

Automated Computational Screening of Temperature-Dependent Cannabinoid Solubility in Industrially Relevant Processing Solvents and Lipid Carrier Oils

Mohammadhu Rifaydheen Sakeer

sakeermr5@gmail.com

Standard Seed Corporation

Katukoliya Gamage Anuththara Samadhi

Standard Seed Corporation

Shashika Kalanaprabhath Kotawalagedara

Standard Seed Corporation

Pamalka Hansani Kaldera

Standard Seed Corporation

Damilola Samuel Bodun

Université Paris Cité

Evan Friedmann

Hellma

Kevin Kimmell

Arvida Labs

Manny Maminakis

Cannamatrix

David Aronson

Standard Seed Corporation

Mike Woster

Standard Seed Corporation

Suliman Sharif

Standard Seed Corporation

Case Report

Keywords: Cannabinoids, Computational screening, Cannabis formulations, Carrier oils, Processing solvents, Solubility prediction, SolProp, Temperature-dependent solubility

Posted Date: July 30th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10131067/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Background

Cannabinoids are highly lipophilic compounds characterized by poor aqueous solubility. This poses significant challenges in cannabinoid extraction, purification, and formulation. They are classified as Biopharmaceutics Classification System (BSC) Class II drugs. Despite their increasing relevance in pharmaceutical and industrial applications, comprehensive data are limited on their solubility across a wide range of solvents, solvent mixtures, and temperatures.

Methods

An automated computational workflow was used to predict the solubility ($\log S$, mol L^{-1}) of 19 cannabinoids and reference compounds across 18 processing solvents and carrier oils at three different temperatures: 0°C (273.15 K), 25°C (298.15 K), and 50°C (323.15 K). This was done using MIT Reaction Mechanism Generator (RMG) Solid Solubility Prediction Tool (SolProp) using an automated Python-based workflow. A total of 1026 $\log S$ data points were generated and analyzed.

Results

The predicted values showed a consistent trend in solubility based on the polarity of the solvent. In our study, the predicted solubility increased consistently with temperature across all solute-solvent systems. This indicates an endothermic dissolution process, which agrees with established thermodynamic findings in the literature. Polar aprotic solvents, such as Dimethyl Sulfoxide (DMSO) and acetone, as well as protic alcohols, such as ethanol, isopropanol, and methanol, showed the highest solubility, while non-polar hydrocarbons, such as hexane and heptane, showed significantly lower solubility. These findings aligned with previous experimental and computational findings. Acid-cannabinoids (Cannabinolic acid, Δ^9 -THCA, 9(R)-HHC Acid, 9(S)-HHC Acid, and Cannabinol acetate) had slightly higher solubility than natural cannabinoids (e.g., Cannabidiol, Δ^9 -THC, Cannabinol, Cannabigerol, Cannabifuran, and Cannabidivarin) at tested temperatures. This is likely due to the presence of a carboxylic acid group, which enhances the interactions with polar solvents. All three hexahydrocannabinol stereoisomers-9(R)-HHC, 9(S)-HHC, and iso-hexahydrocannabinol (iso-HHC)-showed similar solubility, with acetone at 273.15 K indicating predicted $\log S$ values of + 1.62, + 1.62, and + 1.54, respectively. Meanwhile, non-cannabinoid references (acetazolamide and melatonin) showed positive $\log S$ value in polar solvents such as DMSO and ethanol but negative values in nonpolar hydrocarbon solvents.

Conclusion

This study provides a large-scale, temperature-dependent computational cannabinoid solubility dataset across industrial solvents and offers understanding for process optimization in extraction, purification, and formulation.

Background

Cannabis sativa is one of the oldest plants known for medicinal use, and it was incorporated into Western medicine in the early 19th century. It contains a wide variety of natural compounds, including cannabinoids and non-cannabinoid substances such as terpenes, phenolic compounds, and alkaloids. So far, more than 500 compounds have been identified in *C. sativa*, and around 125 of them are cannabinoids. These cannabinoids are C₂₁ terpeno-phenolic compounds unique to this plant (1).

The main cannabinoids found in this plant, such as Cannabidiol (CBD), Δ^9 -Tetrahydrocannabinol (Δ^9 -THC), Cannabinol (CBN), Cannabigerol (CBG), and their carboxylic acid precursors, are naturally available and easily accessible. Subsequently, these compounds were modified to develop novel cannabinoids, such as 9(R)-HHC and 9(S)-HHC. These novel semi-synthetic compounds differ in their functionalities compared to their precursors (2).

Because of the high lipophilicity and significantly lower aqueous solubility cannabinoids were categorized in BCS class II making challenges (3) in product development chain (4). Therefore, extraction of these chemicals from plant materials, purification, formulation into medicinal products and solvent removal is more difficult in product manufacturing (5).

Therefore, solubility prediction for cannabinoids in different solvents in early-stage development can reduce the time and cost in cannabinoid-based product development. This approach is also useful for cannabis-based advanced separation techniques in the hemp processing industry (6). For solubility prediction, different types of experimental methods such as shake-flask method (7), machine learning models and group-contribution thermodynamic models (UNIFAC, COSMO-RS) were used (8, 9). The MIT Reaction Mechanism Generator (RMG) solvation tool is a fully open-source tool that estimates solvation properties for a given solute-solvent pair at a desired temperature. Furthermore, this software does not rely on a single formula alone; instead, it uses separate curated property datasets with correction schemes for liquid-phase solvent-solute effects that can capture broad solvent coverage (10). On the other hand, SolProp is a separate solubility prediction project from MIT's Green Lab, and it is a machine learning-based model with a thermodynamics-based approach to predict log S values by using SMILES for both solute and solvent that can handle arbitrary temperatures for a wide range of organic solvents (11). In 2025, MIT later released another tool called FastSolv that provided more accurate results in benchmarks compared with Solprop (12).

According to the published articles, COSMO-RS predictions and shake-flask measurements for CBD with a broad spectrum of solvents confirm that CBD has higher solubility in DMSO, acetone, and alcohols (13). Furthermore, isothermal equilibration in n-pentane, n-hexane, n-heptane, and isooctane between 253.1 K and 310.1 K measured solubility for CBD, and differential scanning calorimetry identified its

melting points and enthalpy of fusion. Finally, the solubility data were analyzed using Wilson, NRTL, and UNIQUAC activity coefficient equations (14). However, no published work has been found to identify the solubility of cannabinoid compounds across a full industrial solvent spectrum and at different temperatures using a computational framework.

Therefore, in this research project, we used 19 cannabinoids and related compounds across 18 industrially validated solvents and carrier oils at three different temperatures, namely, 273.15 K, 298.15 K, and 323.15 K. For solubility prediction, we used the Solprop computational framework with the Solprop database. We developed the advanced solvent panel through direct consultation with process chemists and industry partners from the cannabis industry. The main objectives of this research project are to generate a temperature-dependent log S dataset across the full industrial solvent panel and to validate the generated predictions against experimental literature and industrial practices.

Phytocannabinoids are a diverse family of terpenoids originally characterized in *Cannabis sativa*, and they are the most important bioactive compounds (15).

However, recent studies have shown that other plants and fungi also produce cannabinoid-like compounds (16). Elucidating the structures of Δ^9 -tetrahydrocannabinol (Δ^9 -THC) and cannabidiol (CBD) by Raphael Mechoulam in the 1960s significantly advanced the field of cannabis chemistry (17). To date, over 140 cannabinoids have been identified, including cannabigerol (CBG), cannabichromene (CBC), cannabinol (CBN), and various minor cannabinoids (13). Interestingly, minor structural changes can yield novel cannabinoids. For example, hexahydrocannabinol (HHC) is a semi-synthetic cannabinoid derived from Δ^9 -THC (18). Studies have shown that the molecular structure of the (9R)-HHC isomer is much more similar to that of Δ^9 -THC than (9S)-HHC. Therefore, (9R)-HHC binds more strongly to cannabinoid receptors than (9S)-HHC (19). Historical texts indicate that cannabis has been used for food, medicine, and spiritual purposes. Medical texts such as 'Shi Liao Ben Cao' and 'Compendium of Materia Medica' in China mentioned the medical importance of cannabis extracts (20, 21). Cannabinoids are useful for producing various drugs. For example, purified CBD was FDA-approved in 2018 for the treatment of refractory pediatric epilepsy (22). Moreover, cannabinoids are used as anti-inflammatory and analgesic agents (23).

Generally, cannabinoids are hydrophobic and difficult to dissolve in water. They have rigid ring systems and higher melting points. High Log P values of CBD, CBN, and THC confirm their lipophilic nature. Also, CBD has been classified as a BCS class II compound based on its low solubility. Therefore, lipid-based formulations are commonly used to enhance their bioavailability (12, 24, 25). Low aqueous solubility presents significant challenges for product design, formulation, and requires higher dosages to achieve therapeutic effects (26, 27). Furthermore, solubility governs crystallization and phase behavior during synthesis and purification (12). In practice, cannabinoids require solubilizers, complexing agents or lipid carriers. As an example, medium-chain triglyceride (MCT) oil is a common carrier for cannabis extracts. It has higher terpene extraction efficiency and better shelf life stability compared to olive oil (28). In

industrial cannabis processing, solvent selection for extraction is a key bottleneck. Therefore, identifying a better solvent using quick assays or models helps to avoid unnecessary trials and costs.

Traditional solvent selection for phytochemical extraction focused on an empirical trial and error method. Industry researchers relied on the Hansen Solubility Parameters (HSP) to map the dispersion (δ_D), polar (δ_P), and hydrogen-bonding (δ_H) forces of solutes. While HSP effectively narrows solvent candidates, it remains a qualitative guide that lacks the ability to predict precise molar concentrations or temperature-dependent fluctuations. In contrast, the "shake-flask" method serves as the experimental gold standard. It provides precise equilibrium solubility data through prolonged isothermal agitation and HPLC analysis. However, the resource-intensive nature of this physical validation creates a bottleneck in high-throughput formulation (29).

Group contribution methods like UNIFAC predict activity coefficients from molecule fragments, while COSMO-RS and COSMO-SAC are solvation models based on electronic surface screening (30). UNIFAC provides initial estimates rapidly, whereas COSMO-RS provides more accurate results but at a higher computational cost. The literature suggests that COSMO-RS has been applied to predict CBD and curcumin solubilities and shows good agreement for the solvation of compounds (13). Recently, data-driven machine learning approaches have emerged. The MIT Reaction Mechanism Generator (RMG) solvation tool is an open-source framework designed to estimate the thermochemical properties of solute-solvent systems at variable temperatures. According to (10), the RMG solvation tool utilizes a modular architecture of curated property datasets and correction schemes instead of simple formulas. The software achieves broad chemical coverage and structural adaptability by integrating solvent libraries with group additivity algorithms (10).

Furthermore, the 'SolProp' model developed by (11) uses neural networks trained on curated solubility data with thermodynamic parameters. The model's predictive accuracy was established by training on the CombiSolv-Exp database, which contains approximately 11,000 experimental solvation free energy measurements. The model then utilized the CombiSolv-Exp database, consisting of approximately 5,000 temperature-dependent solid solubility measurements across 115 unique solutes and 95 organic solvents. The researchers employed a transfer learning protocol. This protocol involved an initial "pre-training" phase using the CombiSolv-QM database with one million solvation free energies calculated via COSMO-RS theory. Moreover, the framework integrates reference data from the AqueousSolv-Exp database to anchor its thermodynamic cycles. SolProp can accept Simplified Molecular Input Line Entry System (SMILES) strings of solutes and solvents. Traditional models, such as UNIFAC, require a researcher to manually identify and parameterize specific functional groups. SolProp bypasses these manual steps by employing Directed Message Passing Neural Networks (D-MPNNs). These networks learn the specific molecular environment of the solute-solvent pair and generate latent embeddings that represent complex intermolecular forces such as hydrogen bonding and van der Waals interactions. This framework strictly adheres to a thermodynamic cycle. The model decomposes the query into two steps: 1) predicting the liquid-phase activity coefficient, and 2) predicting the solid-liquid phase equilibrium. This structural grounding ensures that predictions remain physically plausible across different

temperatures (11). A newer tool named Fastsov is available as an open-source package and web service. It is a further improved prediction tool by an order of magnitude via optimized neural architectures and a large dataset (12).

Although individual cannabinoids such as THC, CBD have been studied in selected solvents, the literature lacks a comprehensive, temperature-dependent solubility dataset for many phytocannabinoids. For example, studies suggested that CBD solubility in DMSO, acetone, alcohol, and similar solvents, but they do not cover less common phytocannabinoids such as HHC, CBN, CBG etc (31). This project addresses this gap by using the SolProp framework to predict log-solubilities (log S) for 19 different cannabinoids in 18 industry-relevant solvents at three temperatures (273.15 K, 298.15 K, 323.15 K). The solvent panel was chosen in consultation with process chemists to reflect real-world processing needs. SolProp provides a high-throughput alternative for industry researchers, rather than expensive and time-consuming laboratory shake-flask testing. To our knowledge, combining prediction models like SolProp with a solvent panel specifically focused on cannabinoids, has not yet been documented in the literature.

Methodology

Study design and objective

This study was designed as an automated computational solubility screen. The objective was to generate a temperature-dependent log S dataset for cannabinoid-related compounds. Furthermore, reference compounds across a solvent panel selected for relevance to industrial cannabinoid extraction, purification and formulation. No new experimental solubility measurements were performed in this study.

Solutes

Nineteen cannabinoids were used as solutes. Six are neutral cannabinoids such as Cannabidiol (CBD) (32), Δ^9 -Tetrahydrocannabinol (THC) (33), Cannabinol (CBN), Cannabigerol (CBG), Cannabifuran (CBF), Cannabinodiol (CBND).

Furthermore, there are three Acid-form cannabinoids : Δ^9 -Tetrahydrocannabinolic Acid A (THCA-A), Cannabinolic Acid (CBNA) (33), Cannabinol Acetate (CBN-OAc); Five were Hexahydrocannabinol stereoisomers and acid forms: (9(R)-Hexahydrocannabinol (9(R)-HHC), 9(S)-Hexahydrocannabinol (9(S)-HHC), iso-Hexahydrocannabinol (iso-HHC) (19), 9(R)-Hexahydrocannabinolic Acid (9(R)-HHC Acid), 9(S)-Hexahydrocannabinolic Acid (9(S)-HHC Acid)) (34).

Moreover, three hydroxylated CBN metabolites were identified such as 11-Hydroxycannabinol (11-OH-CBN), 1'-Hydroxycannabinol (1'-OH-CBN), 8-Hydroxycannabinol (8-OH-CBN) (35). Additionally, two non-cannabinoid reference compounds were included such as Acetazolamide and Melatonin.

Table 1
Canonical SMILES strings for evaluated solutes

Chemical Name	Canonical SMILES
Cannabinol (CBN)	<chem>CCCCC1=CC(=C2C(=C1)OC(C3=C2C=C(C=C3)C)(C)C)O</chem>
Cannabifuran (CBF)	<chem>CCCCC1=CC(=C2C(=C1)OC3=C(C=CC(=C23)C(C)C)C)O</chem>
Cannabinodiol (CBND)	<chem>CCCCC1=CC(=C(C(=C1)O)C2=C(C=CC(=C2)C)C(=C)C)O</chem>
Cannabinolic Acid (CBNA)	<chem>CCCCC1=CC2=C(C3=C(C=CC(=C3)C)C(O2)(C)C)C(=C1C(=O)O)O</chem>
Cannabinol Acetate (CBN-OAc)	<chem>CC(C=C1)=CC2=C1C(C)(C)OC3=C2C(OC(C)=O)=CC(CCCC)=C3</chem>
11-Hydroxycannabinol (11-OH-CBN)	<chem>CCCCC1=CC(=C2C(=C1)OC(C3=C2C=C(C=C3)CO)(C)C)O</chem>
1'-Hydroxycannabinol (1'-OH-CBN)	<chem>CC(C=C1)=CC2=C1C(C)(C)OC3=C2C(O)=CC(C(O)CCCC)=C3</chem>
8-Hydroxycannabinol (8-OH-CBN)	<chem>CC(C(O)=C1)=CC2=C1C(C)(C)OC3=C2C(O)=CC(CCCC)=C3</chem>
Cannabidiol (CBD)	<chem>CCCCC1=CC(O)=C([C@H]2[C@H](C(C)=C)CCC(C)=C2)C(O)=C1</chem>
Acetazolamide	<chem>NS(=O)(=O)c1nnc(s1)NC(=O)C</chem>
Melatonin	<chem>COC1=CC2=C(NC=C2CCNC(C)=O)C=C1</chem>
Cannabigerol (CBG)	<chem>CCCCC1=CC(O)=C(C/C=C(C)/CC/C=C(C)\C)C(O)=C1</chem>
Δ^9 -Tetrahydrocannabinol (THC)	<chem>CCCCC1=CC(=C2[C@@H]3C=C(CC[C@H]3C(OC2=C1)(C)C)C)O</chem>
Δ^9 -Tetrahydrocannabinolic Acid A (THCA-A)	<chem>CCCCC1=CC2=C([C@@H]3C=C(CC[C@H]3C(O2)(C)C)C)C(=C1C(=O)O)O</chem>
9(R)-Hexahydrocannabinolic Acid (9(R)-HHC Acid)	<chem>CC1(C)[C@@]2([H])[C@](C[C@H](C)CC2)([H])C3=C(C=C(C(CCCC)C(C(O)=O)=C3O)O1</chem>
9(S)-Hexahydrocannabinolic Acid (9(S)-HHC Acid)	<chem>CC1(C)[C@@]2([H])[C@](C[C@@H](C)CC2)([H])C3=C(C=C(C(CCCC)C(C(O)=O)=C3O)O1</chem>
iso-Hexahydrocannabinol (iso-HHC)	<chem>C[C@@]12C[C@H]([C@H](C(C)C)CC2)C3=C(C=C(C(CCCC)C=C3O)O1</chem>
9(S)-Hexahydrocannabinol (9(S)-HHC)	<chem>CC1(C)[C@@]2([H])[C@](C[C@@H](C)CC2)([H])C3=C(C=C(C(CCCC)C=C3O)O1</chem>
9(R)-Hexahydrocannabinol (9(R)-HHC)	<chem>CC1(C)[C@@]2([H])[C@](C[C@H](C)CC2)([H])C3=C(C=C(C(CCCC)C=C3O)O1</chem>

Solvents and temperatures

The 18 industrially used solvents were selected based on consultation with process chemists at Arvida Labs. These solvents include DMSO, acetone, acetonitrile, ethanol, methanol, isopropyl alcohol, hexane, heptane, naphtha, butane, propane, CO₂, R134a, MCT oil, coconut oil, butter, olive oil, and sesame oil.

Computational workflow

Log S values ($\log_{10} \text{ mol L}^{-1}$) were predicted using the solubility prediction model SolProp from MIT RMG Solid Solubility Prediction (11) for all solute-solvent pairs evaluated at 273.15K, 298.15K, and 323.15K. Finally, 1026 data points were collected. A Python 3.10-based automated pipeline was implemented and orchestrated through a CI/CD YAML workflow to generate a CSV (Comma-Separated Values) file that contains log S values for each pair.

Data analysis

Predicted values were summarized by solvent, solute class and temperature. Figures were used to visualize the overall distribution of predicted log S values, the temperature dependence of selected cannabinoid classes and class-level comparisons at 298.15 K.

Results

Based on our results (Fig. 1), cannabinoids were predicted to dissolve most effectively in DMSO, acetone, ethanol, isopropyl alcohol, and methanol. Among them, DMSO exhibits the highest solubility due to its strong polarity and its ability to disrupt solute intermolecular forces. Conversely, hexane, heptane, naphtha, butane, propane, and R134a produced negative log S values, indicating poor solubility despite their lipophilicity.

In addition, Coconut oil, Medium-chain triglyceride (MCT) oil, and butter showed medium solubility predictions, similar to methanol and CO₂ based on this model (36). Importantly, the industry uses MCT oil as a common carrier for cannabinoid extracts and those results match the predicted data values from the model. Additionally, predicted solubility increased consistently with temperature across all solute-solvent pairs, indicating an endothermic dissolution process, consistent with thermodynamic behaviors observed experimentally (37). As an example, solubility (log S) for CBD with MCT oil was predicted computationally as 273.15 K = -0.05, 298.15 K = + 1.18, 323.15 K = + 2.23 respectively.

In addition, predictions for sesame oil and olive oil yielded two extreme log S values. CBD in sesame oil at 273.15 K appeared nearly insoluble (log S = -151.21), whereas CBD in olive oil at 273.15 K was predicted to be unrealistically soluble (log S = + 28.74).

According to Figs. 2 and 3, acid-cannabinoids such as Cannabinolic acid (CBNA), Δ^9 -Tetrahydrocannabinolic acid (Δ^9 -THCA), 9(R)-Hexacannabinol acid (9(R)-HHC Acid) and 9(S)-

Hexacannabinol acid (9(S)-HHC Acid) exhibited higher solubility than natural cannabinoids including Cannabidiol (CBD), Δ^9 -Tetrahydrocannabinol (THC), Cannabinol (CBN), Cannabigerol (CBG), Cannabifuran (CBF), and Cannabidivarin (CBND). This is likely due to the presence of a carboxylic acid group, which enhances interactions with polar solvents.

Importantly, Fig. 4 shows that all three Hexahydrocannabinol stereoisomers (9(R)-HHC, 9(S)-HHC, and iso-Hexahydrocannabinol (iso-HHC) exhibited similar solubility. As an example, at 273.15K, 9(R)-HHC, 9(S)-HHC and iso-HHC have had log S values of + 1.62, + 1.62, and + 1.54 with acetone, respectively. The minimal differences in log S values for stereoisomers indicates a minimal stereochemical influence on solubility. Notably, the non-cannabinoid acetazolamide and melatonin showed positive log S value in polar solvents such as DMSO and ethanol but negative values in nonpolar hydrocarbon solvents.

Discussion

For CBD the experimental data indicated a significant increase in solubility in n-hexane and n-heptane over a temperature range between 253K and 310K. Furthermore, this experimental data can be used to estimate the molar solubility at 298K by comparison with the predicted SolProp values for CBD (14). Therefore, the computationally predicted values aligned well with the experimental data according to previous studies.

According to previous literature observations, the solubility trend for CBD was best in DMSO, then acetone, then alcohols, and least in hydrocarbons (DMSO > acetone > alcohols > hydrocarbons). This solubility pattern is mainly due to the differences in solvent polarity and hydrogen bonding (13). Even though DMSO is an aprotic polar solvent, cannabinoids still dissolve. Because it has a high polarity index and dielectric constant, it allows it to effectively dissolve polar inorganic salts and many organic compounds (37). According to the previous studies, CBD was most soluble in MCT oil when compared to other tested plant oils and oils richer in C18:3 fatty acids also indicated high solubility values. Furthermore, even higher solubility values were identified in several hydrophobic eutectic systems and from them L-menthol/C9:phenyl system provided the best results (31). Therefore, by comparing the Solprop predicted solubility values for cannabinoids with previous data, a strong correlation was observed for DMSO, acetone, alcohols, hydrocarbons, and MCT oil.

Even though DMSO achieved the highest log S value of + 4.27, with CBG at 323.15 K, it is not practical for large scale production because its higher toxicity limits its application. On the other hand, ethanol and isopropanol are considered ICH Q3C Class 3 solvents, due to their relatively low toxicity values (38). By comparing the extraction data from studies with the computed values from this experiment, we identified the non-polar solvents that achieved greater cannabinoid selectivity than ethanol. An interview with Arvida Labs indicates that heptane is widely used for CBD crystallization and purification. Industrial observations indicate that because THC is more soluble than CBD in heptane at 298.15 K, CBD crystallizes. Furthermore, CBD was precipitated while THC remained in the liquid phase, maintaining consistent precipitation.

At the interview at the Arvida Labs, they described their experimental conversion of CBD to Δ^8 -THC in heptane at approximately 100°C (373.15 K), both CBD ($\log S = 0.16$ at 323.15 K) and THC ($\log S = 1.6$ at 323.15 K) dissolved readily in heptane at elevated temperature, supporting the predicted computational values in this research; furthermore, previous studies supported the claim (14, 39).

A study evaluating the suitability of carrier oils for cannabinoids suggested that MCT oil is the highest-capacity lipid vehicle among other physically plausible carrier oils. Because the long chain triglycerides produce larger mixed micelles during intestinal digestion (31). However, some studies related to cannabinoid bioavailability showed that a formulation with the highest solubility in vitro does not always yield the best absorption in the body. Therefore, the optimal carrier is not always the one with the highest solubility. Thus, formulation design should consider both thermodynamic solubility and how the oil behaves during digestion through further studies (40, 3)

There are a few limitations related to the Solprop model solvation prediction. For cannabinoids in complex oils, Solprop can yield unrealistic results. This problem applies specifically to olive oil and sesame oil. However, the model is more accurate only when triglyceride-level structures are used, such as triolein for olive oil and trilinolein for sesame oil. Finally, supercritical CO₂ is used as a solvent in some industries. However, Solprop predicted the Log S values at CO₂ gaseous. However, they should not be directly compared with industrial supercritical CO₂ extraction results unless phase behavior is corrected first (41).

Limitations

The study is computational and did not include new shake-flask or chromatographic validation experiments. Furthermore, lipid oils and supercritical CO₂ are complex systems that may fall outside the most reliable use case for SolProp.

Conclusion

This study presents a detailed algorithmic assessment of temperature-dependent solubility for an extensive array of cannabinoids in industrially relevant solvents. Here, a substantial collection of datasets around 1,026 log S values was generated utilizing an automated workflow using the Solprop solvation tool. It allows a systematic evaluation of solute-solvent interactions across different temperatures.

The results indicate that the degree to which cannabinoids dissolve depends on the solvent's polarity and temperature. Furthermore, aprotic polar solvents, such as DMSO and acetone, are very good at dissolving cannabinoids, while non-polar solvents, such as hexane and heptane, indicate poor solubility.

The study also revealed a clear association between temperature and solubility; as temperature increases, cannabinoids dissolve more readily in all systems studied. This indicates that heat promotes

the dissolution process of cannabinoids by facilitating molecular interactions, which aligns with the principles of thermodynamics.

The molecular structure of cannabinoids also affects their dissolution. Acidic cannabinoids dissolve slightly better than neutral cannabinoids due to their increased interaction with polar solvents. In contrast, stereochemistry, the spatial arrangement of molecules, does not seem to play a significant role compared to functional groups. For example, stereoisomers of the HHC molecule dissolve at similar rates. Therefore, among lipid-based systems, MCT oil, coconut oil, and butter exhibit good cannabinoid dissolving capabilities, explaining their frequent use in cannabinoid product formulations.

However, Solprop computational tools had some limitations while predicting how well cannabinoids would dissolve in lipid systems like olive oil and sesame oil. The predictions were not realistic, which shows the importance of development of a more advanced machine learning model, which can be trained and tested as a direct solubility predictor on a much larger, more diverse dataset, rather than predicting several intermediate thermodynamic quantities first.

Overall, the results of this study provide large-scale solubility datasets that are important for cannabinoid industrial applications, including chemical extraction, purification, and formulation process optimization. These results will provide a foundation for future advancement in cannabinoid research and development. Therefore, this study helps determine the solubility of cannabinoids before experimental investigations, reducing cost, time, and effort.

Abbreviations

BCS: Biopharmaceutics Classification System

CBD: cannabidiol

CBF: cannabifuran

CBG: cannabigerol

CBN: cannabinol

CBNA: cannabinolic acid

CBND: cannabinodiol

CBN-OAc: cannabinol acetate

CI/CD: continuous integration/continuous deployment

COSMO-RS: conductor-like screening model for real solvents

DMSO: dimethyl sulfoxide

HHC: hexahydrocannabinol

HSP: Hansen solubility parameters

ICH: International Council for Harmonisation

MCT: medium-chain triglyceride

RMG: Reaction Mechanism Generator

SMILES: Simplified Molecular Input Line Entry System

THC: tetrahydrocannabinol

THCA-A: tetrahydrocannabinolic acid A

UNIFAC: UNIQUAC functional-group activity coefficients

YAML: YAML Ain't Markup Language

Declarations

Corresponding author

Suliman Sharif

Email: sharifsuliman1@gmail.com

Ethics approval and consent to participate

Not applicable to the computational solubility screen. The study did not use human participants, human data, human tissue, animal subjects or animal data.

Consent for publication

Not applicable. The manuscript does not contain individual person's data, images or videos.

Availability of data and materials

Available data as a CSV dataset containing 1,026 generated log S predictions was produced.

Competing interests

Suliman Sharif, Mohammadhu Rifaydheen Sakeer, Katukoliya Gamage Anuththara Samadhi, Shashika Kalanaprabhath Kotawalagedara, Pamalka Hansani Kaldera, David Aronson, Damiola Bodun and, Mike Woster are affiliated with Standard Seed Corporation, which funded this study. Evan Friedmann is affiliated with Helma USA, Kevin Kimmel is affiliated with Arvida Labs, and Manny Maminakis is affiliated with Cannmetrix. These organizations collaborated by providing scientific expertise and technical consultation relevant to industrial cannabinoid processing and solvent selection.

All the authors declare that these affiliations and the funding source did not influence the study design, data analysis, interpretation of the data, or the decision to publish the results. All the findings of this manuscript are based on the computational workflow. All the authors contributed to the study ensuring the accuracy and integrity of the work.

Funding

Funding is provided by the Standard Seed Corporation. The funder supported the research did not influence the computational analysis, result interpretation or neither the decision to publish the study findings.

Authors' contributions

Suliman Sharif conceived the study, secured funding, supervised the project while coordinating the research collaboration. He contributed to the study design, results interpretation while revising the manuscript critically and approved the final version of the manuscript.

Mohammadhu Rifaydheen Sakeer, Katukoliya Gamage Anuththara Samadhi, Shashika Kalanaprabhath Kotawalagedara, and Pamalka Hansani Kaldera contributed to the development of the computational research pipeline, data collection, data curation, literature review, computational solubility prediction, data analysis, result interpretation, figures and tables preparation. They worked together to draft the manuscript, reviewed for final submission and approved the final manuscript.

Damiola Bodun, Mike Woster, and David Aronson contributed their scientific expertise to the research providing research guidance and supervision while reviewing the manuscript critically and approved the final version of the manuscript.

Evan Friedmann (Helma USA), Kevin Kimmel (Arvida labs) and Manny Maminakis (Cannametrix) provided their technical expertise, industrial consultation, solvent selection guidance and knowledge relevant to cannabinoid extraction and processing. They contributed to the study development by giving their feedback.

All the authors contributed intellectually to the study, reviewed and approved the final manuscript. They agree to be accountable for their respective contributions and responsible for ensuring that questions

related to the accuracy and integrity of their work.

Acknowledgments

The authors acknowledge Kevin Kimmell, Manny Maminakis, and Evan Friedmann (Arvida Labs, Fort Lauderdale, Cannmetrix) for expert consultation on industrial cannabinoid solvent usage and process chemistry, forming the basis for the validated solvent panel. The Authors also acknowledge the MIT RMG and SolProp developers for providing the publicly accessible computational tools that supported this work

References

1. Radwan MM, Chandra S, Gul S, ElSohly MA. Cannabinoids, Phenolics, Terpenes and Alkaloids of Cannabis. *Molecules*. 2021;26(9):2774.
2. Graziano S, Vari MR, Pichini S, Busardo FP, Cassano T, Di Trana A. Hexahydrocannabinol Pharmacology, Toxicology, and Analysis: The First Evidence for a Recent New Psychoactive Substance. *Current Neuropharmacology* [Internet]. 2023;21(12):2424–30. Available from: <https://www.ingentaconnect.com/content/ben/cn/2023/00000021/00000012/art00006>
3. Stella B, Baratta F, Della Pepa C, Arpicco S, Gastaldi D, Dosio F. Cannabinoid Formulations and Delivery Systems: Current and Future Options to Treat Pain. *Drugs*. 2021 Sept;81(13):1513–57.
4. Docampo-Palacios R, Tesfatsion GA, Cruces TT. W. Cannabinoid Ethers: Identification, Binding interactions, and their ADME properties. *Research Square* [Internet]. 2024 June 5 [cited 2026 Apr 6];Volume 1. Available from: https://assets-eu.researchsquare.com/files/rs-4523212/v1_covered_0dd806ab-4771-4201-8fc3-5efa21439d70.pdf?c=1717562988
5. Felletti S, Compagnin G, Yannick Krauke, Stephan S, Greco G, Buratti A, et al. Purification and Isolation of Cannabinoids: Current Challenges and Perspectives. *LCGC Europe*. 2023;36(Issue 04):122–31.
6. Aymonier C, Gilles Philippot, Croguennec L, Petit E, Poulon-Quintin A. 20th European Meeting on Supercritical Fluids: Book of Abstracts. HAL (Le Centre pour la Communication Scientifique Directe). 2024 July 24;1st edition.
7. Zarghampour A, Jouyban K, Jouyban-Gharamaleki V, Jouyban A, Rahimpour E. A Description on the Shake-Flask and Laser Monitoring-Based Techniques for Determination of the Drug's Solubility. *Pharmaceutical Sciences* [Internet]. 2024 Feb 14 [cited 2024 Aug 26];30(2):274–8. Available from: <https://ps.tbzmed.ac.ir/Inpress/ps-40027.pdf>
8. Vassileiou AD, Robertson MN, Wareham BG, Mithushan Soundaranathan, Ottoboni S, Florence AJ, et al. A unified ML framework for solubility prediction across organic solvents. *Digit Discovery*. 2023;2(2):356–67.
9. Obaidullah AJ, Mahdi WA. Machine learning-based analysis on pharmaceutical compounds interaction with polymer to estimate drug solubility in formulations. *Scientific Reports*. 2025 July

2;15(1).

10. Johnson MS, Dong X, Alon Grinberg Dana, Chung Y, Farina D, Gillis RJ, et al. RMG Database for Chemical Property Prediction. *J Chem Inf Model*. 2022;62(20):4906–15.
11. Vermeire FH, Chung Y, Green WH. Predicting Solubility Limits of Organic Solutes for a Wide Range of Solvents and Temperatures. *J Am Chem Soc*. 2022 June;10(24):10785–97.
12. Attia L, Burns JW, Doyle PS, Green WH. Data-driven organic solubility prediction at the limit of aleatoric uncertainty. *Nature Communications* [Internet]. 2025;16(1). Available from: <https://www.nature.com/articles/s41467-025-62717-7>
13. Tavčar E, Vidak M. Experimental investigation and thermodynamic modelling of cannabidiol and curcumin in different solvents. *Journal of Molecular Liquids* [Internet]. 2024 July 16;410:125511. Available from: <https://www.sciencedirect.com/science/article/pii/S016773222401568X>
14. Gabriela L, Orozco GA, Chaves G. Gerardo Rodríguez Niño. Experimental Solubility of Cannabidiol (CBD) in Different Organic Solvents. *Industrial Eng Chem Res* 2023 July 25;62(31):12356–64.
15. Kanabus J, Bryła M, Roszko M, Modrzewska M, Pierzgalski A. Cannabinoids—Characteristics and Potential for Use in Food Production. *Molecules* [Internet]. 2021;26(21):6723. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8588477/>
16. Gülck T, Lindberg Møller B, Phytocannabinoids. Origins and Biosynthesis. *Trends in Plant Science* [Internet]. 2020;25(10):985–1004. Available from: <https://www.sciencedirect.com/science/article/pii/S1360138520301874>
17. Appendino G. The early history of cannabinoid research. *Rend Lincei Scienze Fis e Naturali*. 2020;31(4):919–29.
18. Russo F, Maria A, Vandelli, Biagini G, Schmid M, Luongo L, Perrone M et al. The semisynthetic cannabinoid Hexahydrocannabinol (HHC). *Research Square* (Research Square). 2023.
19. Nasrallah DJ, Garg NK. Studies Pertaining to the Emerging Cannabinoid Hexahydrocannabinol (HHC). *ACS Chem Biol*. 2023;18(9).
20. Xie Z, Mi Y, Kong L, Gao M, Chen S, Chen W et al. *Cannabis sativa*: origin and history, glandular trichome development, and cannabinoid biosynthesis. *Horticulture research* [Internet]. 2023 July 26;10(9). Available from: <https://academic.oup.com/hr/article/10/9/uhad150/7231120>
21. Gallardo A, Gutierrez MR, Gomez LA, Arielle P, Dones SA, Dumbrique MM, et al. A Comparative Analysis on the Potential Anticancer Properties of Tetrahydrocannabinol, Cannabidiol, and Tetrahydrocannabivarin Compounds Through In Silico Approach. *Asian Pac J Cancer Prev*. 2024;25(3):839–56.
22. Abu-Sawwa R, Stehling C. Epidiolex (Cannabidiol) Primer: Frequently Asked Questions for Patients and Caregivers. *J Pediatr Pharmacol Ther*. 2020;25(1):75–7.
23. Desa S, Osman A, Hyslop R. Silico Assessment of Drug-Like Properties of Phytocannabinoids in Cannabis Sativa. *EDUCATUM J Sci Math Technol*. 2017;4(2):1–7.

24. Jaipakdee N, Tabboon P, Limpongsa E. Application of a liquisolid technique to cannabis sativa extract compacts: Effect of liquid vehicles on the dissolution enhancement and stability of cannabinoids. *Int J Pharm.* 2022;612:121277.
25. Pankaj Singh Cham, Kotwal P, Sharma K, Dhiman S, Singh L, Varun P, Singh et al. Cannabidiol-Based Prodrugs: Synthesis and Bioevaluation. *ACS Med Chem Lett.* 2024.
26. Tiago FJ, Paiva A, Matias AA, Duarte ARC. Extraction of Bioactive Compounds From Cannabis sativa L. Flowers and/or Leaves Using Deep Eutectic Solvents. *Front Nutr.* 2022;9.
27. Stasiłowicz-Krzemień A, Szulc P, Judyta Cielecka-Piontek. Co-Dispersion Delivery Systems with Solubilizing Carriers Improving the Solubility and Permeability of Cannabinoids (Cannabidiol, Cannabidiolic Acid, and Cannabichromene) from Cannabis sativa (Henola Variety) Inflorescences. *Pharm* 2023 Sept 4;15(9):2280–0.
28. Ramella A, Roda G, Pavlovic R, Dei Cas M, Casagni E, Mosconi G et al. Impact of Lipid Sources on Quality Traits of Medical Cannabis-Based Oil Preparations. *Molecules* [Internet]. 2020;25(13):2986. Available from: <https://www.mdpi.com/1420-3049/25/13/2986>
29. Sumontri S, Eiamart W, Tadtong S, Samee W. Ultra-High-Performance Liquid Chromatography–Tandem Mass Spectrometry Analysis of Δ^9 -Tetrahydrocannabinol and Cannabidiol in Commercial Suk-Saiyasna Herbal Remedy: Applying Hansen Solubility Parameters for Sample Extraction to Ensure Regulatory Compliance. *Pharmaceuticals* [Internet]. 2024 Nov 8 [cited 2026 Apr 6];17(11):1502. Available from: <https://www.mdpi.com/1424-8247/17/11/1502>
30. Moodley K, Rarey J, Ramjugernath D. Model evaluation for the prediction of solubility of active pharmaceutical ingredients (APIs) to guide solid–liquid separator design. *Asian J Pharm Sci.* 2018;13(3):265–78.
31. Alhadid A, Luca SV, Nasrallah S, Minceva M. Experimental investigation and thermodynamic modeling of cannabidiol solubility in plant oils and hydrophobic eutectic systems. *Journal of Molecular Liquids* [Internet]. 2022;372:121172. Available from: <https://www.sciencedirect.com/science/article/abs/pii/S0167732222027118?via%3Dihub>
32. Koll MM, Menardi R, Hughes D, Schwark WS, Wakshlag JJ. Physiological effect and pharmacokinetic evaluation of combined oral administration of cannabidiolic acid and cannabigerolic acid in dogs. *Front Veterinary Sci.* 2026;13.
33. Tahir MN, Shahbazi F, Rondeau-Gagné S, Trant JF. The biosynthesis of the cannabinoids. *J Cannabis Res.* 2021;3(1).
34. Grapp M, Kaufmann C, Peschel A, Potzscher M, Dziadosz M, Marquenie L, et al. Detection of 11-nor-9-carboxy-hexahydrocannabinol (HHC-COOH) as metabolite of both hexahydrocannabinol (HHC) and Δ^9 -tetrahydrocannabinol (Δ^9 -THC) in routine forensic samples. *J Pharm Biomed Anal.* 2025;261:116833.
35. Kundu D, Franchini L, Hasdemir HS, Lloyd E, Maturano J, Rabl K et al. Distinct Interactions of Cannabinol and Its Cytochrome P450-Generated Metabolites with Receptors and Sensory Neurons.

Journal of medicinal chemistry [Internet]. 2025;68(13):13935–53. Available from: <https://pubmed.ncbi.nlm.nih.gov/40568800/>

36. Falcón-Cano G, Molina C, Cabrera-Pérez MA. ADME Prediction with KNIME: In silico aqueous solubility models based on supervised recursive machine learning approaches. *ADMET and DMPK*. 2020.
37. Santos NC, Figueira-Coelho J, Martins-Silva J, Saldanha C. Multidisciplinary utilization of dimethyl sulfoxide: pharmacological, cellular, and molecular aspects. *Biochem Pharmacol*. 2003;65(7):1035–41.
38. López-Olmos C, García-Valverde MT, Hidalgo J, Ferrerio-Vera C, Sánchez de Medina V. Comprehensive comparison of industrial cannabinoid extraction techniques: Evaluation of the most relevant patents and studies at pilot scale. *Front Nat Prod*. 2022;1.
39. Huang S, Van Beek T, Claassen F, Janssen HG, Chen B, Zuilhof H et al. Comprehensive compositional analysis after conversion of CBD to Δ^8 -THC as a 2 controversial ingredient for cannabinoid-infused edibles. 2023 June.
40. Izgelov D, Davidson E, Barasch D, Regev A, Domb AJ, Hoffman A. Pharmacokinetic investigation of synthetic cannabidiol oral formulations in healthy volunteers. *Eur J Pharm Biopharm*. 2020 Sept;154:108–15.
41. Arul J, Tardif R, Boudreau A, McGinnis DS, Lencki RW. Solubility of milk fat triglycerides in supercritical carbon dioxide. *Food Res Int*. 1994;27(5):459–67.

Figures

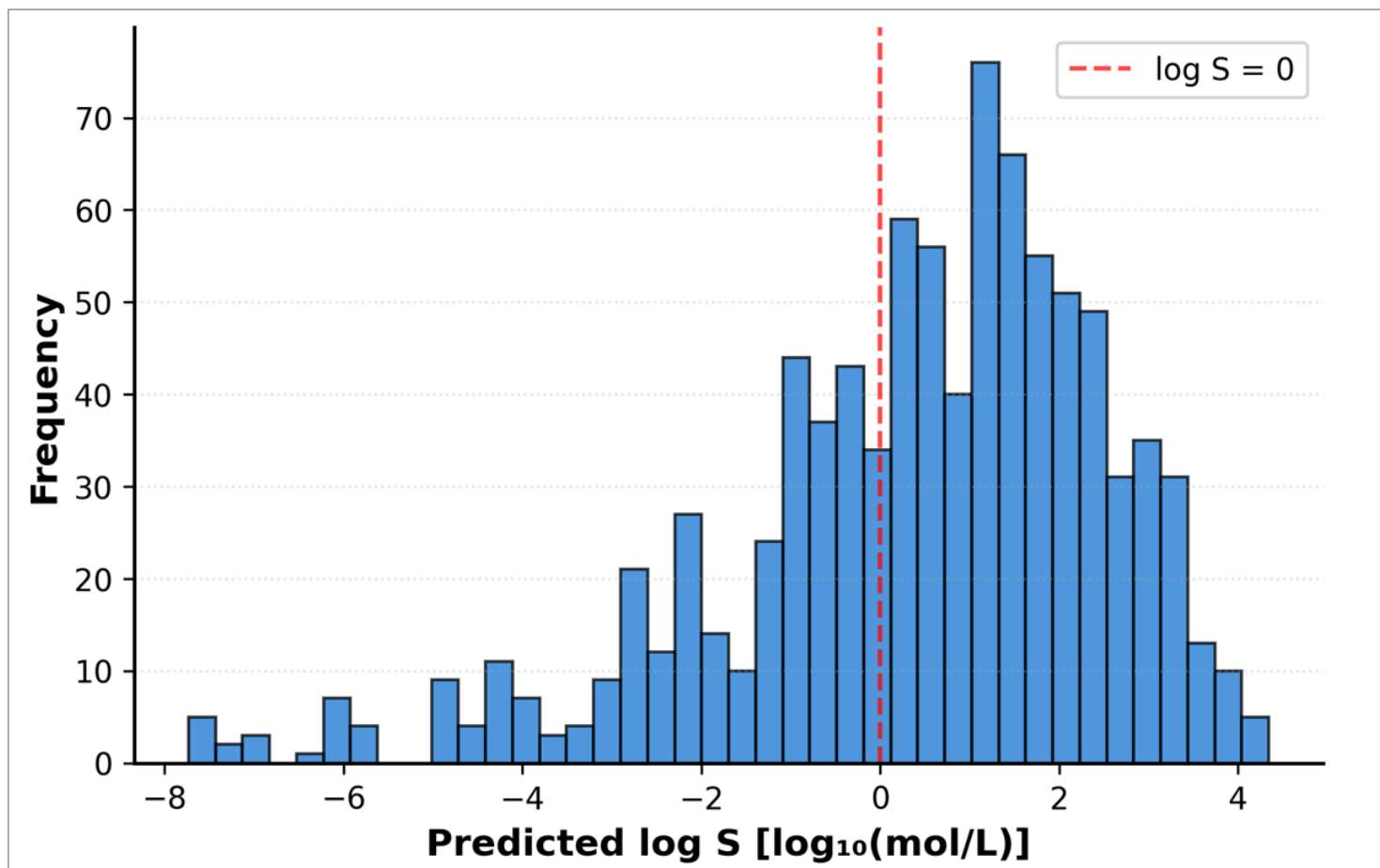


Figure 1

Distribution of solubility

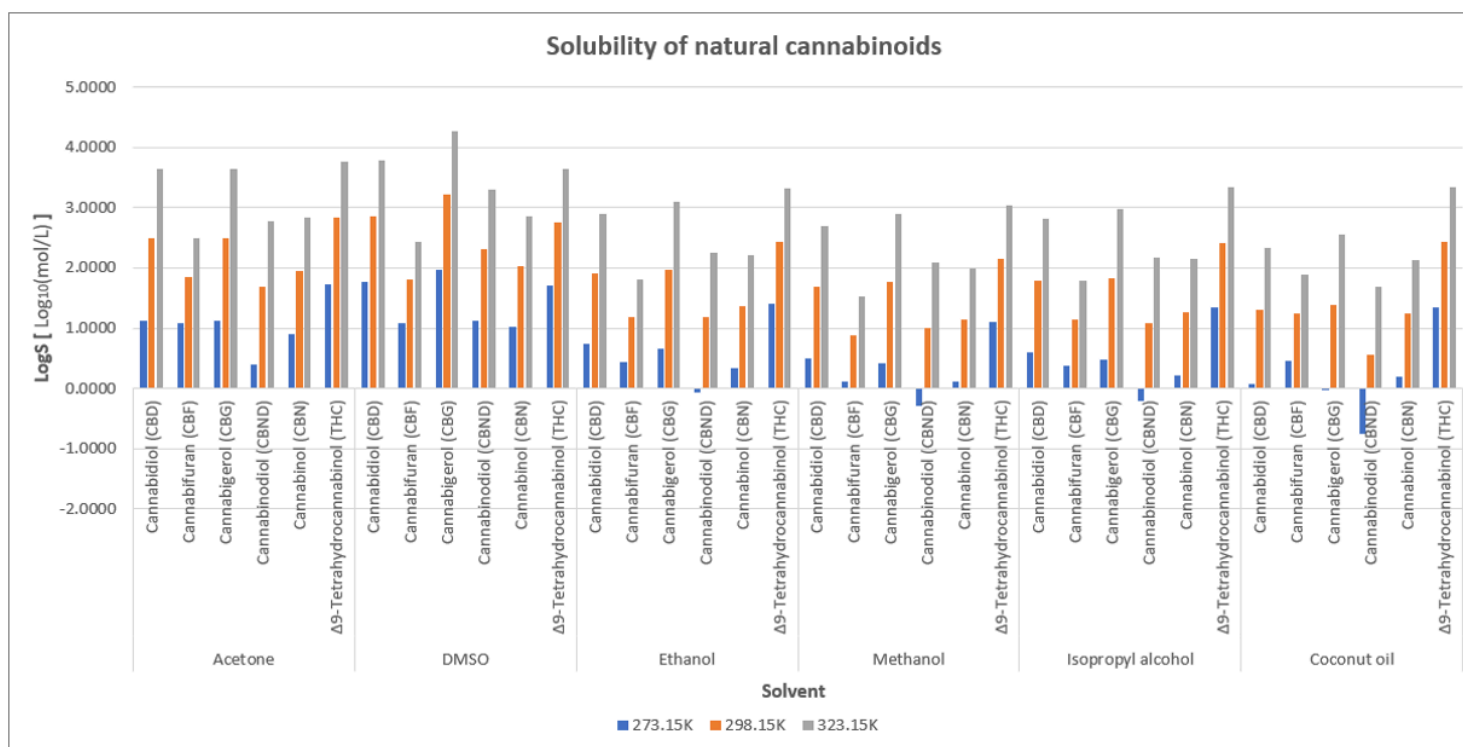


Figure 2

Solubility of natural compounds at temperature; 273.15K, 298.15K and 323.15K

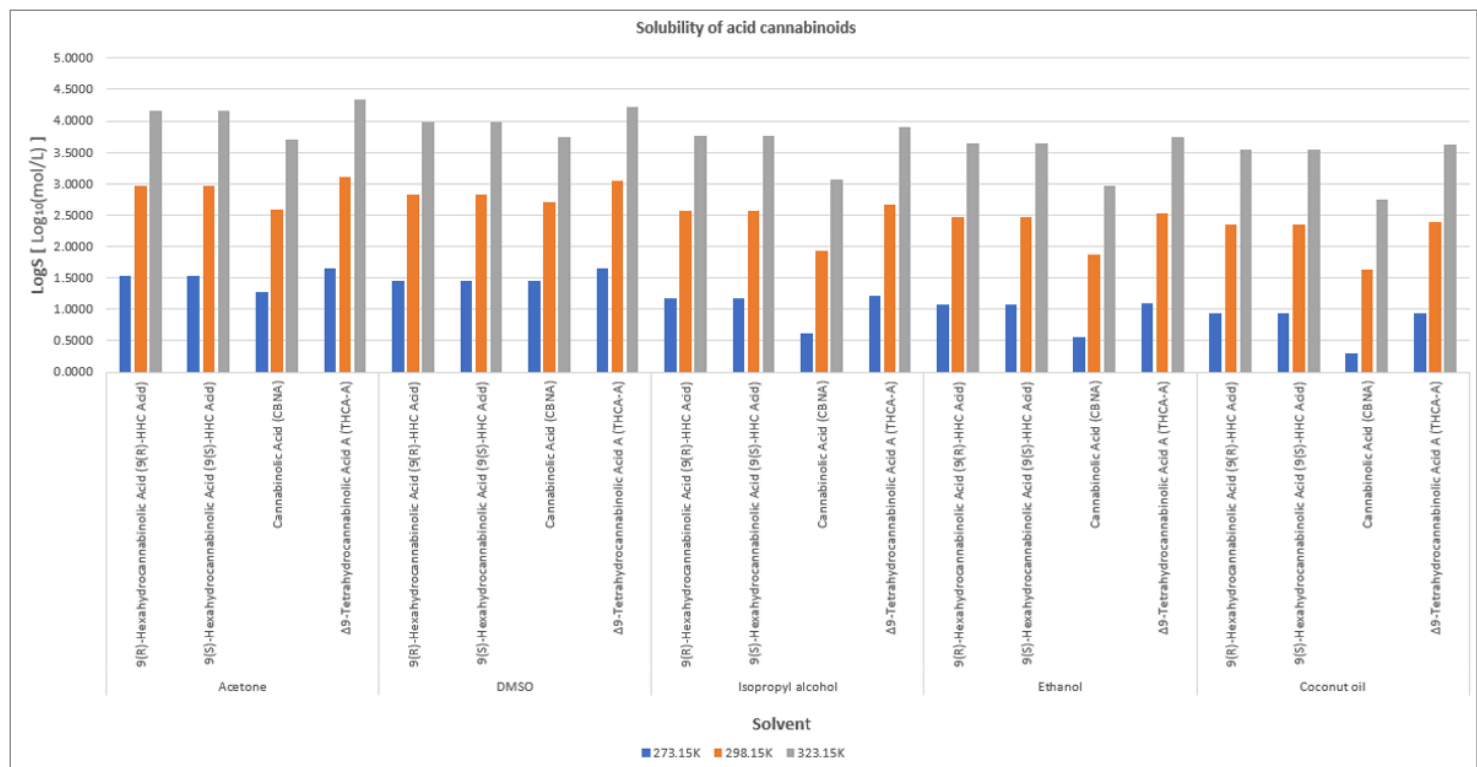


Figure 3

Solubility of acidic compounds at temperatures: 273.15K, 298.15K and 323.15K

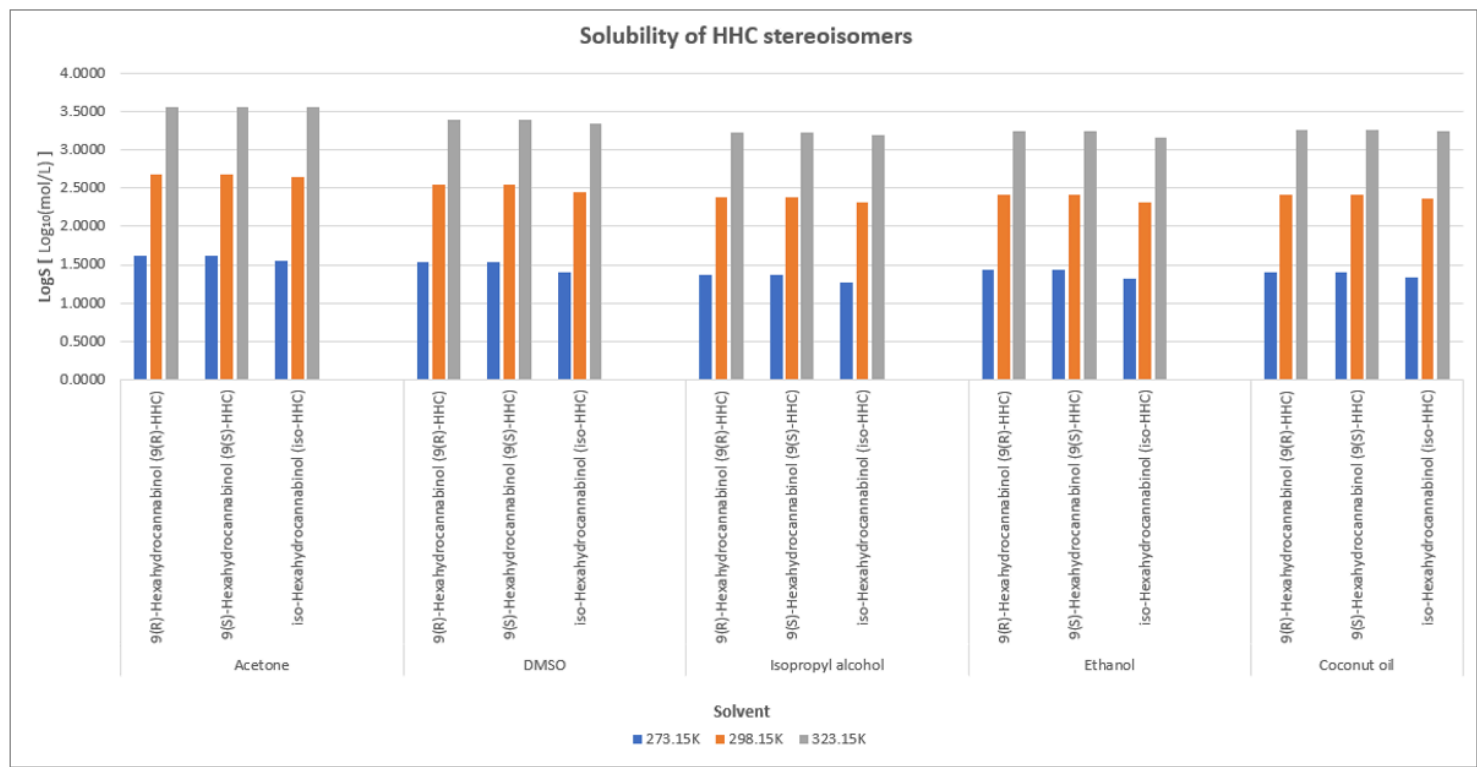


Figure 4

Solubility of HHC isomeric cannabinoids at temperatures: 273.15K, 298.15K and 323.15K

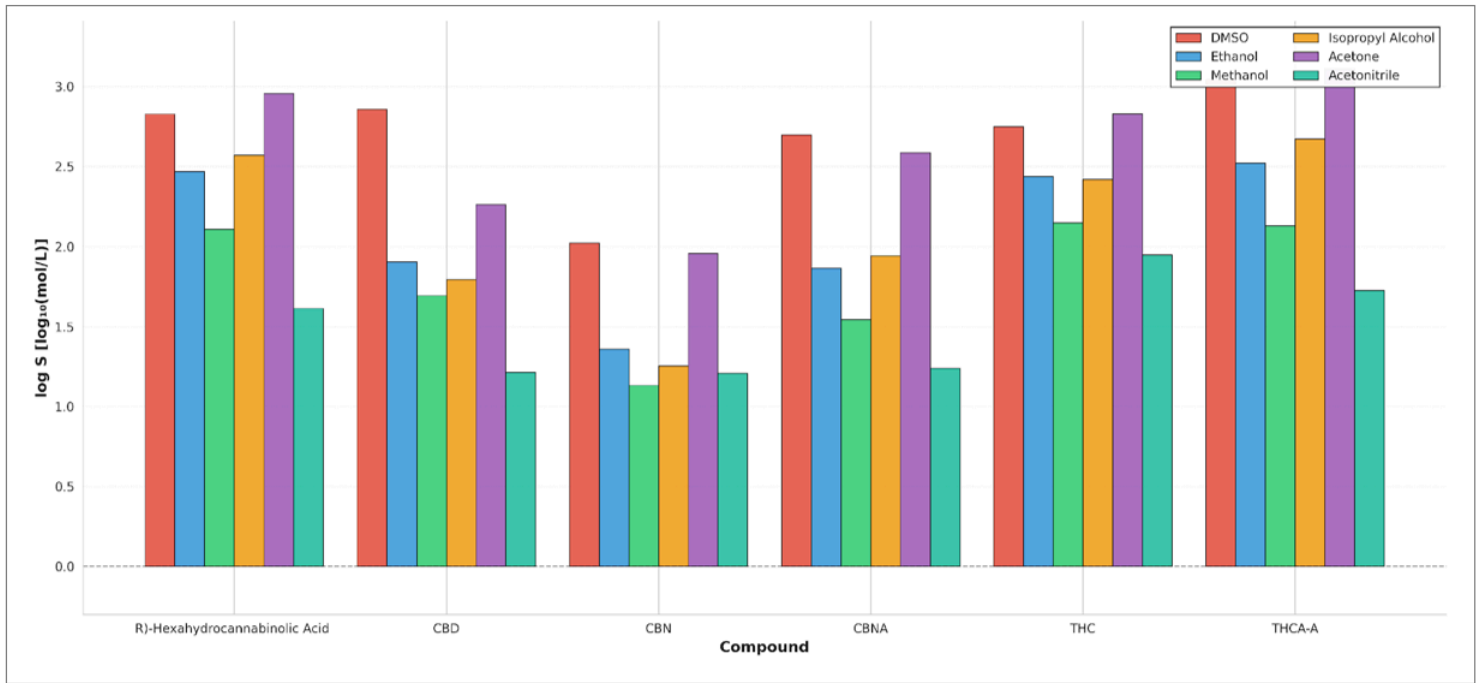


Figure 5

Grouped comparison of mean solubility (log S) for cannabinoid classes at 298.15 K in polar solvents.