerrado brasileiro. **Megadiversidade**, v. 1, n. 1, p. 147–155, 2005.

KOUROU, K. *et al*. Machine learning applications in cancer prognosis and prediction. **Computational and Structural Biotechnology Journal**, v. 13, p. 8–17, 2015.

KUMAR, G. *et al*. Synthesis and biological evaluation of dihydroquinoline carboxamide derivatives as anti-tubercular agents. **European Journal of Medicinal Chemistry**, v. 157, p. 1–13, 2018.

LAMBERTH, C. *et al*. Current Challenges and Trends in the Discovery of Agrochemicals. **Science**, v. 341, n. 6147, p. 742–746, 2013.

LE, T. *et al*. Persistence of a self-complementary N–H···N tape motif in chloro-s-triazine crystals: crystal structures of simazine and atrazine herbicides and their polymorphic and inclusion behaviour. **CrystEngComm**, v. 18, n. 6, p. 962–970, 2016.

LIBBRECHT, M. W.; NOBLE, W. S. Machine learning applications in genetics and genomics. **Nature Reviews Genetics**, v. 16, n. 6, p. 321–332, 2015.

LIU, X.-H. *et al*. Synthesis and in vivo fungicidal activity of some new quinoline derivatives against rice blast. **Pest Management Science**, v. 73, n. 9, p. 1900–1907, 2017a.

LIU, Y. et al. Comprehensive analysis of degradation and accumulation of ametryn in soils and in wheat, maize, ryegrass and alfalfa plants. Ecotoxicology and Environmental Safety, v. 140, p. 264–270, jun. 2017b.

MANN, S. R.; KAUFMAN, E. P. Natural Product Pesticides: Their Development, Delivery and Use Against Insect Vectors. **Mini-Reviews in Organic Chemistry**, v. 9, n. 2, p. 185–202, 2012.

MARCHI, G. Determinação de atrazina em solos e água em áreas de cultivo de milho irrigado do Distrito Federal. **Embrapa**, 2011. Disponível em: <https://www.embrapa.br/busca-de-projetos/-/projeto/22922/determinacao-de-atrazina-em-solos-e-agua-em-areas-de-cultivo-de-milho-irrigado-do-distrito-federal#:~:text=A%20atrazina%20%C3%A9%20o%20segundo,daninhas%20na%20cultura%20do%20milho.>. Acesso em: 16 dez. 2020.

MARELLA, A. *et al*. Quinoline: A versatile heterocyclic. **Saudi Pharmaceutical Journal**, v. 21, n. 1, p. 1–12, 2013.

MARTINS, J. P. A; FERREIRA, M. M. C. Qsar Modeling: a New Open Source Computational Package To Generate and Validate Qsar Models. **Quimica Nova**, v. 36, n. 4, p. 554-U250, 2013.

MEENA, R. *et al*. Impact of Agrochemicals on Soil Microbiota and Management: A Review. **Land**, v. 9, n. 2, p. 34, 2020.

MEHER, P. K. *et al*. DIRProt: a computational approach for discriminating insecticide resistant proteins from non-resistant proteins. **BMC Bioinformatics**, v. 18, n. 1, p. 190, 2017.

MICHELINI, L. J. *et al*. Synthesis, Characterization and Conformational Analysis of Two Novel 4(1H)-Quinolinones. **Journal of the Brazilian Chemical Society**, v. 31, n. 1, p. 66–78, 2020.

MICHELINI, L. J. *et al*. A novel potential anticancer chalcone: Synthesis, crystal structure and cytotoxic assay. **Journal of Molecular Structure**, v. 1168, p. 309–315, 2018.

MICHELINI, L. J. *et al*. Analysis of two novel 1–4 quinolinone structures with bromine and nitrobenzyl ligands. **Journal of Molecular Modeling**, v. 25, n. 3, p. 55, 2019.

MISHRA, P. *et al*. Insecticides Derived from Natural Products: Diversity and Potential Applications. *In*: **Natural Bioactive Products in Sustainable Agriculture**. Singapore: Springer Singapore, 2020. p. 83–99.

MONTAVON, G. *et al*. Machine learning of molecular electronic properties in chemical compound space. **New Journal of Physics**, v. 15, n. 9, p. 095003, 2013.

MOREIRA, C. A. *et al*. A comprehensive study on crystal structure of a novel sulfonamide-dihydroquinolinone through experimental and theoretical approaches. **Journal of Molecular Modeling**, v. 25, n. 7, p. 205, 2019.

MÜLLER, U.; WONDRATSCHEK, H. **International Tables for Crystallography**, Volume A1. Acta Crystallographica Section A Foundations of Crystallography, 2005.

NAPOLITANO, H. B. *et al*. Inovação e Biotecnologia na Biodiversidade do Cerrado. **Fronteiras: Journal of Social, Technological and Environmental Science**, v. 5, n. 2, p. 162–180, 2016.

NAUHA, E.; NISSINEN, M. Co-crystals of an agrochemical active - A pyridine-amine synthon for a thioamide group. **Journal of Molecular Structure**, v. 1006, n. 1–3, p. 566–569, 2011.

NEFERTITI, A. S. G. *et al*. In Vitro and In Vivo Studies of the Trypanocidal Effect of Novel Quinolines. **Antimicrobial Agents and Chemotherapy**, v. 62, n. 2, p. 1-11, 2017.

OHNO, H.; YOSHIZAWA-FUJITA, M.; KOHNO, Y. Design and properties of functional zwitterions derived from ionic liquids. **Physical Chemistry Chemical Physics**, v. 20, n. 16, p. 10978–10991, 2018.

PAUKKU, Y.; HILL, G. Quantum topological molecular descriptors in QSAR analysis of organophosphorus compounds. **International Journal of Quantum Chemistry**, v. 112, n. 5, p. 1343–1352, 2012.

PEIXOTO, J. DE C. et al. Flavonoids from Brazilian Cerrado: Biosynthesis, Chemical and Biological Profile. **Molecules**, v. 24, n. 16, p. 2891, 2019.

PERGAL, M. V. *et al*. Degradation Products, Mineralization, and Toxicity Assessment of Pesticides Malathion and Fenitrothion. **Water, Air, & Soil Pollution**, v. 231, n. 8, p. 433, 2020.

PIMENTEL, A. S.; GUIMARÃES, C. R. W.; MILLER, Y. Molecular Modeling: Advancements and Applications. **Journal of Chemistry**, v. 2013, p. 1–2, 2013.

PRATA, F. *et al*. Degradação e sorção de ametrina em dois solos com aplicação de vinhaça. **Pesquisa Agropecuária Brasileira**, v. 36, n. 7, p. 975–981, 2001.

QUINTELLA, C. M. *et al*. Maturidade Tecnológica: níveis de prontidão TRL. *In*: NÚBIA, A.; RIBEIRO, M. (org.). **Prospecção Tecnológica**. 2 ed. Salvador: 2019

RAVINDRACHARY, V. *et al*. Growth and characterization of chalcone derivative single crystal. **Journal of Crystal Growth**, v. 275, p. 313–318, 2005.

ROSADO-SOLANO, D. N. *et al*. Synthesis, Biological Evaluation and In Silico Computational Studies of 7-Chloro-4-(1 H -1,2,3-triazol-1-yl)quinoline Derivatives: Search for New Controlling Agents against Spodoptera frugiperda (Lepidoptera: Noctuidae) Larvae. **Journal of Agricultural and Food Chemistry**, v. 67, n. 33, p. 9210–9219, 2019.

REIS, M. R. *et al*. **Aproveitamento atual e potencial dos cerrados: base física e potencialidades da região**. Brasília: IPEA, 1973.

SALLUM, L. O. O. *et al*. Synthesis, conformational analysis and molecular docking studies on three novel dihydropyrimidine derivatives. **Journal of Molecular Structure**, v. 1192, p. 274–287, 2019.

SANTOS, D.; VARDANEGA, R.; DE ALMEIDA, M. A. Intensification of bioactive compounds extraction from medicinal plants using ultrasonic irradiation. **Pharmacognosy Reviews**, v. 8, n. 16, p. 88, 2014.

SCHUT, M. *et al*. Systems approaches to innovation in crop protection. A systematic literature review. **Crop Protection**, v. 56, p. 98–108, 2014.

SECHI, M. *et al*. Design and synthesis of novel dihydroquinoline-3-carboxylic acids as HIV-1 integrase inhibitors. **Bioorganic & Medicinal Chemistry**, v. 17, n. 7, p. 2925–2935, 2009.

SINGH, S. B.; PELAEZ, F. Biodiversity, chemical diversity and drug discovery. **Progress in Drug Research**, v. 65, p. 142–174, 2008.

SINHA, K.; GHOSH, J.; SIL, P. C. New pesticides: a cutting-edge view of contributions from nanotechnology for the development of sustainable agricultural pest control. *In*: **New Pesticides and Soil Sensors**. Elsevier, 2017. p. 47–79.

SMART, L. E.; MOORE, E. A. An Introduction to Crystal Structures. *In*: **Solid State Chemistry**: An Introduction. 4 ed. Boca Raton: Taylor & Francis, 2012. p. 1-70.

SPARKS, T. C. Insecticide discovery: An evaluation and analysis. **Pesticide Biochemistry and Physiology**, v. 107, n. 1, p. 8–17, 2013.

STOMPOR, M. *et al*. Insect Antifeedant Potential of Xanthohumol, Isoxanthohumol, and Their Derivatives. **Journal of Agricultural and Food Chemistry**, v. 63, n. 30, p. 6749–6756, 2015.

TILLEY, R. J. D. **Crystals and Crystal Structures**. 1ed. West Sussex: John Wiley & Sons Ltd, 2006.

VALVERDE, C. *et al*. The solid-state structure and environmental polarization effect of a novel asymmetric azine. **New J. Chem**., v. 41, n. 19, p. 11361–11371, 2017.

VAZ, W. F. *et al*. Dihydroquinoline derivative as a potential anticancer agent: synthesis, crystal structure, and molecular modeling studies. **Molecular Diversity**, 2020a.

VAZ, W. F. *et al*. Machine learning prediction of the potential pesticide applicability of three dihydroquinoline derivatives: Syntheses, crystal structures and physical properties. **Journal of Molecular Structure**, v. 1206, p. 127732, 2020b.

VAZ, W. F. *et al*. Structural comparison of five new halogenated dihydroquinoline-4(1H)-ones. **Journal of Molecular Structure**, v. 1219, p. 128559, 2020c.

VAZ, W. F. *et al*. In silico-driven identification and structural analysis of nitrodihydroquinolinone pesticide candidates with antifungal activity. **Journal of Molecular Structure**, v. 1226, p. 129326, 2021.

WANG, Z. Claisen-Schmidt Condensation. *In*: **Comprehensive Organic Name Reactions and Reagents**. Hoboken: John Wiley & Sons, Inc., 2010. p. 660–664.

WEYESA, A.; MULUGETA, E. Recent advances in the synthesis of biologically and pharmaceutically active quinoline and its analogues: a review. **RSC Advances**, v. 10, n. 35, p. 20784–20793, 2020.

WU, G. *et al*. Machine learning models for predicting PAHs bioavailability in compost amended soils. **Chemical Engineering Journal**, v. 223, p. 747–754, 2013.

YANG, J. *et al*. Mechanisms, kinetics and eco-toxicity assessment of singlet oxygen, sulfate and hydroxyl radicals-initiated degradation of fenpiclonil in aquatic environments. **Journal of Hazardous Materials**, p. 124505, 2020.

YURIEV, E.; COOTE, M. L. Molecular Modelling: Advances in Biomolecular and Materials Modelling. **Australian Journal of Chemistry**, v. 64, n. 7, p. 885, 2011.

YUSOFF, S. N. M.; KAMARI, A.; ALJAFREE, N. F. A. A review of materials used as carrier agents in pesticide formulations. **International Journal of Environmental Science and Technology**, v. 13, n. 12, p. 2977–2994, 2016.

ZANNI, R. *et al*. Molecular topology and QSAR multi-target analysis to boost the in silico research for fungicides in agricultural chemistry. **Molecular Diversity**, v. 23, n. 2, p. 371–379, 2019.

ZENG, Y.-M.; ZENG, H.-H.; LIU, F.-M. Synthesis and Structure Characterization of 1,5-Benzothiazepine Derivatives Bearing Quinoline Moiety. **Journal of Heterocyclic Chemistry**, v. 53, n. 3, p. 887–893, 2016.

ZHANG, Y. *et al*. Physicochemical property guidelines for modern agrochemicals. **Pest Management Science**, v. 74, n. 9, p. 1979–1991, 2018.

ZHU, X. *et al*. Manipulating Solid Forms of Contact Insecticides for Infectious Disease Prevention. **Journal of the American Chemical Society**, v. 141, n. 42, p. 16858–16864, 2019.