



**UNIVERSIDADE DE BRASÍLIA
INSTITUTO DE QUÍMICA**

Lucas Siqueira Rodrigues

**PREVISÃO DA PRESSÃO DE VAPOR DE COMPOSTOS ORGÂNICOS AMBIENTALMENTE
RELEVANTES ATRAVÉS DE CÁLCULOS QUANTO-MECÂNICOS**

TRABALHO DE CONCLUSÃO DE CURSO 2

Brasília – DF

2.º/2023



UNIVERSIDADE DE BRASÍLIA
INSTITUTO DE QUÍMICA

Lucas Siqueira Rodrigues

**PREVISÃO DA PRESSÃO DE VAPOR DE COMPOSTOS ORGÂNICOS AMBIENTALMENTE
RELEVANTES ATRAVÉS DE CÁLCULOS QUANTO-MECÂNICOS**

Trabalho de Conclusão de Curso de Bacharelado em Química apresentado ao Instituto de Química da Universidade de Brasília, como requisito parcial para a obtenção do título de Bacharel em Química.

Orientador: Daniel Francisco Scalabrini Machado

Brasília - DF

2.º/2023

RESUMO

A pressão de vapor é um parâmetro que descreve a volatilidade de uma substância, dependendo da magnitude da pressão de vapor, diferentes métodos experimentais são empregados para realizar a sua medição, mas geralmente estão associados a uma grande quantidade de esforço e se tornam menos precisos à medida que a pressão de vapor diminui. Para prever a pressão de vapor, usualmente são empregados algoritmos que geralmente são baseados em relações quantitativas de estrutura-atividade. Neste trabalho é realizada uma abordagem alternativa, menos empírica, onde a mudança na energia livre de Gibbs para a transição da fase condensada para a fase gasosa é obtida a partir de conjuntos conformacionais calculados separadamente para cada fase. Esse fluxo de trabalho é denominado CRENSO, e o mesmo é capaz de fornecer pressões de vapor confiáveis em certa medida para líquidos em uma ampla faixa de pressão. Empregando uma modificação do fluxo CRENSO original em o modelo de solvatação é o SMD, mostramos que esta proposta pode ser utilizada para previsão da pressão de vapor de compostos orgânicos pouco voláteis de interesse ambiental.

Palavras-chave: CREST, CENSO, CRENSO, xTB, COSMO-RS, ALPB, pressão de vapor.

ABSTRACT

Vapor pressure is a parameter that describes the volatility of a substance, depending on the magnitude of vapor pressure, different experimental methods are used to measure it, but they are generally associated with a large amount of effort and become less accurate as vapor pressure decreases. To predict vapor pressure, algorithms have been developed that are generally based on quantitative structure-activity relationships (QSAR). In this work, an alternative, less empirical approach is carried out, where the change in Gibbs free energy for the transition from the condensed phase to the gas phase is obtained from conformational sets calculated separately for each phase. This workflow is called CRENSO, and it is capable of providing reliable vapor pressures for liquids over a wide pressure range. By employing a modification of the original CRENSO workflow, in which the SMD solvation model is used instead, we show that such approach can be used to predict the vapor pressure of low volatile organic compounds with environmental interest.

Keywords: CREST, CENSO, CRENSO, xTB, COSMO-RS, ALPB, vapor pressure.

LISTA DE FIGURAS

Figura 1. Diagrama de fases genérico de uma substância normal. O ponto normal de ebulição b.p. é destacado bem como a curva de coexistência sólido-vapor (Ps) e as condições de metaestabilidade de líquido sub-resfriado (linha tracejada). Extraído da referência [14].	3
Figura 2. Superfície de Energia Potencial genérica diferenciando confórmeros e rotâmeros. Baseado na ref. [23].	1
Figura 3. Fluxo do CRENSO definido por Grimme e colaboradores. Extraído da ref. [14].	4
Figura 4. Estruturas moleculares dos compostos químicos sintéticos das diversas classes de substâncias investigadas neste trabalho.	3
Figura 5. Estruturas dos plastificantes DIN e DET nas formas reduzidas para diminuir o número de conformações acessíveis ao CRENSO.	4
Figura 6. Sobreposição das estruturas mais estáveis na fase de vapor e na fase líquida dos compostos usados para validar o CRENSO com modelo de solvatação SMD. Os confórmeros da fase líquida são os na cor azul claro e em fase gás é representado de forma colorida.	7
Figura 7. Exemplos de comandos usados no CREST para gerar os conjuntos conformacionais, respectivamente na fase gás e na fase líquida. O programa interpreta estruturas com extensão mol, xyz, sdf e coord.	7
Figura 8. Resultado final gerado pelo CENSO, que inclui a energia livre em cada faixa de temperatura selecionada, resultados em Hartree.	9
Figura 9. Sobreposição dos confórmeros de menor energia dos seis compostos selecionados em fase vapor (colorido) e fase líquida (azul claro) empregando o modelo contínuo SMD, obtidos ao final do workflow CRENSO.	3
Figura 10. Logaritmo da pressão de vapor das moléculas da classe dos biocidas, plastificantes e fármacos calculada usando o workflow CRENSO na temperatura de 298,15K e pressão de 1 atm. Tons mais escuros (claros) representam os valores de $\log P$ (de $\log P_{corr}$) sem (com) as correções $avG_{corr}(T)$.	5
Figura 11. Comparação entre as pressões de vapor em escala logarítmica dos compostos selecionados neste trabalho, obtidas usando a média das energias do conjunto (linha de base) e os valores das pressões de vapor de confórmeros selecionados de forma aleatória ($\log P(R)$) e apenas a conformação mais estável ($\log P(L)$)	6
Figura 12. Relação entre a pressão de vapor e a massa dos compostos selecionados neste trabalho obtidos pelo método CRENSO	7

LISTA DE TABELAS

Tabela 1. Tipo de cálculo e nível de teoria utilizado nas diversas partes do algoritmo CENSO.....	2
Tabela 2. Moléculas escolhidas para realizar as simulações CRENSO para previsão da.....	1
Tabela 3. Resultados teóricos para estimativa dos parâmetros da equação de Clausius-Mossotti calculados no nível CAM-B3LYP/aug-cc-pVDZ. O solvente escolhido para conduzir as simulações do workflow CRENSO será aquele com constante dielétrica mais próxima possível.	5
Tabela 4. Comparação entre a pressão de vapor experimental e calculado dos compostos empregados para validação dos parâmetros selecionados no CRENSO.	6
Tabela 5. Energias livre de Gibbs média do ensemble configuracional das moléculas de interesse industrial calculadas com o workflow CRENSO na temperatura de 298,15K e pressão de 1atm.	1
Tabela 6. Número de conformémeros encontrados nas subetapas do workflow CRENSO das moléculas selecionadas em fase vapor e líquido.	2
Tabela 7. Pressão de vapor em Pascal, $\log P$ e desvio padrão do $\log P$ de uma triplicata de simulações CRENSO das moléculas selecionadas neste trabalho, na temperatura de 298,15K e pressão de 1atm...	4
Tabela 8. Pressão de vapor em escala logarítmica dos compostos selecionados neste trabalho obtidos pelo método CRENSO que envolve médias de conformações e os respectivos valores encontrados empregando apenas a conformação mais estável ($\log PL$) e uma conformação aleatória ($\log P(R)$) na temperatura de 298,15K e pressão de 1 atm.	6
Tabela A 1. Solventes disponíveis nos programas gratuitos ORCA e XTb empregando os modelos PCM SM e ALPB. indica como o solvente é declarado no input do CRENSO.	1

LISTA DE ABREVIATURAS E SIGLAS

xTB	<i>Semiempirical Extended Tight-Binding Program Package</i>
CREST	<i>Conformer-Rotamer Ensemble Sampling Tool</i>
CENSO	<i>Commandline ENergetic SOrting</i>
CRENSO	Fluxo de trabalho que engloba o CREST e o CENSO
GFN-FF	<i>GFN Force Field</i>
GFN2-xTB	<i>GFN Semiempirical Extended Tight-Binding</i>
DFT	<i>Density Functional Theory</i>
QM	Mecânica Quântica
LFER	<i>Linear Free-Energy Relationships</i>
SPARC	<i>Sparc Performs Automated Reasoning in Chemistry</i>
ALPB	<i>Analytical Linearized Poisson–Boltzmann</i>
SMD	<i>Solvation Model based on Density</i>
PES	<i>Potencial Energy Surface</i>
RMSD	<i>Root Mean Square Deviation</i>
DFA	<i>Density Functional Approximations</i>
SPH	<i>Single-Point Hessian</i>
COSMO-RS	<i>Conductor-like Screening Model for Realistic Solvents</i>

Sumário

Introdução	1
Revisão Bibliográfica	1
CREST	1
CENSO	2
Workflow CRENSO	3
Objetivos	1
Objetivo Geral	1
Objetivos Específicos	1
Metodologia	1
Programas Utilizados	1
Sistemas Moleculares Escolhidos	1
Seleção dos solventes	4
Comparações com compostos da literatura	5
Parâmetros selecionados no workflow CRENSO	6
Resultados e Discussões	1
Conclusões e Perspectivas	1
Referências Bibliográficas	1
Apêndice A	1
Solventes disponíveis nos modelos ALPB e SMD no CREST/ORCA	1

Introdução

A sociedade moderna produz virtualmente uma infinidade de produtos químicos orgânicos naturais e sintéticos empregados diariamente no estilo de vida contemporâneo. Considere as várias maneiras pelas quais esses produtos químicos podem ser liberados intencionalmente ou acidentalmente no ambiente por nossas atividades humanas. A medida que esses produtos químicos se acumulam na natureza seja na atmosfera, em águas superficiais e sedimentos, assim como em solos e aquíferos, devemos nos perguntar: quais podem ser os diversos efeitos sobre organismos aquáticos e terrestres, ecossistemas inteiros e a saúde humana?

Cerca de 350 mil produtos químicos sintéticos foram produzidos e empregados em larga escala nos últimos 30-40 anos.¹ Globalmente cerca de 4000 substâncias farmacêuticas, para uso humano ou veterinário, 1800 pesticidas foram registrados sob regulamentação.² No entanto, como ficou cada vez mais evidente nas últimas décadas, o uso de todos esses produtos químicos tem seus aspectos negativos. Muitos foram detectados em todo o ambiente, especialmente em corpos d'água.³⁻⁷

Como exemplo histórico, o inseticida organoclorado DDT (Dicloro difenil tricloroetano) foi largamente utilizado no Brasil desde 1945, para fins agrícolas e no controle de vetores. Em 1986 foi proibida a sua utilização na agricultura e em 1997 esta proibição se estendeu também ao uso no controle de vetores. No entanto, a presença de DDT no meio urbano, solos florestais, leite materno e biota aquática ainda é comum especialmente na Amazônia.⁸ Muitos produtos químicos como DDT e PCB's (Bifenil policlorinados) são bastante persistentes no ambiente (ou seja, resistentes à degradação abiótica ou biológica) e podem se acumular nos corpos de organismos, incluindo humanos,^{9,10} que podem representar uma ameaça para ecossistemas, saúde humana ou recursos.

Exceto em locais altamente contaminados, muitos desses produtos químicos estão comumente presentes em concentrações relativamente baixas no ambiente (por exemplo, $ng \cdot L^{-1}$ a $\mu g \cdot L^{-1}$) e são referidos como "micropoluentes orgânicos"^{11,12} e são um dos principais problemas ambientais enfrentados pela humanidade.

Com base em suas propriedades físico-químicas e em sua persistência conhecida, poderia-se antecipar que eles se acumulariam no ambiente e em organismos, incluindo humanos. Um aspecto físico-químico importante na racionalização sobre se um determinado composto orgânico é termodinamicamente favorecido estar na fase gasosa ou em alguma fase

condensada é a pressão de vapor, que fornece informações quantitativas diretas sobre as forças atrativas entre as moléculas do composto em sua fase líquida ou sólida. Além disso, a pressão de vapor não apenas descreve a partição de equilíbrio entre a fase gasosa e uma fase condensada, mas também ajuda a quantificar a taxa de evaporação de um composto de sua fase pura ou de uma mistura.¹³

Desta forma, a pressão de vapor em uma determinada temperatura é de importância central em descrever o comportamento de substâncias no ambiente. O conhecimento desta propriedade auxilia no entendimento sobre as transições de fase envolvendo a fase gasosa, pode ser usada para determinar a constante de Henry, e também compreender a adsorção de substâncias em superfícies.¹⁴

As condições termodinâmicas dos compostos em que iremos realizar as simulações, favorecem a sua existência como sólidos. Desta forma, realizar o cálculo da pressão de vapor destes sólidos, impõe uma dificuldade teórica. Entretanto, é ainda interessante empregar métodos computacionais menos exigentes do ponto de vista de tempo de cálculo. Modelar efeitos do ambiente líquido é atualmente uma tarefa relativamente simplificada com acurácia bastante razoável dependendo da aplicação, quando se empregam os modelos implícitos de solvatação – os chamados modelos contínuos polarizáveis (*Polarizable Continuum Models*, PCM).¹⁵ Os modelos implícitos acoplam a descrição quanto-mecânica do sistema químico de interesse com uma descrição contínua do ambiente em que o “soluto” se encontra. Como mencionamos acima, as moléculas em que este estudo será realizado, estão no estado sólido nas condições termodinâmicas usuais (1 atm e 298,15K), de modo, que o “ambiente químico” ao redor das espécies de interesse é na verdade formado por moléculas da mesma espécie. Ainda assim, podemos nos beneficiar do baixo custo computacional dos modelos PCM para introduzir ainda que aproximadamente, efeitos de fase condensada nas estruturas moleculares e em última instância, na pressão de vapor. Embora tal estratégia possa parecer excessivamente “grosseira”, termodinamicamente, tal estado é factível do ponto de vista experimental. Em um diagrama de fases de uma substância normal pura, as curvas de coexistência se encontram em um único ponto, o ponto triplo, como mostrado na Figura 1.

Se extrapolarmos a curva de coexistência líquido-gás, para regiões abaixo do ponto triplo, temos o estado termodinâmico metaestável *líquido sub-resfriado* (linha tracejada na Fig. 1). A pressão de vapor do líquido sub-resfriado de uma dada substância pode ter importância prática, pois nos informa sobre as interações do composto na sua forma líquida pura em uma temperatura em que o composto é na realidade um sólido. Este conhecimento das propriedades

do líquido sub-resfriado nos ajuda a entender e quantificar interações moleculares em ambientes em que a molécula existe em estado líquido (dissolvido em água, por exemplo) embora fossem sólidos se estivessem puros.¹³ Em suma, podemos trocar o problema de modelar o estado sólido da substância com sua fase líquida ou vapor, estudando um estado metaestável, pois nestas condições o tratamento do meio pode ser realizado através dos métodos PCM de baixo custo para nossas simulações.

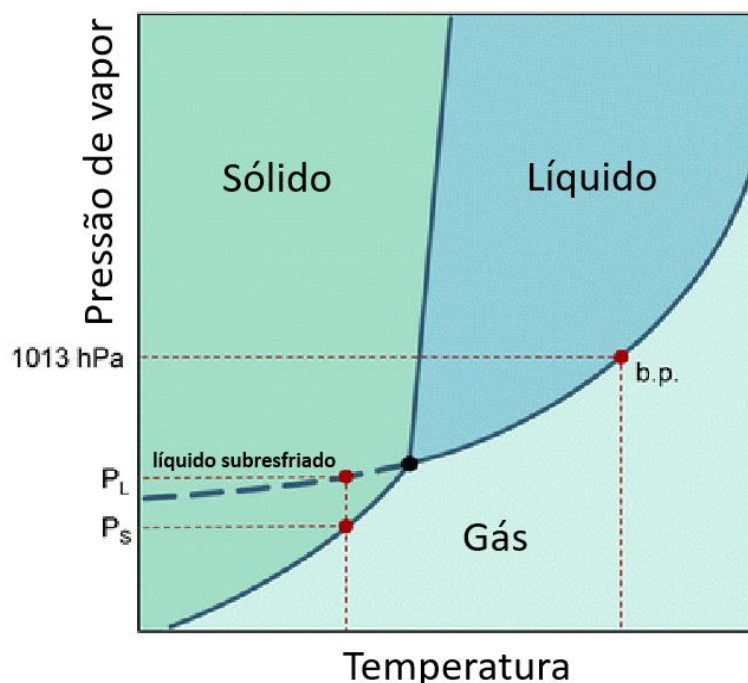


Figura 1. Diagrama de fases genérico de uma substância normal. O ponto normal de ebulição b.p. é destacado bem como a curva de coexistência sólido-vapor (P_s) e as condições de metaestabilidade de líquido sub-resfriado (linha tracejada). Extraído da referência [14].

Para muitos compostos a pressão de vapor pode ser medida com precisão, contudo, a determinação é metodologicamente complexa, e a menos que o método seja cromatográfico,^{16,17} são necessárias substâncias altamente purificadas, e em geral, quanto mais baixa a pressão de vapor, mais sofisticada e propensa a erros se torna a medição. Portanto, é comum os experimentos serem realizados em temperaturas mais altas e o resultado é então extrapolado para a temperatura desejada usando a equação de Antoine,¹⁸ esse procedimento pode ser conveniente desde que não haja transição de fase entre a faixa de medição e a temperatura extrapolada.

Por conta das dificuldades e incertezas experimentais, existe a necessidade de estimar a pressão de vapor usando outros métodos mais facilmente disponíveis. Os modelos teóricos mais

simples se baseiam na regra de Pictet-Trouton, que afirma que a entropia de vaporização atinge um valor constante no ponto de ebulição, $\Delta_{vap}h/T_{vap} \approx 10,5R$. A partir da entalpia de vaporização calculada no ponto de ebulição e com ajuda de fatores de correção de Fishtine,¹⁹ a pressão de vapor de um composto pode ser associada ao ponto de ebulição.²⁰ Uma abordagem alternativa é baseada na estrutura química e calcula o excesso de energia livre de Gibbs de um composto em um líquido a partir das contribuições de seus grupos característicos.^{21,22} Outra possibilidade é o método *Linear Free-Energy Relationship*, LFER, para calcular a pressão de vapor do composto. O método LFER é um conjunto de ferramentas para prever propriedades físicas e termodinâmicas e coeficientes de partição de compostos orgânicos, baseado na determinação de alguns descritores. Porém, o método necessita do ponto de fusão para calcular a pressão de vapor. O algoritmo implementado no SPARC é amplamente usado para prever pressões de vapor, contudo, os valores calculados com SPARC para substâncias com baixa pressão de vapor (pressões abaixo de $10^{-2}Pa$), são pequenos e isso pode ser atribuído à falta de dados na calibração para baixas pressões ou ao fato de que as interações intermoleculares mais significativas envolvidas não podem ser compreendidas com puro empirismo. Tanto o LFER quanto o SPARC usam notação SMILES para construção das moléculas, que descreve apenas a estrutura básica das moléculas, e não conformeros específicos.

Os métodos da mecânica quântica permitem o cálculo das propriedades termodinâmicas fundamentais das moléculas e calculam as pressões de vapor a partir da diferença das energias livres de Gibbs para a fase condensada e a fase gasosa de uma molécula. Embora isso seja simples para moléculas pequenas, que geralmente são bastante rígidas ou apresentam poucos graus de liberdade, surgem questões em aberto para moléculas maiores, normalmente bastante flexíveis. Um ponto importante a se considerar, é a dependência da pressão de vapor no espaço conformacional disponível. A composição dos conjuntos conformacionais e as estruturas tridimensionais dos conformeros são altamente dependentes do ambiente e podem diferir amplamente entre a fase líquida e a fase gasosa, e negligenciar as mudanças conformacionais do estado condensado para o estado gasoso em procedimentos teóricos, pode levar a erros significativos na diferença de energia livre de Gibbs calculada.¹⁴

O grupo de química computacional liderado pelo professor Stefan Grimme explora de forma automatizada e abrangente o espaço conformacional com o programa CREST (*Conformer-Rotamer Ensemble Sampling Tool*)²³ para encontrar diversos conformeros de baixa energia em cada fase, seguido pela reclassificação dos conjuntos conformacionais em níveis mais elevados de teoria DFT usando um algoritmo de classificação e otimização energética, o CENSO

(*Commandline Energetic Sorting*).²⁴ A combinação desses métodos foi chamada de CRENSO,¹⁴ e esse workflow pode ser usado para calcular a pressão de vapor de compostos orgânicos com baixas pressões de vapor com uma boa precisão. O fluxo original da ferramenta CRENSO emprega como modelo de solvatação contínua o método COSMO-RS,²⁵ que está disponível no pacote COSMOTerm,²⁶ que é um software pago. Desta forma, neste trabalho propomos uma alteração no método original para se empregar um modelo de solvatação alternativo, o modelo de solvatação SMD²⁷ (*universal Solvation Model based on solute electron Density*), que está disponível gratuitamente no pacote de livre distribuição ORCA.²⁸

Desta forma, este trabalho busca validar e empregar o método CRENSO propondo uma alteração ao workflow original para estimar a pressão de vapor de algumas substâncias sintéticas de classes de alta relevância para a sociedade moderna: plastificantes, biocidas e substâncias farmacêuticas sintetizadas entre 2022 e 2023, em que dados sobre pressão de vapor experimentais não estão disponíveis, em grande parte, devido a sua baixa volatilidade em condições usuais do ambiente. Esta alteração ambiciosa propicia uma maneira de estimar esta propriedade de alto interesse considerando o problema atual dos contaminantes emergentes na natureza empregando códigos de acesso livre.

Revisão Bibliográfica

CREST

Experimentalmente, muitas observáveis físicas são obtidas como uma média temporal sobre as diferentes estruturas de baixa energia de uma molécula. O conjunto termicamente acessível de estruturas de energia mínima geralmente consiste em conformeros e rotâmeros, ou seja, isômeros espaciais de uma molécula que apresentam conectividade e topologia covalentes idênticas.²³ Seguindo a hipótese ergódica, substituir a média temporal por uma média sobre esse chamado conjunto de conformadores/rotâmeros (CRE) teoricamente fornecerá as mesmas observáveis, transformando a complicada média temporal em um problema de amostragem gerenciável. Portanto, na química computacional, pode ser importante realizar cálculos na conformação preferida ou em um conjunto de conformações.

Por definição, confórmeros são os mínimos em uma superfície de energia potencial (SEP) para uma dada topologia molecular como mostrado na Figura 2. Devido à rotação em torno de ligações químicas covalentes, surgem os chamados rotâmeros. Rotâmeros têm mínimos degenerados na SEP e, portanto, são indistinguíveis. Essas definições são representadas abaixo na Figura 2.

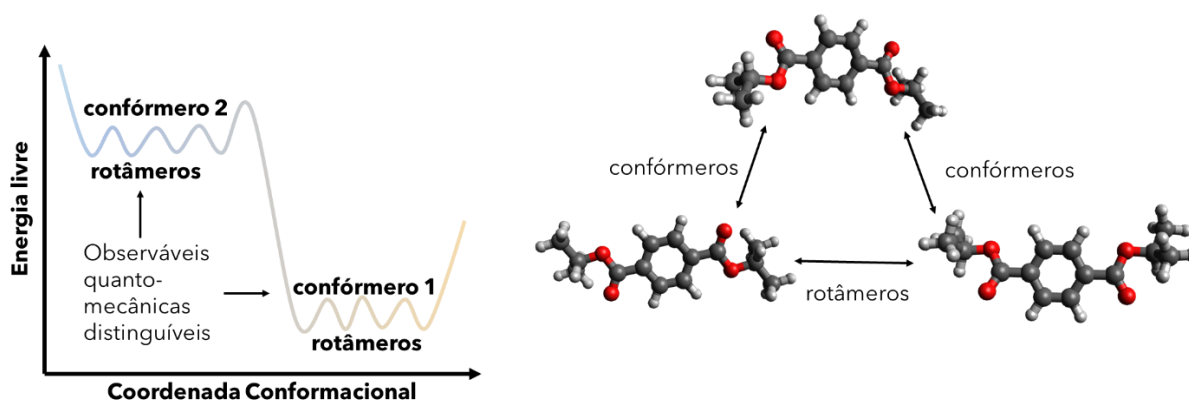


Figura 2. Superfície de Energia Potencial genérica diferenciando confórmeros e rotâmeros. Baseado na ref. [23].

O CREST utiliza três descritores para distinguir entre confórmeros e rotâmeros: (i) a energia total, (ii) as constantes rotacionais da molécula e (iii) o desvio médio quadrático das

posições atômicas RMSD entre duas moléculas. Se duas estruturas têm energias completamente diferentes, são, pela definição acima da SEP, diferentes conforméromos. Caso contrário, se tiverem a mesma energia dentro de um limite escolhido, as estruturas podem ser dois conforméromos diferentes com energias semelhantes, dois rotâmeros da mesma molécula ou uma duplicata do mesmo rotâmero. No primeiro caso (dois conforméromos com energias semelhantes), o RMSD será alto e a constante rotacional será diferente. Para dois rotâmeros, o RMSD será alto, mas a constante rotacional será a mesma (dentro de um limite escolhido). Apenas para duplicatas do mesmo rotâmero que precisam ser excluídas, a energia e as constantes rotacionais serão idênticas, enquanto o RMSD será baixo.

CENSO

A etapa CENSO, *Commandline Energetic Sorting* consistem em um algoritmo de classificação para avaliar eficientemente um conjunto de estruturas. O conjunto de conformações originado da etapa CREST serve como entrada (também funciona em uma única estrutura) pode ser ranqueada por valores de energia livre calculadas em nível DFT e/ou as geometria moleculares podem ser otimizadas usando DFT. A classificação é baseada em “janelas” de energia (livre) ou energias de corte, dentro dos quais todos os conforméromos/estruturas são considerados. Através de uma escolha apropriada de valores de corte as seguintes “propriedades”/resultados podem ser obtidos: (i) conforméromo de mais baixa energia e energia livre correspondente, (ii) conjunto conformacional totalmente otimizado e determinação da energia livre do conjunto (ensemble) e (iii) população de Boltzmann do ensemble em uma dada temperatura.²⁴

O algoritmo CENSO de classificação de estruturas segue algumas etapas sequenciais que listamos abaixo na Tabela 1.

Tabela 1. Tipo de cálculo e nível de teoria utilizado nas diversas partes do algoritmo CENSO.

Parte 0 – Triagem “Barata” do Ensemble
Cálculos single point no nível de teoria B97-D3/def2-SV(P) // GFN2-xTB no programa ORCA. Energia de corte: $G_{thr}(0) = 4kcal \cdot mol^{-1}$
Parte 1 – Triagem do Ensemble
Cálculo single point no nível de teoria r2scan-3c + GmRRHO(GFN2) // GFNn-xTB no programa ORCA.

Energia de corte: $G_{thr}(1) = 3,5kcal \cdot mol^{-1}$
Parte 2 – Otimização do Ensemble
Cálculo de otimização geométrica no nível de teoria r2scan-3c + GmRRHO(GFN2) // r2scan-3c Energia de corte: $G_{thr}(2) = 2,5kcal \cdot mol^{-1}$ Nesta parte, somente as conformações que somadas resultem em 99% da população de Boltzmann são levadas adiante.
Parte 3 - Refinamento
Cálculo single point no nível de teoria pw6b95-d4/def2-TZVPD + GmRRHO(GFN2) // r2scan-3c Nesta parte, somente as conformações que somadas resultem em 99% da população de Boltzmann são consideradas para média final das propriedades.

Na Tabela 1, a sigla GmRRHO representa uma modificação nas correções térmicas vibracionais para os modos normais de baixa frequência dentro da aproximação do rotor rígido e oscilador harmônico (mRRHO) na energia livre de Gibbs. A modificação trata os modos de baixa frequência que contribuem amplamente para a entropia (e são numericamente problemáticos) e é utilizado como padrão no grupo do Prof. Grimme.²⁹

Workflow CRENSO

No workflow original,¹⁴ também adotado neste trabalho, a ferramenta CREST, utiliza os métodos GFN-FF e GFN2-xTB³⁰ para explorar extensivamente o espaço químico e criar um conjunto conformacional inicial. Então, com o CENSO, usando métodos DFT de nível superior que foram cuidadosamente selecionados, esse conjunto inicial é filtrado energeticamente, e os conformeros mais elevados são classificados para a próxima etapa. Em todas as etapas, são usados métodos apropriados de solvatação para os cálculos das fases condensadas, que são selecionados com base na respectiva precisão e esforço computacional do método subjacente. As contribuições térmicas são incluídas usando a abordagem Hessiana de ponto único.³¹ O Fluxograma CRENSO original pode ser visto na Figura 3.

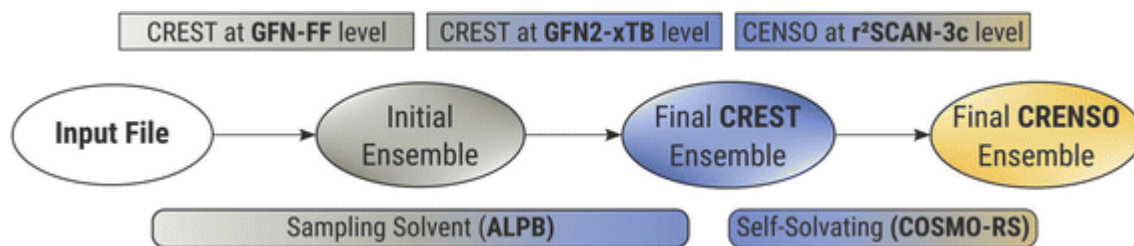


Figura 3. Fluxo do CRENSO definido por Grimme e colaboradores. Extraído da ref. [14].

A energia livre de uma molécula em cada fase pode então ser determinada a partir de uma média ponderada sobre as energias livres dos conformémeros no conjunto selecionado ao final do CRENSO, assim obtemos a energia livre para o estado líquido e a energia livre para o estado de vapor, ambos determinados na condição padrão de 1 mol/L. Segundo Ben-Naim,³² a pressão de vapor (P) pode ser obtida a partir do equilíbrio entre o estado líquido e de vapor por meio da equação,

$$P = \frac{RT\rho_{liq}}{MW} e^{(g_{liq}^* - g_{vap}^*)/RT}, \quad (1)$$

em que, ρ_{liq} é a densidade de massa do líquido, MW é a massa molar e g_{liq}^* e g_{vap}^* , representam, respectivamente, a energia livre de Gibbs molar do líquido e vapor puro. Na equação (1), as energias livres de Gibbs molar serão oriundos da média do conjunto conformacional de cada molécula. Embora o fluxo de trabalho original use o modelo de solvatação COSMO-RS,²⁵ que está disponível no pacote COSMOTerm, que é pago, neste trabalho optamos por usar o modelo de solvatação SMD na etapa final do CRENSO (ver Fig. 3),²⁷ que está disponível gratuitamente no pacote ORCA.²⁸

Na etapa CREST do CRENSO, emprega-se o modelo de solvatação implícita linearizada analítica de Poisson Boltzmann (ALPB)^{33,34} usado para a geração inicial de conjuntos de conformémeros e rotâmeros. Porém, em nosso procedimento o ALPB é parametrizado apenas para um conjunto fixo de solventes comuns, e um tratamento de auto-solvatação não é viável. No entanto, negligenciar totalmente a influência do solvente ao criar o conjunto conformacional pode levar a erros significativos na previsão da pressão de vapor.

Portanto, nas primeiras etapas dos cálculos, utilizamos solventes, que foram parametrizados para o modelo de solvatação ALPB e possuem constante dielétrica semelhante ao composto alvo. Na Tabela A1, na seção de apêndices, apresentamos os solventes atualmente parametrizados para empregar o ALPB e SMD no workflow CRENSO. Com esta aproximação, ainda somos capazes de criar diferentes conjuntos de conformações para a fase condensada e

gasosa nos estágios iniciais do fluxo de trabalho, sem a necessidade de parametrizar explicitamente o modelo de solvatação ALPB e SMD para cada novo composto.

Objetivos

Objetivo Geral

Validar e aplicar uma alteração no tratamento da fase líquida no fluxo de trabalho CRENSO proposto por Grimme e colaboradores para calcular de maneira computacionalmente barata a pressão de vapor de compostos orgânicos sintéticos da classe dos fármacos, biocidas e plastificantes.

Objetivos Específicos

Validar a alteração no workflow CRENSO substituindo o modelo COSMO-RS pelo SMD disponível em pacotes de química computacional gratuitos como o ORCA, calculando a pressão de vapor de algumas moléculas com dados experimentais, também investigados pelos autores na publicação original.¹⁴

Calcular as pressões de vapor de 6 compostos de interesse químico moderno que foram recentemente publicados, que estão divididos em 3 categorias, plastificantes, biocidas e fármacos utilizando o workflow CRENSO modificado.

Temos também como objetivo, realizar a construção de um código na linguagem Python para facilitar os cálculos e automatizar todo o processo para que futuros estudantes possam empregá-lo de maneira simplificada.

Metodologia

Programas Utilizados

A primeira parte envolve a instalação dos programas necessários, primeiramente temos que instalar o xTB³⁵ o CREST e o CENSO,³⁶ também precisamos instalar o ORCA,³⁷ pois ele é usado para realizar os cálculos de nível DFT, e o modelo de solvatação usado está disponível no ORCA. Para o cálculo das constantes dielétricas dos compostos químicos selecionados utilizamos o software Gaussian16,³⁸ disponíveis nos clusters da Divisão de Físico-Química do Instituto de Química da UnB.

Sistemas Moleculares Escolhidos

Nesta seção, abordaremos sobre os métodos computacionais adotados para realização deste trabalho. Biocidas, Fármacos e Plastificantes são três classes de compostos químicos produzidos em larga escala pela indústria química moderna. Desta forma, selecionamos duas estruturas de cada classe por serem, em geral, substâncias que apresentam baixíssima volatilidade, de modo que propriedades físico-químicas como a pressão de vapor e a entalpia de vaporização não podem ser determinadas facilmente com grande acurácia.

Estes seis compostos estão destacados na Tabela 2 e suas estruturas moleculares reunidas na Figura 4. Na Tabela 2, sugerimos siglas para se referir a cada estrutura de maneira mais breve.

Tabela 2. Moléculas escolhidas para realizar as simulações CRENSO para previsão da pressão de vapor.

Composto	Sigla	Classe
DINCH	DIN	Plastificante
Di-2-ethylhexil terephthalate	DET	Plastificante
Transketolase-IN-3	TKT	Biocida
Mefentrifluconazole	MEF	Biocida

Abiraterone acetate	ABA	Fármaco
Palovarotene	PAV	Fármaco

Na classe dos plastificantes, as estruturas DIN e DET, Figura 4, apresentam ramificações com cadeias alifáticas relativamente longas. Esta característica química de suas estruturas confere aos dois compostos um elevado número de graus de liberdade e muita flexibilidade conformacional. Na prática, o espaço conformacional acessível para os dois compostos mostrou-se excessivamente elevado para que pudéssemos executar as simulações nas duas estruturas dentro do prazo previsto para conclusão deste Trabalho de Conclusão de Curso. Desta forma, propomos estudar uma forma estrutural truncada dos dois plastificantes DET e DIN reduzindo o número de átomos de carbono nas cadeias alifáticas. As estruturas das versões truncadas ou simplificadas da DET são apresentadas na Figura 5, e todos os resultados que apresentaremos sobre os dois compostos referem-se a estas estruturas.

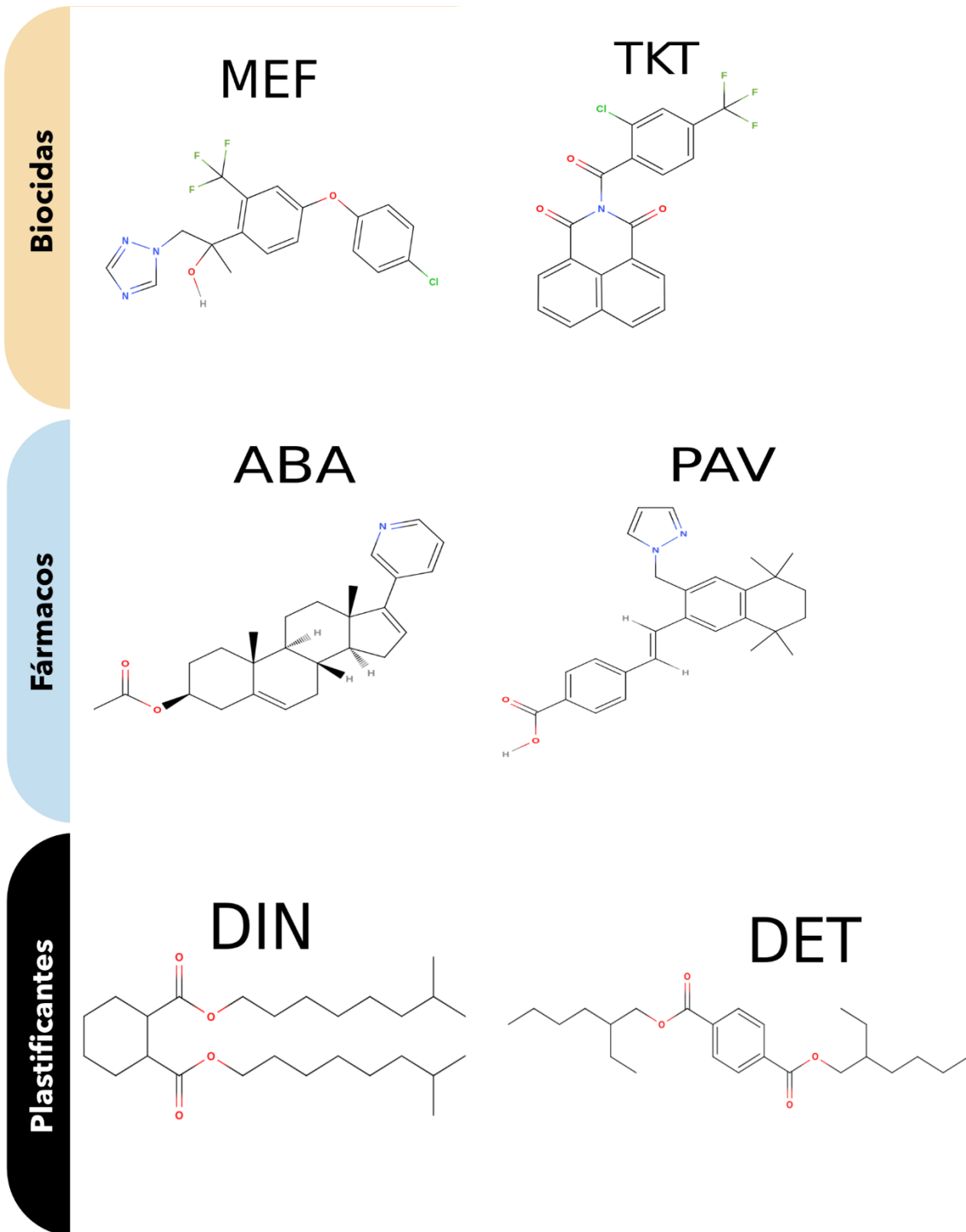


Figura 4. Estruturas moleculares dos compostos químicos sintéticos das diversas classes de substâncias investigadas neste trabalho.

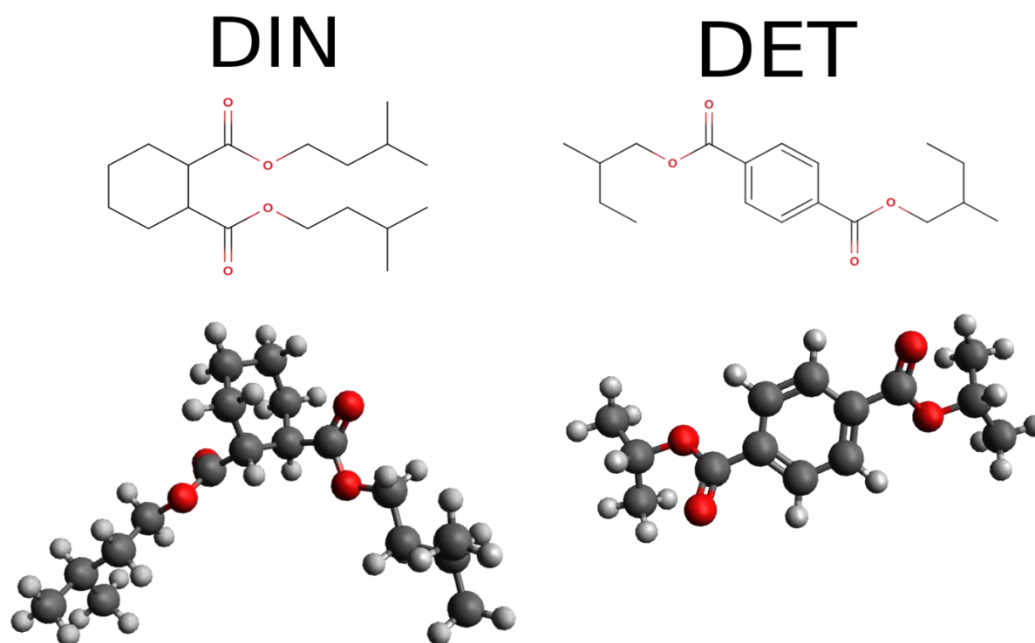


Figura 5. Estruturas dos plastificantes DIN e DET nas formas reduzidas para diminuir o número de conformações acessíveis ao CRENSO.

Seleção dos solventes

Para determinar a energia livre de Gibbs em fase líquida dos 6 compostos, primeiramente precisamos definir qual solvente será selecionado para cada composto baseado na constante dielétrica. Como os compostos da Tabela 3 não apresentam valores de constante dielétricas na literatura, optamos por determiná-los de maneira aproximada através de cálculos quanto-mecânicos baseados na Teoria do Funcional de Densidade (DFT). Para determinação da constante dielétrica estática, empregamos a equação de Clausius-Mossotti,

$$\frac{\varepsilon - 1}{\varepsilon + 2} = \frac{4\pi N_A \alpha}{3V_m} \quad (2)$$

em que ε é a constante dielétrica estática, N_A é a constante de Avogadro, α é a polarizabilidade estática e V_m é o volume molecular. O volume molecular foi calculado com a Keyword *polar* no programa Gaussian16³⁸ e a polarizabilidade foi calculada no nível CAM-B3LYP/aug-cc-pVDZ que mostrou desvio inferior a 5% em relação aos valores experimentais de moléculas orgânicas.³⁹

Na Tabela 3 apresentamos os cálculos obtidos para estimar a constante dielétrica dos seis compostos de interesse no Gaussian16. As constantes dielétricas ficaram entre 2 e 3,34,

assim os solventes escolhidos foram Hexadecano no caso do composto DIN e Dissulfeto de Carbono (cs2) para os demais casos (ver Tab. A1 para comparar com os demais solventes disponíveis no ORCA).

Tabela 3. Resultados teóricos para estimativa dos parâmetros da equação de Clausius-Mossotti calculados no nível CAM-B3LYP/aug-cc-pVDZ. O solvente escolhido para conduzir as simulações do workflow CRENSO será aquele com constante dielétrica mais próxima possível.

Molécula	$\alpha \times 10^{-29} (m^3)$	$V_m (cm^3 \cdot mol^{-1})$	ϵ	Solvente (xtb)
ABA	4,529782	287,025	2,98	cs2
DET	4,682266	318,467	2,77	cs2
DIN	4,908252	472,713	2,06	Hexadecano
MEF	3,711938	234,110	3,00	cs2
PAV	5,485737	315,631	3,34	cs2
TKT	3,971561	253,449	2,96	cs2

Comparações com compostos da literatura

O artigo usado como referência para esse trabalho valida o fluxo de trabalho CRENSO, realizando comparações entre a pressão de vapor calculada de diversos compostos com seus respectivos valores na literatura. Como há uma mudança no modelo de solvatação desse trabalho, é necessário realizar novamente uma validação, para verificar se o cálculo usando o modelo de solvatação SMD é capaz de apresentar resultados satisfatórios quando comparado ao modelo de solvatação COSMO-RS e com os valores tabelados na literatura.

Foram realizados cálculos com três das mesmas moléculas que o artigo original de Stahn e colaboradores⁴⁰ usou para realizar comparações e assim validar o fluxo de trabalho. Selecionamos o glicerol, n-decano e naftaleno como sistemas moleculares para validação dos parâmetros selecionados do fluxo CRENSO. Na Tabela 4, são apresentados os valores da pressão de vapor obtidos a partir das simulações na temperatura de 298,15K e pressão de 1 atm. No glicerol, empregou-se o solvente implícito etanol para incorporar pelo menos parcialmente o efeito da fase líquida neste composto (na Tab. A1 é possível verificar que não há parâmetros ALPB e SMD para o glicerol, n-decano e naftaleno). Este solvente foi selecionado em virtude

de sua constante dielétrica ($\epsilon = 24.3$) não ser tão diferente que a do próprio glicerol ($\epsilon = 20$). Similarmente, para o n-decano e naftaleno foram conduzidos usando solvatação implícita em n-hexadecano e benzeno, respectivamente para capturar a energia livre de Gibbs da fase líquida de maneira aproximada. Observa-se pela Tabela 4 que o modelo SMD retornou resultados bastante razoáveis da pressão de vapor (na escala logarítmica) para o glicerol n-decano e o naftaleno.

De acordo com os resultados obtidos e apresentados na Tabela 4, a alteração do modelo de solvatação COSMO-RS para SMD na etapa CENSO do workflow CREST retornou resultados inferiores frente aos resultados experimentais nos casos dos líquidos n-decano e naftaleno. Porém, no caso do glicerol, de maneira interessante, o SMD retornou resultados melhores para a pressão de vapor. Estes resultados nos encorajaram a manter o protocolo alterado do CRENSO para prever, ainda que em caráter qualitativo, a pressão de vapor de outras substâncias químicas de interesse industrial em que medidas experimentais não são tão simples de serem realizadas.

Tabela 4. Comparação entre a pressão de vapor experimental e calculado dos compostos empregados para validação dos parâmetros selecionados no CRENSO.

Molécula	$\log P$ CRENSO (SMD) ^a	$\log P$ CRENSO (COSMO-RS) ^b	$\log P$ (Exptl.) ^c
Glicerol	-1,72	-0,69	-1,6 ^{d,41}
n-Decano	2,48	2,28	2,26 ⁴²
Naftaleno	1,32	1,48	1,57 ⁴³

^aEste Trabalho. ^bCRENSO original.⁴⁰ ^cValor experimental.

Parâmetros selecionados no workflow CRENSO

A primeira etapa do fluxo de trabalho do CRENSO é usar o CREST para criar conjuntos conformacionais, que são variações na estrutura de uma molécula selecionando apenas os confórmeros e rotâmeros mais estáveis como indicado na Figura 6. O CREST combina cálculos de mecânica clássica utilizando o campo de força parcialmente polarizável GFN-FF e o método semi-empírico GFN2-xTB com simulações de metadinâmica baseadas no desvio médio quadrático das geometrias para encontrar os confórmeros e rotâmeros:

$$RMSD = \sqrt{\frac{1}{n} \sum_{i=1}^n [(x_{i1} - x_{i2})^2 + (y_{i1} - y_{i2})^2 + (z_{i1} - z_{i2})^2]} \quad (3)$$

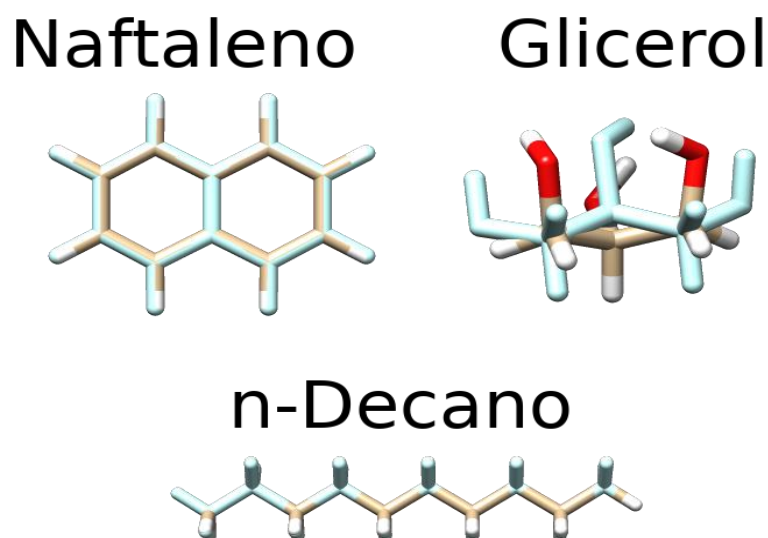


Figura 6. Sobreposição das estruturas mais estáveis na fase de vapor e na fase líquida dos compostos usados para validar o CRENSO com modelo de solvatação SMD. Os conformémeros da fase líquida são os na cor azul claro e em fase gás é representado de forma colorida.

O conjunto conformacional é gerado a partir de um único arquivo contendo a estrutura da molécula de interesse, também podemos indicar um solvente, que será usado para gerar o conjunto conformacional como mostrado nos comandos de terminal ilustrados na Figura 7.

```
crest gly.mol --gfn2//gfnff -T 6 > crest.out
```

```
crest gly.mol --alpb methanol --gfn2//gfnff -T 6 > crest.out
```

Figura 7. Exemplos de comandos usados no CREST para gerar os conjuntos conformacionais, respectivamente na fase gás e na fase líquida. O programa interpreta estruturas com extensão mol, xyz, sdf e coord.

Durante a geração das estruturas, o CREST, nas etapas finais realiza o cruzamento de estruturas, nessa etapa o RMSD é calculado para estruturas que possuem energias semelhantes

dentro de um limiar, assim evitando que o programa retorne conformeros e rotâmeros duplicados. Em seguida, o ensemble configuracional encontrado pelo CREST pode ser refinado e as energias livres determinadas para o ensemble usando o CENSO. A seguir indicamos os métodos empregados em cada etapa do algoritmo CENSO utilizado no CRENSO.

Parte 0 - Pré-seleção barata:

Nessa parte os conformeros com energia eletrônica muito alta são removidos. Aqui, a descrição da energia eletrônica é melhorada em relação à energia inicial, realizando cálculos de energia (single point) B97-D3/def2-SV(P)+gcp que são considerados bastante rápidos. O termo gcp representa a correção da base *geometric counterpoise correction* que busca mitigar o erro do conjunto de funções de base nos cálculos DFT empregando bases locais melhorando a geometria molecular.⁴⁴

Se a molécula estiver em fase de solução, a solvatação foi tratada no nível de GFN2-xTB[ALPB]. A classificação é baseada em uma janela energética em relação ao conformero mais estável. Seguindo a sugestão original, mantivemos a janela em 4 kcal/mol.

Parte 1 - Pré-seleção:

Nessa parte, é feita a seleção dos conformeros que possuem energia elevada com base na melhoria da energia eletrônica e nas contribuições de solvatação, essa classificação é baseada na energia livre total, incluindo contribuições termostáticas.

Energias eletrônicas e de solvatação precisas são calculadas. Para serem eficientes, as contribuições de solvatação são calculadas no nível r2SCAN-3c, em que tanto a energia eletrônica da fase gasosa quanto a contribuição de solvatação são obtidas e nenhum cálculo adicional é necessário. Conformeros com energias acima de 4 kcal/mol em relação ao mais estável, são descartados. Após a classificação, as contribuições termostáticas são calculadas no nível GFN2-xTB[ALPB] usando cálculo Hessiano de Ponto Único (SPH)³¹. As energias livres totais são calculadas e a classificação é realizada dentro da janela energética selecionada.

Parte 2 - Otimização e Cálculo da Energia Livre:

Os conformeros/estruturas relevantes devem ser otimizados no nível DFT (r2SCAN-3c) e usando solvatação implícita (neste trabalho, empregamos o SMD). Um otimizador do ensemble realiza otimizações de geometria para 8 iterações. Esta escolha foi sugerida pelos autores originais, e truncar os ciclos de otimização atenua o tempo de cálculo desta etapa. Em

seguida, um coeficiente de correlação de Spearman é calculado para verificar superfícies de energia potencial paralelas. Se a paralelidade puder ser assumida, o limite de classificação é diminuído e os confórmeros acima do limite são descartados, se sua norma de gradiente estiver abaixo de um limite predefinido. Para conjuntos grandes, isso diminui rapidamente o número de confórmeros elevados. A otimização em lote é repetida até que todos os confórmeros dentro da janela de energia 2,5 kcal/mol. Nas geometrias otimizadas no nível DFT, as energias livres são calculadas como na parte 1, com COSMO-RS(r2SCAN-3c) ou SMD(r2SCAN-3c) e GFN2-xTB[ALPB]-SPH. Os pesos de Boltzmann são calculados e uma energia livre média do conjunto pode ser obtida.

Parte 3 - Refinamento (opcional em nível superior de DFA híbrido):

Se for necessário refinar os pesos calculados de Boltzmann em um nível DFT mais alto (híbrido), isso pode ser feito na parte 3. Também deve ser notado que os pesos de Boltzmann calculados a partir de energias r2SCAN-3c bastante precisas são confiáveis o suficiente para a maioria das aplicações.⁴⁵ Como na parte 2, as contribuições de energia livre são calculadas em geometrias otimizadas por DFT, embora usando um nível de DFT mais alto (funcional híbrido). Os confórmeros abaixo do limite da soma de Boltzmann (soma populacional de 99%) são considerados para determinação das propriedades termodinâmicas médias. Na Figura 8, são mostradas estas propriedades no arquivo de saída do CENSO.

```
temperature /K:  avE(T) /a.u.  avGmRRHO(T) /a.u.  avGsolv(T) /a.u.  avG(T) /a.u.  avGcorrection(T)* /a.u.
-----
298.15           -345.3341583      0.0813121      -0.0241241     -345.2769703      0.0010410
-----
avGcorrection(T)* = Correction to free energy, which can be added (manually) to avG(T).
                    If only a small ensemble is evaluated in part3
                    this can incorporate information of higher lying
                    conformers of a more complete ensemble from part2.
```

Figura 8. Resultado final gerado pelo CENSO, que inclui a energia livre em cada faixa de temperatura selecionada, resultados em Hartree.

Resultados e Discussões

Abordaremos agora os resultados encontrados para as moléculas das três classes de compostos químicos de relevância atual, os fármacos, os biocidas e os plastificantes (somente a forma truncada). Na Tabela 5, são apresentados os valores da energia livre de Gibbs em fase líquida (na realidade, líquido sub-resfriado) das seis substâncias químicas. A energia $avGcorr(T)$ representa a correção na energia livre que deve ser manualmente adicionada na energia livre de Gibbs média $avG(T)$. Se somente um pequeno ensemble foi avaliado na etapa 3 do CENSO, esta correção pode incorporar informação sobre os confórmeros de mais alta energia de um conjunto mais completo da parte 2. Notamos que em todos os casos a correção $avGcorr(T)$ foi muito pequena e não deve afetar significativamente a previsão da pressão de vapor.

Tabela 5. Energias livre de Gibbs média do ensemble configuracional das moléculas de interesse industrial calculadas com o workflow CRENSO na temperatura de 298,15K e pressão de 1atm.

Propriedade	ABA	DET	DIN	MEF	PAV	TKT
Gás						
$avG(T)$	-1217.442058	-846.4333544	-1007.3904423	-1771.6502164	-1308.6911943	-1810.1733490
$avGcorr(T)$	0.0004382	0.0004262	0.0006361	0.0010108	0.0006632	0.0004606
Líquido						
$avG(T)$	-1217.4680718	-846.4510136	-1007.4073435	-1771.6802533	-1308.7226952	-1810.2010211
$avGcorr(T)$	0.0006583	0.0002959	0.0018278	0.0003009	0.0004886	0.0000861

O número de confórmeros encontrados ao final do CRENSO é listado na Tabela 6, discriminando as subetapas CREST e CENSO em fase vapor e líquido. Estes valores nos dá uma informação a respeito do número de graus de liberdade que é refletido diretamente no número de confórmeros encontrados. O plastificante DIN, como esperado, é o mais flexível e retornou a maior quantidade de confórmeros. O biocida TKT, por outro lado, é o menos flexível e conseqüentemente apresentou o menor número de confórmeros dentro da janela de energia selecionado no workflow deste trabalho. É interessante notar o efeito de solvente no número de estruturas. Em todos os casos, a inclusão do meio na busca conformacional, fez com que mais confórmeros fossem acessíveis nas condições de temperatura e pressão simuladas. Isto foi

particularmente mais dramático no plastificante DIN em que o número de conformémeros encontrados no líquido foi mais de 200% maior que na fase gasosa. Note ainda como o plastificante DIN apresenta um número bastante elevado de conformémeros que são aceitos e classificados pelo CREST/CRENDO, mesmo que sua estrutura seja uma forma truncada (ver Figura 5 e a discussão associada). Isto mostra que a simplificação estrutural dos plastificantes DIN e DET foi fundamental para pudéssemos concluir as simulações dos compostos dentro do período de um semestre letivo da Universidade de Brasília.

Tabela 6. Número de conformémeros encontrados nas subetapas do workflow CRENSO das moléculas selecionadas em fase vapor e líquido.

Molécula	CREST (vap)	CENSO (vap)	CREST (líq)	CENSO (líq)
ABA	113	27	116	34
DET	19	9	17	7
DIN	961	30	1059	128
MEF	218	46	532	49
PAV	86	32	122	49
TKT	6	5	13	7

Na Figura 9, é mostrada a sobreposição das estruturas de menor energia dos compostos químicos de cada classe obtidos ao final do workflow CRENSO em fase vapor e líquida. Esta sobreposição estrutural revela como em fase líquida as mudanças geométricas podem ser mais ou menos dramáticas em relação à fase vapor. Notamos que na classe dos fármacos, ABA e PAV apresentam grande similaridade estrutural nas duas fases. Por outro lado, os biocidas e os plastificantes são bastante afetados pela inclusão do meio líquido (lembre que ainda é uma mera aproximação, o solvente é tratado como contínuo, e os parâmetros SMD é para outro solvente de constante dielétrica similar). O plastificante DIN consegue acessar conformações mais torcidas quando em solução, justificando o número elevado de conformémeros obtidos na etapa CREST e CENSO do CRENSO (ver Tabela 6).

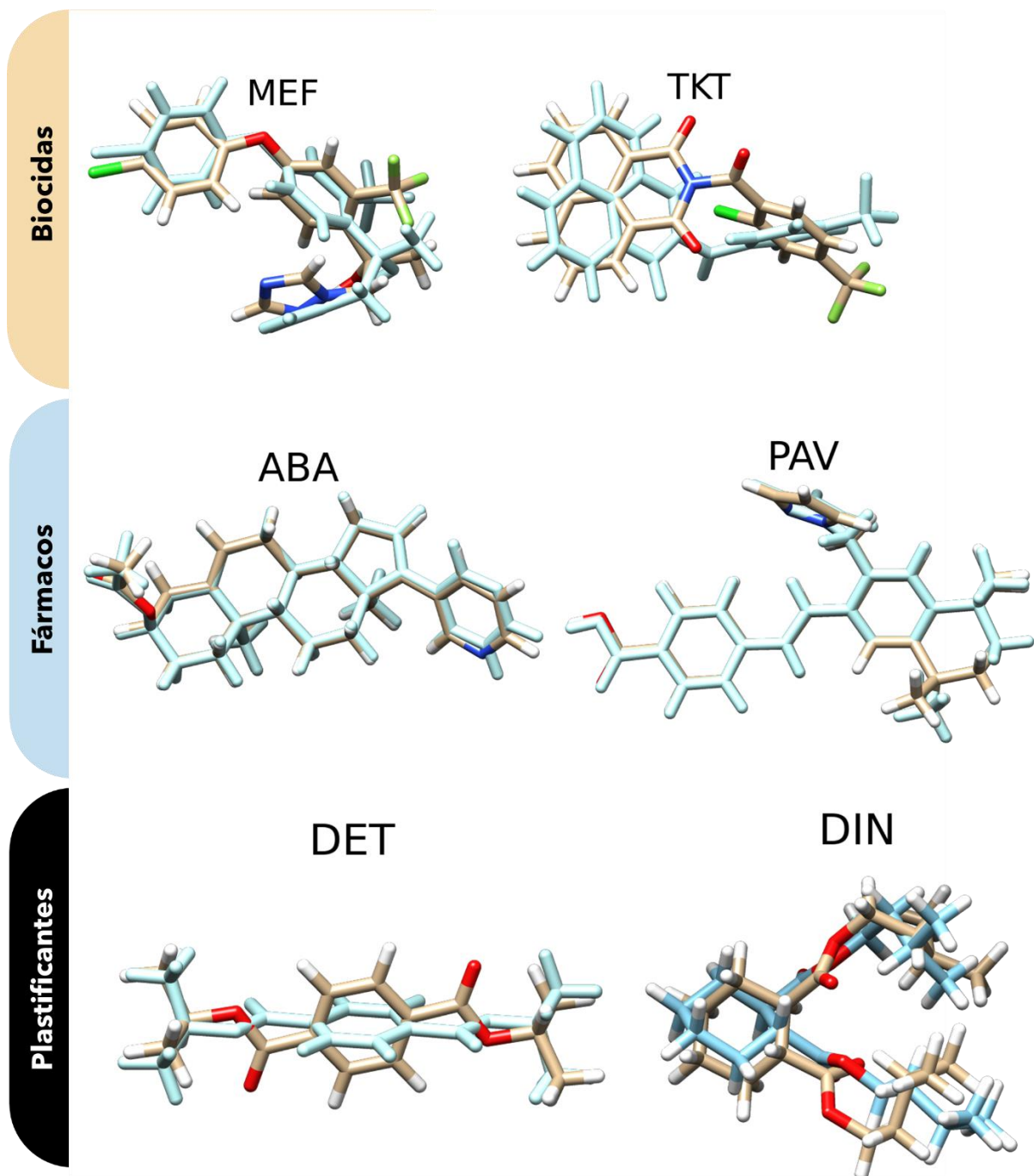


Figura 9. Sobreposição dos confórmeros de menor energia dos seis compostos selecionados em fase vapor (colorido) e fase líquida (azul claro) empregando o modelo contínuo SMD, obtidos ao final do workflow CRENSO.

Foram realizados os cálculos de energia livre e consequentemente da pressão de vapor (ver Eq. 1) para as 6 moléculas de interesse. Na Tabela 7, são apresentados os valores da pressão de vapor, P e seu valor em escala logarítmica $\log P$. Para alguns compostos, ABA, DET e TKT, foi possível realizar a tempo, simulações em triplicada e para estes casos apresentamos a média

da triplicata bem como o desvio padrão da pressão de vapor sobre as três simulações CRENSO. Os dados Tabela 7 mostram como a pressão de vapor calculada é muito baixa, exceto para os plastificantes DET e DIN que teve valores de $\log P$ comparáveis ao glicerol (ver Tab. 4). Todos os demais casos demais casos apresentaram valores de $\log P$ algumas ordens grandeza menores (em unidades logarítmicas). Isto nos leva a crer que esta pressão de vapor ordens de grandeza maior dos plastificantes se deve a estratégia de simplificação de suas estruturas, que potencialmente tornou os compostos muito mais voláteis do que seria na prática. Desta forma, os valores de $\log P$ deve ser tomado com cautela e não podemos esperar correlação quantitativa com valores experimentais desta classe de compostos.

Para demonstrar como as correções para energia livre de Gibbs ($avG_{corr}(T)$) apresentados na Tabela 5 tiveram impacto marginal e podem ser seguramente ignorados, construímos o gráfico de $\log P$ sem e com a correção $\log P_{corr}$ para os compostos selecionados e os resultados são mostrados na Figura 10. Note que na escala logarítmica, a pressão de vapor corrigida é idêntica a pressão de vapor usual calculada pelo CRENSO.

Tabela 7. Pressão de vapor em Pascal, $\log P$ e desvio padrão do $\log P$ de uma triplicata de simulações CRENSO das moléculas selecionadas neste trabalho, na temperatura de 298,15K e pressão de 1atm.

Molécula	P (Pa)	$\log P$	Desvio Padrão $\log P$
ABA	$7,767911 \times 10^{-6}$	-5,117257	0,100674
DET	$6,292974 \times 10^{-2}$	-1,202157	0,041957
DIN	$8,816532 \times 10^{-2}$	-1,054702	
MEF	$1,617334 \times 10^{-7}$	-6,791200	
PAV	$2,544763 \times 10^{-8}$	-7,594353	
TKT	$2,186798 \times 10^{-6}$	-5,666092	0,101477

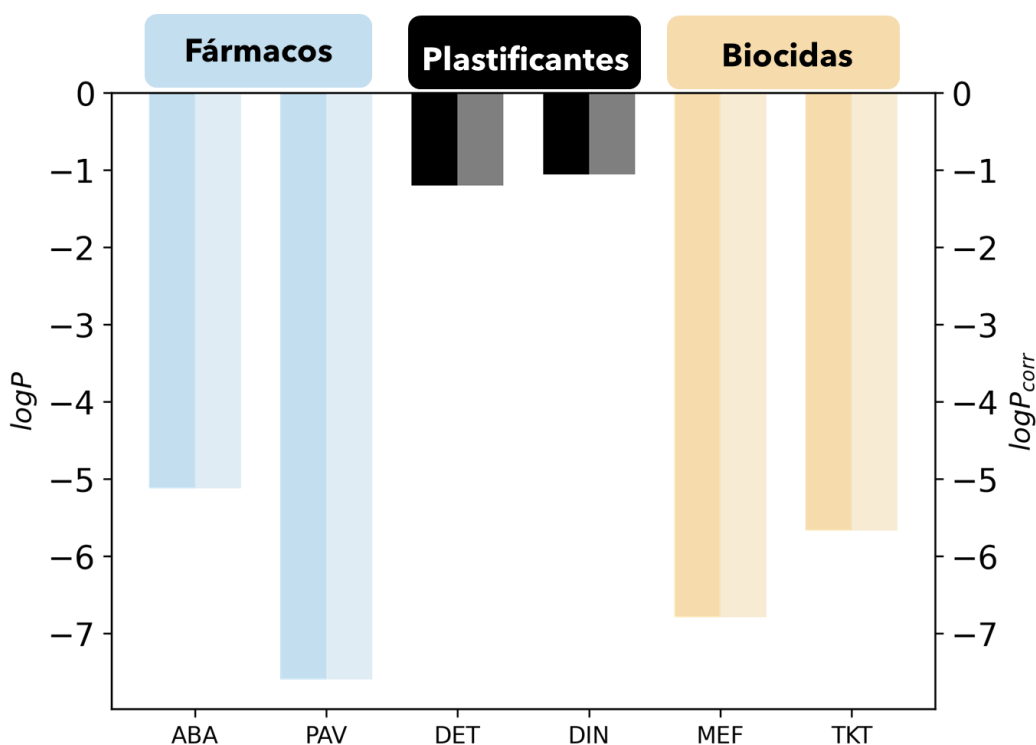


Figura 10. Logaritmo da pressão de vapor das moléculas da classe dos biocidas, plastificantes e fármacos calculada usando o workflow CRENSO na temperatura de 298,15K e pressão de 1 atm. Tons mais escuros (claros) representam os valores de $\log P$ (de $\log P_{corr}$) sem (com) as correções $avGcorr(T)$.

Com o intuito de averiguar a influência das conformações para o cálculo da média de Boltzmann nas energias livres que são essenciais para determinação de pressão de vapor, Eq. (1), nós calculamos $\log P$ dos seis compostos selecionando aleatoriamente uma conformação (empregamos o módulo *Random* da linguagem Python3) que denominaremos de $\log P(R)$ (para indicar a conformação aleatória), bem como o conformero mais estável em cada fase, que denominaremos de $\log P(L)$ (para indicar a conformação mais estável, *Lowest Conformation*). Desta forma veremos o impacto nos valores de pressão de vapor se negligenciarmos a distribuição configuracional das estruturas na pressão e temperatura de interesse. Os resultados podem ser encontrados na Tabela 8.

De acordo com os dados da Tab. 8, para alguns casos, como no fármaco ABA, ignorar a flexibilidade molecular introduz desvios da ordem de apenas 0,1 unidades de log. No caso do plastificante DIN os desvios podem ser maiores que 0.5 unidades de log, o que passa a ser um desvio considerável. Em suma, a inclusão do workflow CRENSO para prever a pressão de

vapor de compostos pouco voláteis passa a ser fundamental quanto mais flexível for o composto químico sob análise.

Tabela 8. Pressão de vapor em escala logarítmica dos compostos selecionados neste trabalho obtidos pelo método CRENSO que envolve médias de conformações e os respectivos valores encontrados empregando apenas a conformação mais estável ($\log P(L)$) e uma conformação aleatória ($\log P(R)$) na temperatura de 298,15K e pressão de 1 atm.

Composto	$\log P$ CRENSO (SMD)	$\log P(L)$	$\log P(R)$
MEF	-6.791200163	-6.430079538	-6.780206935
TKT	-5.666092307	-5.73379955	-6.030754683
ABA	-5.117256942	-5.125291044	-5.087297714
PAV	-7.594352591	-7.611785366	-6.922293816
DET	-1.2021566	-1.182239999	-1.179802171
DIN	-1.054702215	-1.451746315	0.172859526

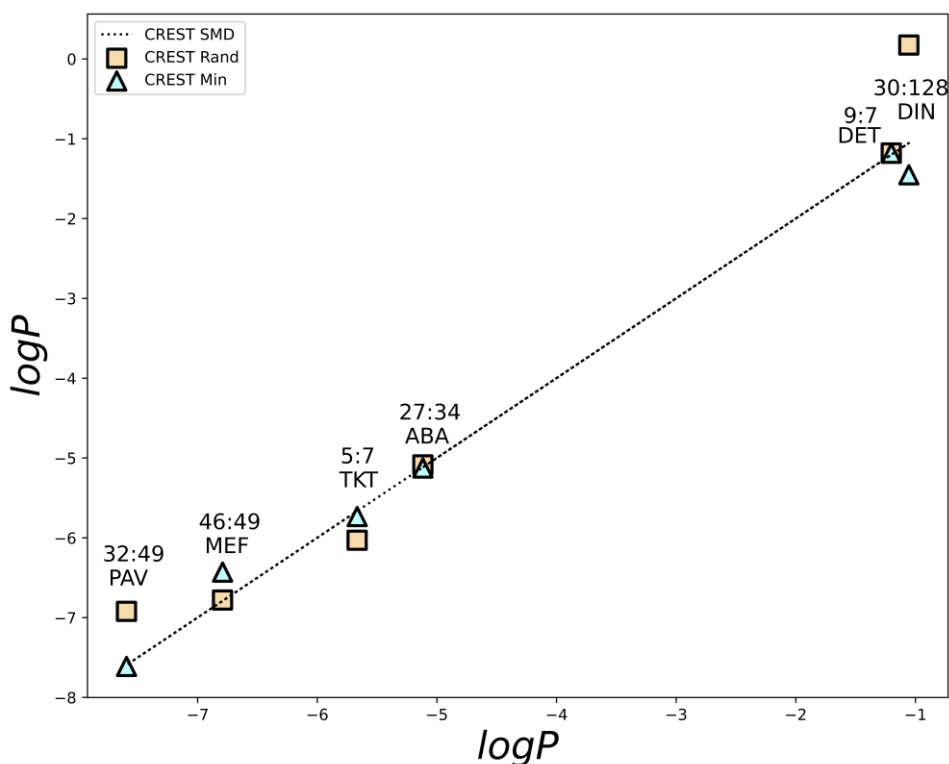


Figura 11. Comparação entre as pressões de vapor em escala logarítmica dos compostos selecionados neste trabalho, obtidas usando a média das energias do conjunto (linha de base) e

os valores das pressões de vapor de conformémeros selecionados de forma aleatória ($\log P(R)$) e apenas a conformação mais estável ($\log P(L)$).

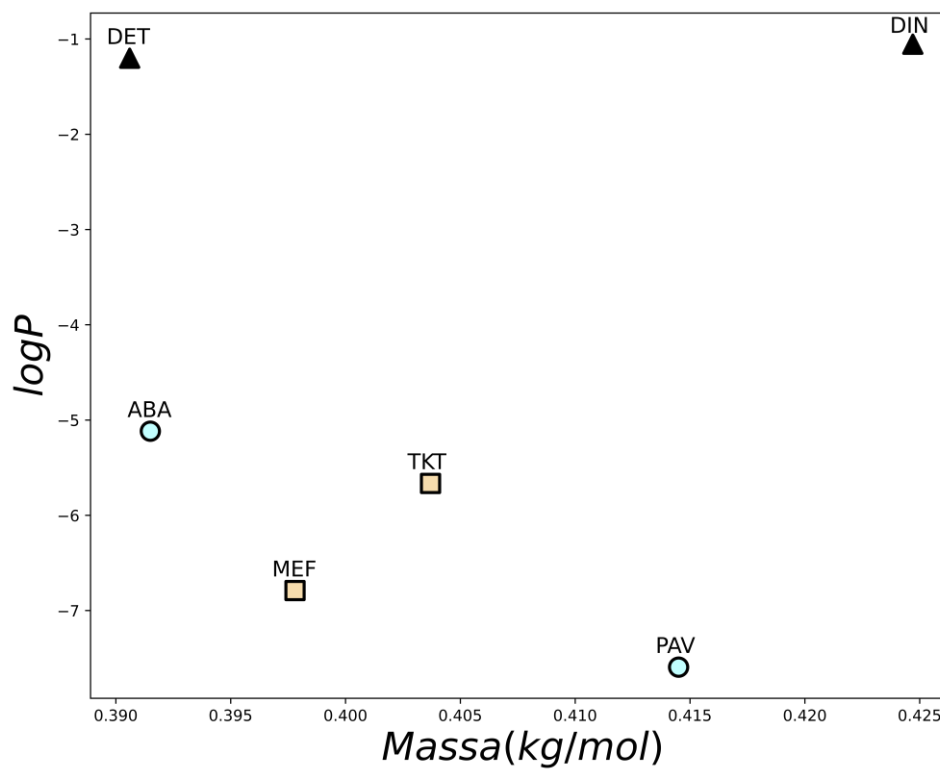


Figura 12. Relação entre a pressão de vapor e a massa dos compostos selecionados neste trabalho obtidos pelo método CRENSO.

Conclusões e Perspectivas

Neste trabalho, devido às dificuldades e incertezas experimentais na determinação da pressão de vapor de substâncias de baixíssima volatilidade, propomos a utilização do workflow CRENSO para estimar a pressão de vapor usando métodos de química teórica de relativo baixo custo. Para tal, propusemos uma alteração ao método CRENSO original, em que na etapa CENSO de refinamento das estruturas conformacionais do CREST, o solvente é originalmente tratado com o modelo de solvatação COSMO-RS disponível no pacote CosmoTherm que é um software pago. Alteramos o modelo de solvatação para o modelo SMD, que é um modelo disponível no pacote gratuito ORCA.

Para validação desta troca de modelo de solvatação no CRENSO, realizamos uma comparação com os compostos de teste, em que dados experimentais estão disponíveis e foram também testados no trabalho original do CRENSO.¹⁴ Com este intuito de validação, simulamos a pressão de vapor do glicerol, n-decano e naftaleno com o CRENSO-SMD nas condições de temperatura e pressão empregados no artigo original. Os resultados ficaram próximos e dentro da mesma ordem de grandeza dos valores usando o CRENSO-COSMO-RS e dos valores de referência. Isto nos mostrou que o CRENSO-SMD é uma alternativa para usuários que não tenham acesso ao CosmoTherm, além de se mostrar como sendo uma boa ferramenta para calcular a pressão de vapor de compostos orgânicos ambientalmente relevantes, sua instalação e uso são simples, não demandando um grande conhecimento de computação.

Verificamos que quando se deseja obter resultados rápidos, há algumas limitações, sendo a principal a quantidade de memória RAM disponível. Se o composto de interesse for relativamente grande com muitas ligações sigma o número de graus de liberdade pode ser excessivo, e ao gerar o conjunto conformacional com o CREST, o programa pode encerrar a execução por falta de memória RAM, esse uso intenso de memória ocorre durante o cruzamento de estruturas, etapa onde é realizado o cálculo do RMSD. Ainda que se disponibilize mais memória para a simulação, esta levará um tempo longo tornando-o não exequível a depender do prazo de estudo do usuário.

Realizamos a aplicação do CRENSO-SMD em moléculas da classe dos Biocidas, Fármacos e Plastificantes, pois são três classes de compostos químicos produzidos em larga escala pela indústria química moderna e a previsão de propriedades físico-químicas como a pressão de vapor pode ser correlacionada com sua persistência no ambiente, podendo em última

instância antecipar se estas substâncias se acumulariam no ambiente e/ou em organismos, incluindo humanos. As simulações retornaram pressões de vapor extremamente baixas (relativo). Para os plastificantes, a pressão de vapor simulada ficou várias ordens de grandeza acima dos biocidas e fármacos, mas isto pode ser atribuído à simplificação estrutural imposta nestes casos para diminuir o tempo de simulação destas moléculas flexíveis, onde podemos ver também que ao comparar a pressão de vapor com a massa do composto na Figura 12, os pontos relativos aos plastificantes nos mostram que foram os compostos que retornaram a menor pressão de vapor, e isso pode ser consequência do truncamento das cadeias, em estudos posteriores, pode-se verificar como o truncamento de estruturas podem afetar no resultado final da pressão de vapor.

Em resumo, este trabalho mostra que é viável a utilização da metodologia alternativa CRENSO-SMD para simular propriedades físico-químicas como a pressão de vapor de compostos orgânicos mesmo em computadores com configurações não robustas e pode ser uma ferramenta de interesse para química orgânica ambiental.

Como perspectivas futuras, outros tipos de estudos e validações podem ser realizados com o CRENSO, como o fluxo de trabalho para calcular o coeficiente de partição de compostos, onde os valores de energia livre podem ser obtidos para um determinado composto, realizando cálculos com o CRENSO para o composto usando água e octanol como solvente.

Referências Bibliográficas

- (1) Wang, Z.; Walker, G. W.; Muir, D. C. G.; Nagatani-Yoshida, K. Toward a Global Understanding of Chemical Pollution: A First Comprehensive Analysis of National and Regional Chemical Inventories. *Environ Sci Technol* **2020**, *54* (5). <https://doi.org/10.1021/acs.est.9b06379>.
- (2) Muir, D. C. G.; Getzinger, G. J.; McBride, M.; Ferguson, P. L. How Many Chemicals in Commerce Have Been Analyzed in Environmental Media? A 50 Year Bibliometric Analysis. *Environmental Science and Technology*. 2023. <https://doi.org/10.1021/acs.est.2c09353>.
- (3) Bayabil, H. K.; Teshome, F. T.; Li, Y. C. Emerging Contaminants in Soil and Water. *Frontiers in Environmental Science*. 2022. <https://doi.org/10.3389/fenvs.2022.873499>.
- (4) Chen, Y.; Lin, M.; Zhuang, D. Wastewater Treatment and Emerging Contaminants: Bibliometric Analysis. *Chemosphere*. 2022. <https://doi.org/10.1016/j.chemosphere.2022.133932>.
- (5) Morin-Crini, N.; Lichtfouse, E.; Fourmentin, M.; Ribeiro, A. R. L.; Noutsopoulos, C.; Mapelli, F.; Fenyvesi, É.; Vieira, M. G. A.; Picos-Corrales, L. A.; Moreno-Piraján, J. C.; Giraldo, L.; Sohajda, T.; Huq, M. M.; Soltan, J.; Torri, G.; Magureanu, M.; Bradu, C.; Crini, G. Removal of Emerging Contaminants from Wastewater Using Advanced Treatments. A Review. *Environmental Chemistry Letters*. 2022. <https://doi.org/10.1007/s10311-021-01379-5>.
- (6) Khan, S.; Naushad, M.; Govarthanam, M.; Iqbal, J.; Alfadul, S. M. Emerging Contaminants of High Concern for the Environment: Current Trends and Future Research. *Environ Res* **2022**, *207*. <https://doi.org/10.1016/j.envres.2021.112609>.
- (7) Morin-Crini, N.; Lichtfouse, E.; Liu, G.; Balaram, V.; Ribeiro, A. R. L.; Lu, Z.; Stock, F.; Carmona, E.; Teixeira, M. R.; Picos-Corrales, L. A.; Moreno-Piraján, J. C.; Giraldo, L.; Li, C.; Pandey, A.; Hocquet, D.; Torri, G.; Crini, G. Worldwide Cases of Water Pollution by Emerging Contaminants: A Review. *Environmental Chemistry Letters*. 2022. <https://doi.org/10.1007/s10311-022-01447-4>.
- (8) Torres, J. P. M.; Lailson-Brito, J.; Saldanha, G. C.; Dorneles, P.; e Silva, C. E. A.; Malm, O.; Guimarães, J. R.; Azeredo, A.; Bastos, W. R.; da Silva, V. M. F.; Martin, A. R.; Cláudio, L.; Markowitz, S. Persistent Toxic Substances in the Brazilian Amazon: Contamination of Man and the Environment. *J Braz Chem Soc* **2009**, *20* (6). <https://doi.org/10.1590/S0103-50532009000600024>.

- (9) Melymuk, L.; Blumenthal, J.; Sáňka, O.; Shu-Yin, A.; Singla, V.; Šebková, K.; Pullen Fedinick, K.; Diamond, M. L. Persistent Problem: Global Challenges to Managing PCBs. *Environ Sci Technol* **2022**, *56* (12). <https://doi.org/10.1021/acs.est.2c01204>.
- (10) Crinnion, W. J. Polychlorinated Biphenyls: Persistent Pollutants with Immunological, Neurological, and Endocrinological Consequences. *Alternative Medicine Review*. 2011.
- (11) Gruchlik, Y.; Linge, K.; Joll, C. Removal of Organic Micropollutants in Waste Stabilisation Ponds: A Review. *Journal of Environmental Management*. 2018. <https://doi.org/10.1016/j.jenvman.2017.10.020>.
- (12) Madadi, R.; Bester, K. Fungi and Biochar Applications in Bioremediation of Organic Micropollutants from Aquatic Media. *Marine Pollution Bulletin*. 2021. <https://doi.org/10.1016/j.marpolbul.2021.112247>.
- (13) Schwarzenbach, R. P.; Gschwend, P. M.; Imboden, D. M. *Environmental Organic Chemistry*; 2005. <https://doi.org/10.1002/0471649643>.
- (14) Stahn, M.; Grimme, S.; Salthammer, T.; Hohm, U.; Palm, W. U. Quantum Chemical Calculation of the Vapor Pressure of Volatile and Semi Volatile Organic Compounds. *Environ Sci Process Impacts* **2022**, *24* (11). <https://doi.org/10.1039/d2em00271j>.
- (15) Mennucci, B. Polarizable Continuum Model. *Wiley Interdisciplinary Reviews: Computational Molecular Science*. 2012. <https://doi.org/10.1002/wcms.1086>.
- (16) Hinckley, D. A.; Bidleman, T. F.; Foreman, W. T.; Tuschall, J. R.; Hinckley, D. A. Determination of Vapor Pressures for Nonpolar and Semipolar Organic Compounds from Gas Chromatographic Retention Data. *J Chem Eng Data* **1990**, *35* (3). <https://doi.org/10.1021/je00061a003>.
- (17) Koutek, B.; Hoskovec, M.; Vrkočová, P.; Feltl, L. Gas Chromatographic Determination of Vapour Pressures of Pheromone-like Compounds IV. Acetates, a Reinvestigation. *J Chromatogr A* **1997**, *759* (1–2). [https://doi.org/10.1016/S0021-9673\(96\)00775-3](https://doi.org/10.1016/S0021-9673(96)00775-3).
- (18) Yaws, C. L. *The Yaws Handbook of Vapor Pressure : Antoine Coefficients*; 2015.
- (19) Fishtine, S. H. Reliable Latent Heats of Vaporization. *Ind Eng Chem* **1963**, *55* (6). <https://doi.org/10.1021/ie50642a008>.
- (20) Lyman, W. J.; Reehl, W. F.; Rosenblatt, D. H. *Handbook of Chemical Property Estimation Methods: Environmental Behavior of Organic Compounds*; Washington, DC (United States); American Chemical Society: United States, 1990.
- (21) Asher, W. E.; Pankow, J. F.; Erdakos, G. B.; Seinfeld, J. H. Estimating the Vapor Pressures of Multi-Functional Oxygen-Containing Organic Compounds Using Group

- Contribution Methods. *Atmos Environ* **2002**, *36* (9). [https://doi.org/10.1016/S1352-2310\(01\)00564-7](https://doi.org/10.1016/S1352-2310(01)00564-7).
- (22) Asher, W. E.; Pankow, J. F. Vapor Pressure Prediction for Alkenoic and Aromatic Organic Compounds by a UNIFAC-Based Group Contribution Method. *Atmos Environ* **2006**, *40* (19). <https://doi.org/10.1016/j.atmosenv.2005.12.004>.
- (23) Pracht, P.; Bohle, F.; Grimme, S. Automated Exploration of the Low-Energy Chemical Space with Fast Quantum Chemical Methods. *Physical Chemistry Chemical Physics* **2020**, *22* (14). <https://doi.org/10.1039/c9cp06869d>.
- (24) Grimme, S.; Bohle, F.; Hansen, A.; Pracht, P.; Spicher, S.; Stahn, M. Efficient Quantum Chemical Calculation of Structure Ensembles and Free Energies for Nonrigid Molecules. *Journal of Physical Chemistry A*. 2021. <https://doi.org/10.1021/acs.jpca.1c00971>.
- (25) Klamt, A.; Schüürmann, G. COSMO: A New Approach to Dielectric Screening in Solvents with Explicit Expressions for the Screening Energy and Its Gradient. *Journal of the Chemical Society, Perkin Transactions 2* **1993**, No. 5. <https://doi.org/10.1039/P29930000799>.
- (26) Eckert, F.; Klamt, A. Fast Solvent Screening via Quantum Chemistry: COSMO-RS Approach. *AIChE Journal* **2002**, *48* (2). <https://doi.org/10.1002/aic.690480220>.
- (27) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *Journal of Physical Chemistry B* **2009**, *113* (18). <https://doi.org/10.1021/jp810292n>.
- (28) Neese, F.; Wennmohs, F.; Becker, U.; Riplinger, C. The ORCA Quantum Chemistry Program Package. *Journal of Chemical Physics* **2020**, *152* (22). <https://doi.org/10.1063/5.0004608>.
- (29) Spicher, S.; Grimme, S. Efficient Computation of Free Energy Contributions for Association Reactions of Large Molecules. *Journal of Physical Chemistry Letters* **2020**, *11* (16). <https://doi.org/10.1021/acs.jpcllett.0c01930>.
- (30) Bannwarth, C.; Caldeweyher, E.; Ehlert, S.; Hansen, A.; Pracht, P.; Seibert, J.; Spicher, S.; Grimme, S. Extended Tight-Binding Quantum Chemistry Methods. *Wiley Interdisciplinary Reviews: Computational Molecular Science*. 2021. <https://doi.org/10.1002/wcms.1493>.
- (31) Spicher, S.; Grimme, S. Single-Point Hessian Calculations for Improved Vibrational Frequencies and Rigid-Rotor-Harmonic-Oscillator Thermodynamics. *J Chem Theory Comput* **2021**, *17* (3). <https://doi.org/10.1021/acs.jctc.0c01306>.

- (32) Ben-Nairn, A.; Marcus, Y. Solvation Thermodynamics of Nonionic Solutes. *J Chem Phys* **1984**, *81* (4). <https://doi.org/10.1063/1.447824>.
- (33) Sigalov, G.; Fenley, A.; Onufriev, A. Analytical Electrostatics for Biomolecules: Beyond the Generalized Born Approximation. *Journal of Chemical Physics* **2006**, *124* (12). <https://doi.org/10.1063/1.2177251>.
- (34) Ehlert, S.; Stahn, M.; Spicher, S.; Grimme, S. Robust and Efficient Implicit Solvation Model for Fast Semiempirical Methods. *J Chem Theory Comput* **2021**, *17* (7). <https://doi.org/10.1021/acs.jctc.1c00471>.
- (35) Bannwarth, C.; Caldeweyher, E.; Ehlert, S.; Hansen, A.; Pracht, P.; Seibert, J.; Spicher, S.; Grimme, S. Extended Tight-Binding Quantum Chemistry Methods. *Wiley Interdisciplinary Reviews: Computational Molecular Science*. 2021. <https://doi.org/10.1002/wcms.1493>.
- (36) Pracht, P.; Bohle, F.; Grimme, S. Automated Exploration of the Low-Energy Chemical Space with Fast Quantum Chemical Methods. *Physical Chemistry Chemical Physics* **2020**, *22* (14). <https://doi.org/10.1039/c9cp06869d>.
- (37) Neese, F.; Wennmohs, F.; Becker, U.; Riplinger, C. The ORCA Quantum Chemistry Program Package. *Journal of Chemical Physics* **2020**, *152* (22). <https://doi.org/10.1063/5.0004608>.
- (38) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. Gaussian 16, Rev. C.01. *Gaussian 16, Rev. C. 01* **2016**.
- (39) Robitaille, P. O.; Abou-Rachid, H.; Brisson, J. Dielectric Constant Predictions for Energetic Materials Using Quantum Calculations. *Defence Technology*. 2021. <https://doi.org/10.1016/j.dt.2020.09.022>.
- (40) Stahn, M.; Grimme, S.; Salthammer, T.; Hohm, U.; Palm, W. U. Quantum Chemical Calculation of the Vapor Pressure of Volatile and Semi Volatile Organic Compounds. *Environ Sci Process Impacts* **2022**, *24* (11). <https://doi.org/10.1039/d2em00271j>.

- (41) Cammenga, H. K.; Schulze, F. W.; Theuerl, W. Vapor Pressure and Evaporation Coefficient of Glycerol. *J Chem Eng Data* **1977**, *22* (2).
<https://doi.org/10.1021/je60073a004>.
- (42) R'žika, K.; Majer, V. Simultaneous Treatment of Vapor Pressures and Related Thermal Data Between the Triple and Normal Boiling Temperatures for N-Alkanes C5–C20. *J Phys Chem Ref Data* **1994**, *23* (1). <https://doi.org/10.1063/1.555942>.
- (43) Lei, Y. D.; Chankalal, R.; Chan, A.; Wania, F. Supercooled Liquid Vapor Pressures of the Polycyclic Aromatic Hydrocarbons. *J Chem Eng Data* **2002**, *47* (4).
<https://doi.org/10.1021/je0155148>.
- (44) Witte, J.; Neaton, J. B.; Head-Gordon, M. Effective Empirical Corrections for Basis Set Superposition Error in the Def2-SVPD Basis: GCP and DFT-C. *Journal of Chemical Physics* **2017**, *146* (23). <https://doi.org/10.1063/1.4986962>.
- (45) Grimme, S.; Hansen, A.; Ehlert, S.; Mewes, J. M. R2SCAN-3c: A “Swiss Army Knife” Composite Electronic-Structure Method. *Journal of Chemical Physics* **2021**, *154* (6).
<https://doi.org/10.1063/5.0040021>.

Apêndice A

Solventes disponíveis nos modelos ALPB e SMD no CREST/ORCA

Tabela A 1. Solventes disponíveis nos programas gratuitos ORCA e XTB empregando os modelos PCM SMD e ALPB.

Solvente ^a	Constante Dielétrica	Nome no input
Hexano	1.9	[N-HEXANE, N-HEXANE]
Octano	1.94	[N-OCTANE, N-OCTANE]
Ciclohexano	2.0	[CYCLOHEXANE, CYCLOHEXANE]
Hexadecano	2.1	[N-HEXADECANE, N- HEXADECANE]
Tetracloro de Carbono	2.2	[CARBON TETRACHLORIDE, CARBON TETRACHLORIDE]
Dioxano	2.2	[1,4-DIOXANE, 1,4- DIOXANE]
Benzeno	2.3	[BENZENE, BENZENE]
Tolueno	2.4	[TOLUENE, TOLUENE]
Dissulfeto de carbono	2.6	[CARBON DISULFIDE, CARBON DISULFIDE]
Furano	3.0	[None, THF]
chbr3	4.25	[BROMOFORM, BROMOFORM]
diethylether	4.4	[DIETHYL ETHER, DIETHYL ETHER]
chcl3	4.8	[CHLOROFORM, CHLOROFORM]
dibromoethane	4.9	[1,2-DIBROMOETHANE, 1,2-DIBROMOETHANE]
ethylacetate	5.9	[ETHYL ETHANOATE, ETHYL ETHANOATE]

aniline	6.9	[ANILINE, ANILINE]
thf	7.6	[TETRAHYDROFURAN, TETRAHYDROFURAN]
phenol	8.0	[None, THIOPHENOL]
woctanol	8.1	[None, 1-OCTANOL]
ch2cl2	9.1	[DICHLOROMETHANE, DICHLOROMETHANE]
octanol	9.9	[1-OCTANOL, 1-OCTANOL]
dichloroethane	10.125	[1,2-DICHLOROETHANE, 1,2-DICHLOROETHANE]
benzaldehyde	18.2	[BENZALDEHYDE, BENZALDEHYDE]
acetone	20.7	[ACETONE, ACETONE]
ethanol	24.6	[ETHANOL, ETHANOL]
methanol	32.7	[METHANOL, METHANOL]
acetonitrile	36.6	[ACETONITRILE, ACETONITRILE]
nitromethane	38.2	[NITROMETHANE, NITROMETHANE]
dmf	38.3	[N,N- DIMETHYLFORMAMIDE, N,N- DIMETHYLFORMAMIDE]
dmsO	47.2	[DIMETHYLSULFOXIDE, DIMETHYLSULFOXIDE]
h2o	80.1	[WATER, WATER]