



**HAL**  
open science

# Gelation properties of various long chain amidoamines: Prediction of solvent gelation via machine learning using Hansen solubility parameters

Frederic Delbecq, Guillaume Adenier, Yuki Ogue, Takeshi Kawai

## ► To cite this version:

Frederic Delbecq, Guillaume Adenier, Yuki Ogue, Takeshi Kawai. Gelation properties of various long chain amidoamines: Prediction of solvent gelation via machine learning using Hansen solubility parameters. *Journal of Molecular Liquids*, 2020, 303, pp.112587 -. 10.1016/j.molliq.2020.112587 . hal-03489684

**HAL Id: hal-03489684**

**<https://hal.science/hal-03489684>**

Submitted on 21 Jul 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

# Gelation Properties of Various Long Chain Amido Amines: Prediction of Solvent Gelation via Machine Learning using Hansen Solubility Parameters

Frederic Delbecq<sup>a\*</sup>, Guillaume Adenier<sup>b</sup>, Yuki Ogue<sup>c</sup>, Takeshi Kawai<sup>c</sup>

<sup>a</sup> Ecole Supérieure de Chimie Organique et Minérale (ESCOM), EA TIMR 4297 UTC/ESCOM, 1 allée du Réseau Jean-Marie Buckmaster, F-60200 Compiègne, France. \*E-mail : [f.delbecq@escom.fr](mailto:f.delbecq@escom.fr)

<sup>b</sup> 2 rue des Pommerelles, F-60200 Compiègne, France.

<sup>c</sup> Department of Industrial chemistry, Tokyo University of Science, 1-3 Kagurazaka, Shinjuku, Tokyo 162-8614, Japan

**ABSTRACT:** Four new amphiphilic long chain amidoamine derivatives displaying different structure variations are synthesized and tested in 27 liquids and compared to the study of two similar molecules already reported in the literature. In many cases, these compounds can act as low molecular weight gelators to form a three-dimensional network in organic liquids or water, which can be confirmed by FE-SEM observations and rheology measurements. For each sample, XRD diffraction of the corresponding xerogel and FT-IR analysis of native supramolecular gels reveal that they can self-assemble into lamellar-like aggregates or in pseudo-cubic structures, depending on the alkyl chain length and the steric hindrance of the polar head. The number of amide bonds and their positions inside gelator structures are determinant for the nature of the packing. For each gelator, we perform a series of gelation tests in each of the solvents and show that Hansen parameters, which are known characteristics of each liquid, can be used to successfully predict their gelation properties via machine learning in the vast majority of liquids at a concentration of 4 wt %.

**Key-words:** Amphiphilic compounds, supramolecular gelators, Hansen solubility parameters, gelation prediction models.

## 1. Introduction

During the last decades, the interest for the production of low molecular weight gelators has increased rapidly. These gelators are molecules that can self-assemble in liquids to form fibers of sheet-like aggregates when subjected to unidirectional non-covalent forces such as hydrogen bonds,  $\pi$ - $\pi$  stacking or Van der Waals hydrophobic interactions [1]. The resulting supramolecular fibers get entangled into a 3D-network that entraps the solvent (organic liquid or water) within its interstitial spaces. It is important to remark that the gelation phenomenon often occurs in an intermediate state between solubility and insolubility in a given solvent. In fact, independently of the relative hydrophilic-lipophilic balance (HLB) values of a good gelator, many conditions and factors are responsible for its gelation ability, such as the nature of the solvent, the possible supersaturation of the liquid, which often cause aggregation-type variations, the kinetics of the gelation phenomenon, and of course the molecular structure of the gelator [2].

Among gelators, modified small lipopeptides [3,4] or long amidoamines displaying a large variety of different polar heads have been studied for various applications such as drug delivery systems, lithium-ion batteries, cosmetics, or oil-spill treatment. Gelators are sometimes intended as templates for the generation and stabilization of nanosized materials of noble metals [5-11]. For industrial purposes, it is essential to be able to understand and predict the gelation abilities of gelator candidates in liquids of particular interest for specific applications.

In this work, we investigate the impact of small and localized modifications of two lead compounds **1** and **2** recognized as long amidoamine gelators [5,7], which have good relative gelation capability on the packing nature of formed aggregates in standard liquids such as toluene or water. The structure of all molecules is summarized in Fig. 1.

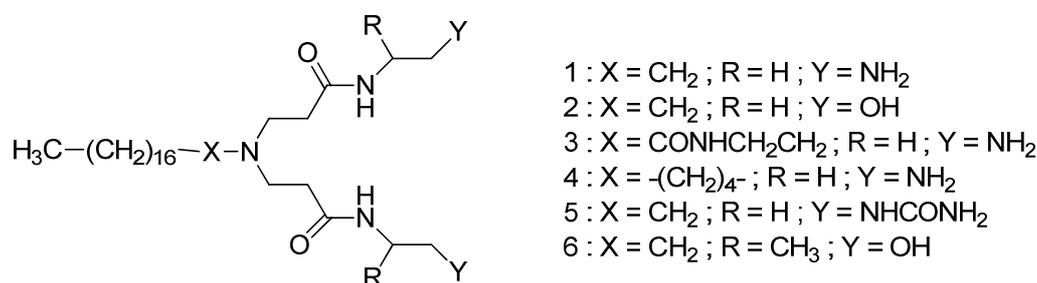


Fig. 1. Structure of organogelators 1 to 6.

Looking at the structure of the new compounds, that is, gelators **4-6**, the most striking features are the following. The elongation of the fatty tail by the replacement of the current C18 alkyl chain with a longer alkyl chain of 22 atoms of carbon for gelator **4**, the conversion of terminal primary amine groups of gelator **1** into two urea moieties for gelator **5** and the introduction of a methyl group on both sides of the polar head of gelator **2**, closed to the OH extremities in gelator **6**, but also the presence of an additional amide bond inserted in the structure of gelator **3**.

After performing gelation tests on gelators **1-6** and confirmed the real gel nature of selected sample by rheology, we observe the morphology of each gel sample using field emission scanning microscopy (FE-SEM). Studying the respective air-dried xerogel of toluene gel samples provides information on the strength of the intermolecular interaction by Fourier Transform Infrared Spectroscopy (FT-IR). X-ray diffraction (XRD) of xerogels is expected to provide enough indications on the possible packing nature formation. To the best of our knowledge, no comprehensive and predictive study has been performed to date on families of gelators with long fatty alkyl chains except in the works of Bouteiller et al [12] or Rogers [22], where he tried to predict the gelation behavior in edible oils of 34 compounds grafted with at least one long alkyl chain, using the Hansen Solubility Parameters via a predictive model.

The second part of this study is inspired by an article by Yan *et al* [13] and by other recent works [14-24] that have demonstrated how 3-D spherical Hansen plots can be used to predict variations in gelation capabilities inside a family of compounds. Hansen solubility parameters (HSPs) have already been employed to predict the solubility of different material

in liquids such as polymers [25], small organic compounds [26-30], organometallic complexes [31] and fatty-based biodiesels [32] through a machine learning approach. These parameters provide a remarkable tool to build a comprehensive understanding of our series of gelators and to predict their gelation ability before being tested. Given that HSPs are known for all solvents, in particular for those reported in this work, it is thus possible to quantify their potential to interact with gelators using their dispersion ( $\delta_d$ ), dipole-dipole or polar ( $\delta_p$ ), and hydrogen bonding ( $\delta_H$ ) interactions.

Three types of region of interest can be schematically identified in HSP space: solubility (S), gelation (G) and insolubility (I) regions. They are empirically determined by gelation tests consisting of mixing a small amount of the gelator being considered with various solvents. The result of each gelation test (S, G, or I) with a particular solvent is marked in HSP space and placed according to the HSP coordinates of the solvent. The set of all gelation tests thus reveal regions in HSP space with similar gelation results. The results of a gelation test with an unknown solvent can thus be determined according its location inside a previously identified region in HSP space. These regions are typically represented as spheres, which can be concentric or not.

However, in many cases, this spherical approach is too rigid to represent complex gelating features, in particular for high polar solvents such as polyols or diols which are often found outside the main gelation sphere. A question thus remains unsolved. Is it possible to give a complete prediction of the gelation ability of a family of compounds without synthesizing each of these derivatives? A good matching between real tests and simulation could provide an answer to this question.

## 2. Experiment

### 2.1. Materials

All chemicals that we used were reagent grades obtained from Aldrich, Acros Chemical or TCI Japan. Commercially available reagents were used without further purification, except in the case of methyl acrylate furnished by Kanto Chemicals, which was purified by distillation under reduced pressure in a nitrogen atmosphere, and in the case of octadecylamine, which was recrystallized two times starting from hexane.

### 2.2. Organic synthesis

Gelators **1** and **2** were synthesized using common procedures reported in the literature [5, 7]. Gelator **3** was prepared using the same protocol as reported in a previous paper [8].

NMR spectra were recorded in deuterated solvents such as DMSO-*d*6 or CDCl<sub>3</sub> with a Bruker 400 Ultrashield Spectrometer operating at 400 MHz. (See Figure S1 of the ESI for <sup>1</sup>H NMR charts of compounds 3-6). Chemical shifts ( $\delta$ ) are quoted in ppm and are referenced to TMS (tetramethylsilane) as an internal standard. Coupling constants (*J*) are quoted in Hz, common splitting patterns and their abbreviations were s (singlet), d (doublet), t (triplet), q (quartet), and m (multiplet). Mass spectra were recorded on Bruker Daltonics micro TOF Focus mass spectrometer.

For the gelators 1 and 2, their purities were confirmed by  $^1\text{H}$  NMR and mass spectroscopy.

#### 2.2.1. Synthesis of *N*-(2-amino-ethyl)-3-[[2-(2-amino-ethylcarbamoyl)-ethyl]-behenylamino] propionamide 4

To 50 mL of toluene were added 5.5 g (16.17 mmol) of behenic acid. Then, to the resulted dispersion was added dropwisely 10.26 g (80.78 mmol) of oxalyl chloride and the complete mixture was stirred refluxing for two hours under argon atmosphere. The solvent and unreacted reagents were both removed under reduced pressure to produce colored oil which was introduced in 50 mL of THF without further purification. In the next step, 20 mL of 25 %  $\text{NH}_4\text{OH}$  aqueous solution was added to the previous clear yellow solution and a white solid started to precipitate which was successively filtrated on Buchner, washed with a slight volume of water and THF. (mass = 5.43 g ; yield = 99 %) The recovered behenamide powder was dried in a dessicator and remained insoluble in common organic solvents. FT-IR: 3396  $\text{cm}^{-1}$  (NH) ; 2849-2916  $\text{cm}^{-1}$  ( $\text{CH}_2$ ) ; 1647 and 1526  $\text{cm}^{-1}$  (amide). MS : Calcd ( $\text{M}+\text{H}^+$ ) 339 found 339.23.

In the next stage, 2.66 g (7.84 mmol) of the previous behenamide was dispersed in 40 mL of dry THF, followed by addition of 0.60 g (15.68 mmol) of Lithium aluminium hydride ( $\text{AlLiH}_4$ ). Thus, the resulted dispersion was allowed to stir refluxing for 1 day. After returned to the room temperature, the solution was poured in water and extracted with several volumes of  $\text{CHCl}_3$  and diethyl ether. All organic layers was separated and combined before the drying on  $\text{MgSO}_4$ . The final organic solution was evaporated in vacuum to produce a white solid. (mass = 1.1 g ; yield = 43 %).  $^1\text{H}$  NMR ( $\text{CDCl}_3$ ) :  $\delta$  0.85 (t, 3H,  $J = 1.62$  Hz) ; 1.1 (m, 40 H,  $\text{CH}_2$ ) ; 3.50 (t, 2H,  $J = 6.3$  Hz,  $\text{CH}_2\text{NH}_2$ ). MS : Calcd ( $\text{M}+\text{H}^+$ ) 325 found 325.13.

By using the synthesis protocol reported for **1**, the diester intermediates was obtained as follow: 1.1 g (3.38 mmol) of the previously synthesized amine was introduced in 30 mL of MeOH, followed by the addition of 1.45 g (16.86 mmol) of methylacrylate. The resulted solution was stirred at 60°C for 4 hours, a slightly amount of white solid precipitated which was later filtrated. Both solvent and unreacted acrylate was removed by distillation in normal pressure to produce clear oil that recrystallized at room temperature. The generated diester was introduced without further purification in 30 mL of MeOH containing 7.0 g (0.11 mol) of ethylene diamine. The resulted solution was heated up at 70°C for one day, all liquids were removed under vacuum to give an oil which crystallized into a white cream solid. (mass : 1.02 g ; yield = 55 % for two steps)

$^1\text{H}$  NMR ( $\text{CDCl}_3$ ):  $\delta$  0.88 (t, 3H,  $J = 1.44$  Hz) ; 1.25 (br, 36 H,  $\text{CH}_2$ ) ; 1.42 (br, 4H,  $\text{CH}_2\text{CH}_3$ ,  $\text{CH}_2\text{CH}_2\text{CH}_2\text{N}$ ) ; 2.37-2.82 (m, 14 H,  $\text{CH}_2\text{CH}_2\text{CO}$ ,  $\text{CH}_2\text{N}$ ,  $\text{NCH}_2\text{CH}_2$ ,  $\text{CH}_2\text{NH}_2$ ) ; 3.30 (m, 4H,  $\text{NHCH}_2$ ). MS : Calcd ( $\text{M}+\text{H}^+$ ) 553 found 553.17.

#### 2.2.2. Synthesis of [(octadecyliminobis)bis(ethylenecarbonylaminoethylene)]-1,1'-diurea 5

The diamine **1** (1.0 g, 2.01 mmol) and 0.60 g (4.42 mmol) of commercial phenylcarbamate (97 % Aldrich) were both dissolved in 40 mL of THF. The resulted solution was stirred at 60°C and 0.44 g (4.42 mmol) of triethylamine (TEA) was added before carry on the stirring for 4

hours. Once the heating was stopped, the solvent was removed in vacuum and the crude product was recrystallized from a mixture of MeOH and toluene to produce a yellow solid. (mass = 0.85 g ; yield = 72 %).

$^1\text{H}$  NMR ( $\text{CDCl}_3$ ) :  $\delta$  0.88 (t, 3H,  $J = 7.0$  Hz,  $\text{CH}_2\text{CH}_3$ ) ; 1.25-1.63 (br, 24H,  $\text{CH}_2$ ), 2-16-2.73 (m, 14H,  $\text{CH}_2\text{CH}_2\text{CONH}$ ,  $\text{NCH}_2\text{CH}_2\text{CO}$ ,  $\text{CH}_2\text{CH}_2\text{N}$  and  $\text{NCH}_2\text{CH}_2\text{CO}$ ) ; 3.27 (m, 8H,  $\text{CH}_2\text{CH}_2\text{NHCONH}_2$  and  $\text{NHCH}_2\text{CH}_2$ ). MS : Calcd ( $\text{M}+\text{H}^+$ ) 583.2 found 573.17.

### 2.2.3. Synthesis of *N*-(2-hydroxy-1-methylethyl)-3-[[2-(2-hydroxy-1-methylethylcarbamoyl)-ethyl]octadecylamino]propionamide 6

The diester precursor or more exactly the 3-[(2-methoxycarbonyl-ethyl)-(octadecylamino)]propionic acid methyl ester (3.0 g, 6.80 mmol) was introduced in 15 mL of MeOH, then 1.03 g (13.80 mmol) of (*S*)-(+)-2-aminopropanol was added and the resulted mixture was stirred for one week at 60°C. The methanol was removed by evaporation and the resulted solid was triturated in acetone and recovered by filtration on buchner before drying in vacuum. (mass = 2.8 g ; yield = 78 %).

$^1\text{H}$  NMR ( $\text{CDCl}_3$ ) :  $\delta$  0.88 (t, 3H,  $J = 7.2$  Hz,  $\text{CH}_3$ ) ; 1.16 (d, 6H,  $\text{CHCH}_3$ ) ; 1.25-1.42 (br, 32 H,  $\text{CH}_2$ ) ; 2.36-2.73 (m, 10 H,  $\text{CH}_2\text{CH}_2\text{CO}$ ,  $\text{CH}_2\text{N}$  and  $\text{NCH}_2\text{CH}_2\text{CO}$ ) ; 3.40-3.68 (2m, 4H,  $\text{CH}_2\text{OH}$ ) ; 4.96 (t, 2H,  $\text{NHCH}(\text{CH}_3)\text{CH}_2$ ). MS : Calcd ( $\text{M}+\text{H}^+$ ) 527 found 528.45.

### 2.3. Gelation tests

Gelators **1-6** taken at different concentrations were mixed in a test tube with a precise volume of a selected solvent and heated up until a clear transparent solution was obtained. The resulting solution was left to cool down at room temperature. After a few minutes, if the initially viscous or liquid solution became a gel-like material, the sample was turned upside down for confirmation. The sample was considered as a gel if the sample remained perfectly immobilized. In other cases, it was recognized as soluble, insoluble or simply a gelatinous precipitate.

### 2.4. Characterizations of gels

Rheology measurements were realized using a Physica HCR 301 (Anton Paar) rotational rheometer equipped with a disk and a cone forming the measuring system suited for high viscosity samples. The rheometer is also equipped with a Peltier temperature control system that can regulate the temperature in a range from - 40 to 200°C. The measurements were monitored by the Rheoplus software. All experiments were performed at 20°C and the viscosity was checked for frequencies ranging from 0.1 to 10 Hz. The viscoelasticity of selected samples was characterized in terms of the elastic modulus ( $G'$ ) and the loss modulus ( $G''$ ). X-ray diffraction (XRD) measurements were performed with a Rigaku Ultima IV diffractometer using an air-dried sample of 2 wt % supramolecular gels. An FT-IR was conducted using a spectrometer (Thermo-Fisher Scientific Nicolet 6700 FT-IR) equipped with an MCT detector with a resolution at 4  $\text{cm}^{-1}$ . A liquid cell with ZnSe windows was employed for organogels, and a Harrick GATR attenuated total reflection (ATR) attachment was used for hydrogels. A demountable liquid cell with a  $\text{CaF}_2$  window was employed for solvent solution FT-IR measurements. Then, one thousands scans were recorded on average for

each sample spectrum. Scanning electron microscopy was carried out using a Jeol JCM-5700 microscope. The xerogels were subjected to Au-Pd Sputtering (Hitachi E-1010 sputter coater) before starting the experiments. Differential Scanning Calorimetry (DSC) was conducted on a DSC 8MC Mettler-Toledo using aluminium pans. Scans were conducted under nitrogen with a heating rate of 10°C/min in the temperature range of 40-250°C.

### 3. Results and discussion

#### 3.1. Evaluation of gelation ability for all gelators (1-6)

The gelation ability of gelators **1-6** was examined for a large panel of organic liquids, as well as for water. (Some gel samples are given as example in the ESI, Fig. S2) The gelation tests were calibrated as follows. For each sample a mass of 0.04 g was introduced in 1.0 mL of solvent and heated up until the material was entirely dissolved. When the sample remained insoluble in the liquid, it was recorded as such. When dissolved, the resulting transparent or clear yellow solution was left to cool down at room temperature. It then turned into a more viscous or turbid solution until a gel started to appear, within a short period of time. Depending on the gelator, the formed gel could be stable for up to three months without syneresis, or could simply precipitate within a few hours. In order to confirm the gelation ability of each gelator 1-6, the corresponding 4 wt % gel samples in cyclohexane were employed for testing their rheological properties. All samples were observed as strong elastic gels throughout the entire frequency range with higher  $G'$  values always over 10000. Here,  $G'$  values were systematically higher than  $G''$  and always parallel with the respect to the angular frequency. The gelation ability of gelators **1-6** was tested in a large variety of solvents, as reported in Table 1. It should be mentioned that the gelation ability of gelators **1, 2** and **3** has already been partially described in the literature [5,7, 8].

Table 1. Gelation ability for gelators 1-6 in water and various organic liquids introduced in a liquid at 4 wt % with their respective Hansen solubility parameters values (MPa<sup>1/2</sup>).

|                  | $\delta_d$ | $\delta_p$ | $\delta_H$ | 1 | 2 | 3 | 4 | 5 | 6 |
|------------------|------------|------------|------------|---|---|---|---|---|---|
| Toluene          | 18.0       | 1.4        | 2.0        | G | G | G | I | G | G |
| 2-Methylfuran    | 17.3       | 2.8        | 7.4        | I | S | I | G | I | S |
| Benzene          | 18.4       | 0          | 2.0        | G | G | G | I | G | G |
| Pyridine         | 19.0       | 8.8        | 5.9        | G | S | S | S | I | S |
| Acetophenone     | 19.6       | 8.6        | 3.7        | S | S | S | G | I | I |
| Aniline          | 19.4       | 5.1        | 10.2       | S | S | I | G | S | S |
| m-Cresol         | 18.0       | 5.1        | 12.9       | S | S | S | S | S | S |
| Cyclohexane      | 16.8       | 0          | 0.2        | G | G | G | G | G | G |
| n-Hexane         | 14.9       | 0          | 0          | G | I | G | G | I | G |
| Chloroform       | 17.8       | 3.1        | 5.7        | I | S | S | S | G | S |
| CCl <sub>4</sub> | 17.8       | 0          | 0.6        | G | G | G | G | G | G |
| THF              | 16.8       | 5.7        | 8.0        | S | S | I | S | G | S |
| Dibutylether     | 14.4       | 2.9        | 5.1        | G | I | I | G | I | I |
| Acetone          | 15.5       | 10.4       | 7.0        | I | S | I | I | I | I |
| MIBK             | 15.3       | 6.1        | 4.1        | S | S | S | I | I | I |
| Acetonitrile     | 15.3       | 18.0       | 6.1        | I | I | I | G | I | I |
| DMF              | 17.4       | 13.7       | 11.3       | I | S | I | G | G | S |
| DMSO             | 18.4       | 16.4       | 10.2       | I | S | I | G | G | S |

|                     |      |      |      |   |   |   |   |   |   |
|---------------------|------|------|------|---|---|---|---|---|---|
| Ethylacetate        | 15.8 | 5.3  | 7.2  | G | I | I | G | I | I |
| Ethyllevulinate     | 16.5 | 7.8  | 6.8  | I | I | I | G | I | I |
| Propylene carbonate | 20.0 | 18.0 | 4.1  | I | I | I | G | I | I |
| Methanol            | 15.1 | 12.3 | 22.3 | S | S | S | I | S | S |
| Butanol             | 16.0 | 5.7  | 15.8 | S | S | S | S | I | S |
| Ethylene glycol     | 17.0 | 11.0 | 26.0 | I | G | I | I | S | I |
| Glycerol            | 17.4 | 12.1 | 29.3 | G | G | G | G | G | G |
| Triethylamine       | 17.8 | 0.4  | 1.0  | G | I | I | I | I | i |
| Water               | 15.5 | 16.0 | 42.3 | S | G | S | I | I | G |

MIBK = Methylisobutyl ketone; THF = Tetrahydrofuran; DMF = Dimethyl formamide; DMSO =Dimethyl Sulfoxide ; S : soluble ; G : gel or opaque gel; I : insoluble.

### 3.2. Machine learning classification of HSP space

Predicting the formation of gels, such as displayed in Table 1, is a daunting task [13]. Hansen solubility parameters (HSP) are widely used to help classify how specific compounds react when immersed into solvents [14], and more specifically to attempt to predict the formation of gels [15].

As we will see, for gelators **1-6** the regions defining all three relevant classes (soluble, gel, insoluble) in HSP space, usually established by the calculation of a specific interaction radius (Ra) [32], are generally too complex to be able to be sketched with a simple spherical approach, be it concentric or not [16,17]. In order to provide useful insight into gel formation, we therefore chose to turn to *machine learning* approaches to classify HSP space for each gelator, whenever possible. This approach has already been used successfully to predict gel formation [18].

For this purpose, several machine learning methods have been compared and we have evaluated their performances not only by measuring their predictive abilities but also by assessing their interpretative usefulness. A common drawback of machine learning approaches is indeed that although they can be excellent at predicting classes of previously unseen input states, they are essentially black boxes providing results without explanation as to why a specific input state is predicted to belong to a certain class rather than to another. We thus require that the model should provide visualizations that favor human interpretation.

### 3.3. Predictive performances of various machine learning methods

We implement the machine learning classification with a code written in Wolfram Mathematica, with which we compare the performances of several methods: nearest neighbors (k-NN), Naïve Bayes, Random Forest, Logistic Regression, Neural Network, Support Vector Machine (SVM), Decision Tree, and Gradient Boosted Tree.

As can be seen in Table 1, the gelation properties of each gelator have been tested for each of the 27 potential solvents. The input data are the Hansen solubility parameters ( $\delta_d$ ,  $\delta_p$  and  $\delta_H$ ) in Table 1, and the output is the corresponding result of the gelation test in the column of the gelator being considered.

```

for each algorithm  $A \in \{\text{SVM}, \text{RF}, \text{NN}, \dots\}$ 
  initialize the number of correct predictions  $n_{\text{correct}}$ 
  for each solvent  $k$ :
    discard solvent data  $(\mathbf{X}_k, y_k)$  from dataset  $\{(\mathbf{X}_i, y_i), \forall i\}$ 
    train new classifier  $c_k$  on remaining data  $\{(\mathbf{X}_i, y_i), \forall i \neq k\}$ 
    predict class  $\tilde{y}_k$  of discarded solvent  $k$  as  $\tilde{y}_k = c_k(\mathbf{X}_k)$ 
    if  $\tilde{y}_k = y_k$  increment number of correct predictions  $n_{\text{correct}}$ 
  calculate accuracy as  $n_{\text{correct}}/N_{\text{solvents}}$ 

```

Fig. 2. Pseudocode for LOOC algorithm comparison, where the HSP  $(\delta_d, \delta_p, \delta_H)$  of each solvent  $i$  are taken as input  $\mathbf{X}_i$  for the classifiers and the 2-class response (Gel or non-Gel) of the gelator is taken as output  $y_i$ .

To mitigate the risk of over-fitting, we evaluate the predictive performances of each method using a leave-one-out cross-validation (LOOC) procedure, with a two-class distinction (gel or non-gel). It means that we take out one solvent from this list and train the machine learning algorithm with the remaining 26 solvents. Once the classifier has been trained with the set of 26 solvents, we test whether it can correctly predict the class of the solvent that was left out. We repeat this procedure with all possible 26 to 1 partitions and count how many times the method correctly predicted the class of the solvent that was left out. The pseudocode for this procedure is described in Fig. 2. The predictive performance displayed in Fig. 3 is the ratio of the number of correct predictions over the number of partitions (that is, 27). We use this procedure for all gelators and for all classification methods.

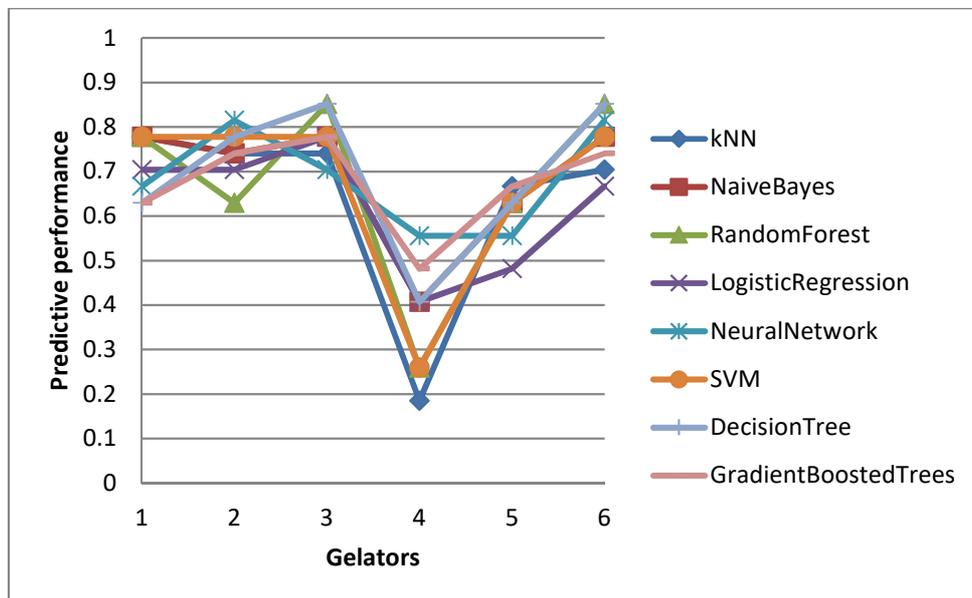


Fig. 3. Predictive performances of several machine learning methods, for all gelators 1-6. The performances are measured with a leave-one-out cross-validation (LOOC) method, with a two-class distinction (gel or non-gel). The performance of a fully random classifier would be equal to 0.5.

The performances of the different methods vary with each gelator (see Fig. 3). Since we consider only two output classes (gel or non-gel), the predictive performances should be

compared to that of a fully random classifier, that is, one that would attribute classes at random with a probability 1/2. As can be seen in Fig. 3, the performances of the classification methods are only consistently better than a random classifier for gelator **1, 2, 3** and **6**.

The predictive performance is particularly poor for gelator **4**. It can be attributed to the fact that the classes for this gelator are not grouped as clusters but rather intricately intertwined, with a probability to find a closest neighbor of the same class of only 0.37, very close to that of a random probability of 1/3.

The performances of the different methods appear to be very similar. For a specific gelator, the fact that one particular method performs better than others may very well be a mere coincidence. Choosing a different method for each gelator to improve performances might very well be equivalent to over-fitting. Similarly, adjusting the machine learning settings to maximize the classification accuracy and predictive performance for each gelator might very well lead to an overestimation of the performances of the classifiers.

### 3.4. Interpretation of results using Support Vector Machine method

Since all methods perform more or less equally well as predictive tools, we therefore chose the method that provides the most natural looking and easily interpretable visualizations. In this respect, the best choice appears to be that of Support Vector Machine (SVM). Other methods tend to either look blocky with straight lines that look rather contrived, or blurred with regions that are not easy to distinguish or interpret.

Since we want to be able to visualize not only the gelating region (G) but also the soluble (S) and insoluble (I) regions in HSP space, we train the SVM as a 3-class problem. We have measured the performance of the SVM classifier with LOOC validation and obtained an average accuracy of 54% for the six gelators (using a one-vs-all method, where each class is trained against all other classes). This lower accuracy of 54% should be compared with the 33% performance of a random classifier, so that the performance of the classifier is arguably not fundamentally reduced with respect to the 2-class performances. The pseudocode for the visualizing procedure is described in Fig. 4.

```

initialize 3D grid of  $n \times n \times n$  points  $\{X_{k,l,m}\}$  in HSP space
train SVM classifier on gelator dataset  $\{(X_i, y_i), \forall i\}$ 
get class probabilities  $p_S, p_G, p_I$  from SVM classifier for each grid point  $X_{k,l,m}$ 
for each class probability  $p \in \{p_S, p_G, p_I\}$ :
    generate 3D density plot for points  $X_{k,l,m}$  with  $p >$  threshold ( $\sim 0.33$ )
  
```

Fig. 4. Pseudocode for visual representation of gelating regions in HSP space, where the HSP ( $\delta_d, \delta_p, \delta_H$ ) of each solvent  $i$  are taken as input  $X_i$  for the classifier and the 3-class response (S, G, or I) of the considered gelator as output  $y_i$ .

It should be noted that some features are beyond any classification method and show the intrinsic limitations of the HSP method. The example of triethylamine is particularly telling. Indeed, for several gelators **2, 3, 5** and **6**, triethylamine is tightly surrounded by close neighbors in HSP space that are all pointing to a different class than that of triethylamine. There is strictly no way that a classification method, however advanced, would predict such

a feature that would almost appear to be an anomaly if the same behavior was not observed with multiple gelators.

Nearly all gelators, except gelator **4** and perhaps gelator **3**, exhibit two disconnected regions of gelation, which are very remote from one another in HSP space and separated by large regions associated with different classes. One gelation regions is typically associated with low polarity and low hydrogen bonding (near benzene and toluene), whereas the other is associated with high polarity and high hydrogen bonding (near water and glycerol).

Besides, compared to the classic approach of Hansen space found in the literature [26, 33], due to non-spherical nature of each zone on our graphics, it appears not easy to interpret the relationship between volume expressed by each respective gelation zone and some different empiric results such as kinetic or undercooling without consistent and specific values of Ra.

### 3.4.1. Gelator 1

Gelator **1** is known for its strong gelating capability in aromatic solvents such as benzene or toluene. This behavior is confirmed in Fig. 5 (left). With high polar solvents, for example dimethylsulfoxide or DMF, the solid remains insoluble even when introduced at a lower concentration of 0.2 wt %. Interestingly, it is able to form a gel in hydrogen-bond acceptor liquids, such as triethylamine or pyridine that can also be considered as aromatic solvents. Surprisingly, even though ethylene glycol (EG) and glycerol are both water soluble polyols, gelator **1** can form a gel in the former but not in the latter. This phenomenon might be due to the smaller number of OH groups in EG, leading to more interactions between gelator **1** and its environment. In addition, a second gelation region that could correspond to unknown solvents or liquid synthetic polyols appears in Fig. 5, but given that the classifier seem to identify this region based on the known class of glycerol alone, it should rather be interpreted as a potential region of interest that would call for further tests rather than as a prediction.

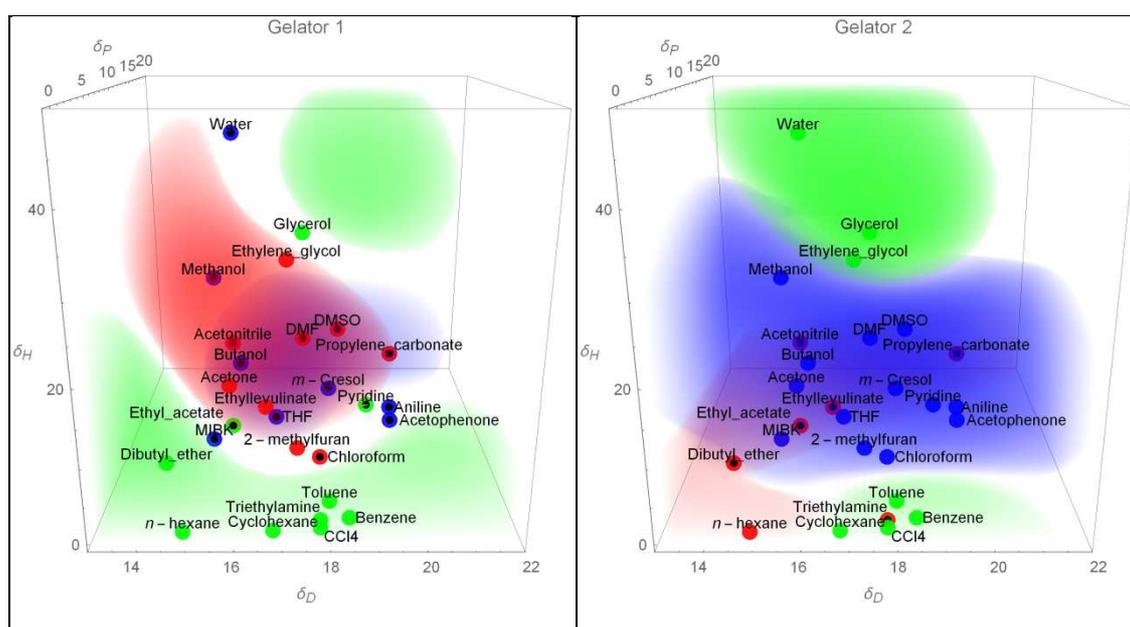


Fig. 5. Gelation classification in HSP space for gelator 1 (left) and gelator 2 (right). Solvents are marked with different colors depending on their gelation properties: soluble (blue), gel (green) or insoluble (red). The regions predicted by the classifier (SVM) are displayed in the same colors (blue, green and red) as probability clouds for all three classes. The opacity of the clouds intensifies with the predicted probability above that of a fully random classifier. Solvents that are misclassified by the classifier are marked with a center black disk.

#### 3.4.2. Gelator 2

Gelator **2** is shown in Fig. 5 (right). It is characterized by the replacement of the two terminal primary amine groups by OH moieties, which impacts only slightly the performances of the derived gelator. This molecule is almost insoluble in common organic solvents, although it can produce a gel in ethylene glycol, glycerol and water at a lower concentration of 1 wt %. In addition, gelator **2** is capable of gelating non H-donor aromatic solvents such as benzene, as well as more hydrophobic liquids such as short alkanes or CCl<sub>4</sub>.

The particular case of triethylamine is interesting. Given its HSPs values, one would have expected that it behaves like other apolar liquids. The contribution of its non-binding pair of nitrogen atom could however explain the differences observed with its neighbors in HSP space, with the idea that different binding modes of a gelator could occur inside gel networks, depending on the nature of the solvent employed for each test. By contrast with gelator **1**, with its strong hydrogen bonding between adjacent OH groups, gelator **2** could self-assemble into some type of lipidic bilayer exposing either its hydrophobic tail or polar head according the nature of the host liquid. Furthermore, these two possible modes of assembly could arguably explain the presence of two distinct three dimensional gelation spaces in the corresponding HSP solubility plot.

#### 3.4.3. Gelator 3

The introduction of an additional amide group in gelator **3** can be interesting on several accounts, as can be seen in Fig. 6 (left). Firstly, this gelator is expected to provide superior hydrophilic-lipophilic balance value (HLB) compared to gelator **1**, which could provide a simple explanation to the better water solubility of the solid. Secondly, the gelation abilities of gelators **1** and **3** are similar, except in nitrogen rich liquids where gelator **3** shows difficulties in being solvated, which might be caused by the inhibition of the formation of successful hydrogen bonds by these solvents. Thirdly, gelator **3** is unexpectedly almost insoluble in all water miscible solvents, but it can nevertheless effectively act as a gelator in glycerol at a lower concentration of 2 wt %. This phenomenon could be explained by the normal viscosity of glycerol, which means that gelator **3** could possibly be acting as a thickener, without playing the traditional role of a gelator.

#### 3.4.4. Gelator 4

As shown in Fig. 6 (right), gelator **4** is obviously more hydrophobic than gelator **1** due to the longer C22 alkyl chain in its structure. Unlike with previous gelators, we were able to obtain opaque gel samples that were stable in unusual solvents such as acetonitrile, acetophenone, 2-methylfuran or DMF, even with only a small amount of this gelator. Curiously, this molecule seems to be a good candidate to produce gels in the most polar water soluble

solvents (DMF, DMSO, and acetonitrile) but also in non-water miscible liquids, especially in ketone, as well as in the smallest ethyl esters. In this case, the total HLB value was heavily decreased. We will discuss later in the paragraph 3.5 of the real contribution of the polar head in the packing, but the balance was definitively broken for this material. However, if we superpose both models for gelators **1** and **4**, it seems clear that the gelation ability of gelator **4** is complementary to that of gelator **1**.

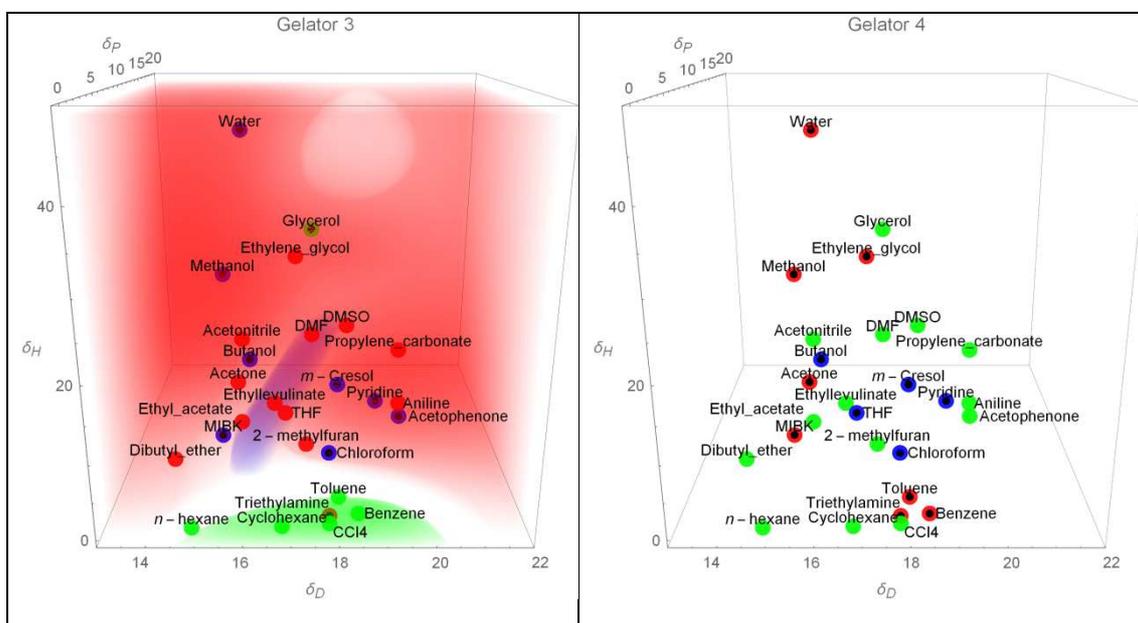


Fig. 6. Gelation classification in HSP space for gelator 3 (left) and gelator 4 (right). Solvents are marked with different colors depending on their gelation properties: soluble (blue), gel (green) or insoluble (red). Here, the classifier experiences some difficulties tackling with the complexity of the gelation properties in HSP space, especially for gelator 4, as it misclassifies the whole HSP space as a gel.

### 3.4.5. Gelator 5

As shown in Fig. 7 (left), gelator **5** exhibits three distinct and nested regions in HSP space. This behavior is arguably a consequence of the introduction of urea moieties on both extremities of the polar head of the gelator. Due to its high probability to form a large number of additional intermolecular hydrogen bonding in complicated networks, gelator **5** is surprisingly not a good candidate to form a gel in the majority of the liquids used in this study, even in MeOH or pyridine. As was the case with gelator **3**, the HSP solubility plots of gelator **5** shows that increasing the amide bonds inside the structure of the gelator leads to a drastic decrease of its solubility in various types of solvents. Obviously, not only the number but their positions in relation to the hydrophilic head are responsible of the serious solubility changes that could be potentially controlled.

### 3.4.6. Gelator 6

Finally, the solubility of gelator **6** shown in Fig. 7 (right) is affected by the presence of two methyl groups on the periphery of the polar head, near the OH groups. The head of the molecule is more hydrophobic, and overall bulkier. This behavior is however different than that of gelator **2**, as it might prevent the formation of head-to-head complexes between two

molecules promoted by intermolecular hydrogen bonding, which is the main cause of hydrogen bond weakening. Fig. 5 (right) shows that, independently of the chirality of the extremities of gelator **6**, the held methyl groups have only a small influence on the gelation ability of this gelator, and that although the solubility region is somewhat reduced, it remains overall consistent with that of gelator **2**, for which two clear gelation regions are visible.

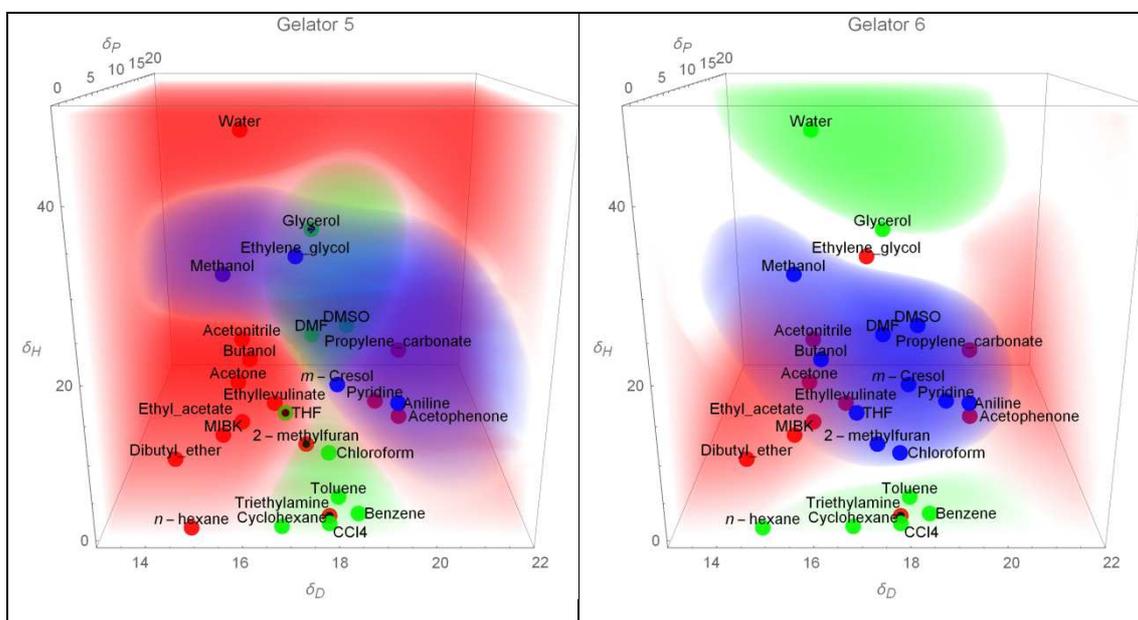


Fig. 7. Gelation classification in HSP space for gelator 5 (left) and gelator 6 (right). Solvents are marked with different colors depending on their gelation properties: soluble (blue), gel (green) or insoluble (red). The regions predicted by the classifier (SVM) are displayed in the same colors (blue, green and red) as probability clouds for all three classes. The opacity of the clouds intensifies with the predicted probability above that of a fully random classifier. Solvents that are misclassified by the classifier are marked with a center black disk.

### 3.5. FE-SEM observations of the xerogels

Using gelators **1-6** at a concentration near their Gelation Minimum Concentration (GMC), we obtained electron microscope images for air-dried self-assembled xerogels in toluene organogels, except for gelator **4** where ethylacetate was used. In Fig. 8, in all images, the 3D-networks are made of long thin fibers with average diameter of 1  $\mu\text{m}$ , except for gelator **4**.

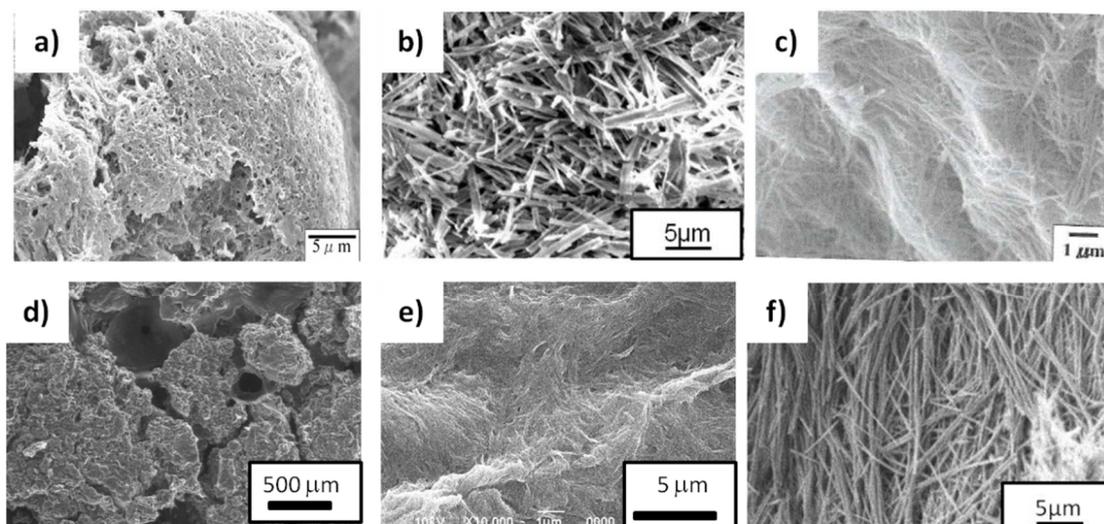


Fig. 8. FE-SEM images of air-dried organogels in toluene obtained for gelators 1-3, 5, 6, in images a-c) to e) and f) respectively, and d) in ethylacetate for gelator 4.

For gelators **1** and **3**, the fibers are very thin compared to those obtained from gelators **2** and **6**. For these two diol-type gelators, it is reasonable to assume that two different binding modes occur inside their gel matrix. Indeed, microscopic observations of the xerogels obtained from gelators **2** and **6** in water show alternative surface morphology. That is, gelator **2** gives a visibly weaker and altogether different structure network than the long crystalline rods observed with gelator **6**.

For gelator **4**, the difficulties that we experienced in producing a gel from toluene convinced us to use an ethylacetate (AcOEt) xerogel instead. Fig. 8 d) shows a gel matrix networks with a visibly thick enough net to entrap the solvent, which could be directly connected to the length of the fatty alkyl chain grafted on the remaining unchanged polar heads. Finally, for gelator **5**, a typical entangled three-dimensional  $\beta$ -sheet network was formed in toluene, with less apparent cavities.

In a recent work, Rogers and his collaborators have attempted to develop a correlation between fiber lengths and the Hansen sphere radius or the  $\delta_H$  values of respective solvents [33]. In our case, it appears difficult to get a specific exploitable value like radius  $R_a$  and we can only interpret the volume of the gelation area surrounding the toluene location on the graphics, especially for gelators **2**, **3**, **5** and **6**. The values of the volumes are given from our calculation models in arbitrary unity with values of 78.2, 42.6, 115.3 and 74.6 respectively for **2**, **3**, **5**, and **6**. Interestingly, the values obtained from **5** were recorded as higher than the others because of two nearest gelations lobes connected via a narrow area. In regards to these results, it seems to be possible that the volume of a gelation area could become enough informative. Thus, it is possible to imagine a direct connection with the thickness and length of the networks fiber. For example, for gelators **2** and **6**, we got the closest values and similarities on the respective SEM pictures. By decreasing the value of the gelation space around the toluene position, we observed thinner and longer fibers for **3**.

### 3.6. IR spectroscopy of gel samples

As depicted in Fig. 9, FT-IR spectra of gel samples were measured in toluene- $d_8$  to determine the representative bands of physical gel, except for the more hydrophobic sample **4** where DMSO- $d_6$  was used because of its lack of solubility in apolar benzene-type solvents.

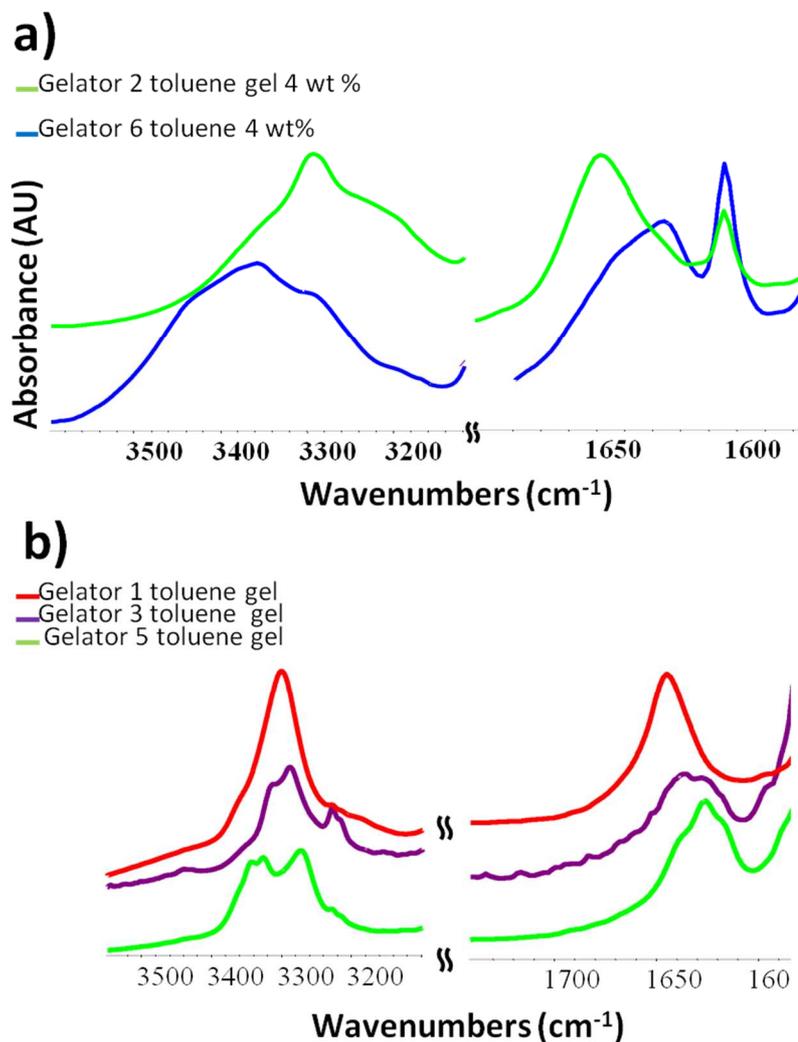


Fig. 9. FT-IR spectra: a) Toluene- $d_8$  gels of gelators 2 and 6 (on left: NH bands variations; on right: amide (I) band variations); b) Toluene- $d_8$  gels of gelators 1, 3 and 5 (on left: NH bands variations ; on right: amide (I) band variations).

For all gels samples, typical absorption bands that are characteristic of trans alkyl chains were generally observed at a value superior to  $2900\text{ cm}^{-1}$ , which might be considered as evidence that hydrophobic packing participates in the stabilization of the entire structure of xerogels. Compared to gelator **6**, gelator **2** shows lower OH vibration bands located at  $3294\text{ cm}^{-1}$ , with a broad band found between  $3300$  and  $3400\text{ cm}^{-1}$ . For gelator **6**, this result can be caused by the methyl group on intermolecular H-bonding networks, precisely between OH extremities of two adjacent molecules. Moreover, lower value of amide (I) bond vibration also confirms the creation of strong packing between compounds, which is a consequence of stronger hydrogen bonding networks. This situation results in higher **crystallinity** of the material associated with an increase of its GMC.

For gelator **3**, the introduction of another inserted amide bond played an important role in gel formation. Compared to **1**, the lower signal at  $1637\text{ cm}^{-1}$  can be seen as evidence of a stronger packing and of a tendency to form crystals in a large variety of organic liquids. Furthermore, the contribution of the additional amide group appears with two distinct bands located at  $3360$  and  $3282\text{ cm}^{-1}$ , respectively. However, in case of gelator **5** the signal was decomposed into distinct bands that also attributed to urea moieties. For the amine signal, there are three consecutive NH bands located at  $3453$ ,  $3325$  and  $3283\text{ cm}^{-1}$ . For the amide (I), they are located at  $1659$  and  $1639\text{ cm}^{-1}$ , which suggests that a part of the terminal amide function, remains free and that the gelator only gets involved in a simple interdigitated structure.

For gelator **4**, the NH vibration band is interestingly as high as  $3458\text{ cm}^{-1}$ . This value provides evidence of the lower contribution to the networks of terminal amine group components of polar heads. Hydrophobic interactions dominate superficial hydrogen bonds.

### 3.7. Calorimetric characterization of formed supramolecular gels

On the Table 2, the phase transition temperature  $T_{\text{gel}}$  was determined in toluene for gelators **1**, **2**, **3**, **5**, and **6**, and in ethylacetate for gelator **4**. It was measured by differential scanning calorimetry (DSC) in a range of temperatures between  $15$  and  $80^\circ\text{C}$ . The gel-to-sol melting temperature ( $T_{\text{g-s}}$ ) was measured for toluene organogels at  $3\text{ wt } \%$ . Gelators **1**, **3**, **4**, and **5** gave values of  $46$ ,  $43$ ,  $68$  and  $80^\circ\text{C}$ , respectively. For gelators **4** and **5**, due to the evidently greater packing strength, the transition requires more energy. Against all expectation, the  $T_{\text{g-s}}$  value of organogel **3** is not much higher than that of gelator **1**, despite its additional inserted amide group and its possibly higher number of intermolecular hydrogen bonds. That said, with a value of  $66^\circ\text{C}$ , gelator **2** has a smaller melting temperature than that of gelator **6**, whose melting temperature of  $77^\circ\text{C}$  can be seen as a sign of higher crystallinity of the material inside the gel matrix. Herein, despite our efforts, for this series of gelators, we were not able to measure a coherent value of sol-to-gel crystallization ( $T_{\text{s-g}}$ ) values. On the other hand, the gelation space volume does not increase linearly with the measured melting temperature values.

Table 2. Gel melting temperature from DSC measurements (from toluene).

| Compound                                 | <b>1</b> | <b>2</b> | <b>3</b> | <b>4</b> | <b>5</b> | <b>6</b> |
|--|----------|----------|----------|----------|----------|----------|
| $T_{\text{g-s}}\text{ (}^\circ\text{C)}$ | 46       | 66       | 43       | 68       | 80       | 77       |

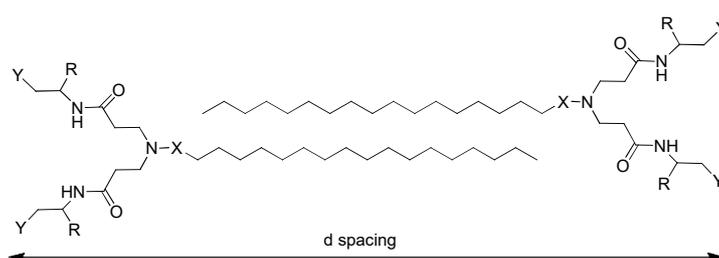
### 3.8. Molecular packing and gel network organization

X-Ray diffraction (XRD) analysis of the xerogels was performed to study their network structures. For all gelators, the reflection peaks observed at small angle regions are reported in Table 3, together with the associated d-spacing values. The 3D hypothetical models of organogels **1-6** were subjected to a simple empirical method for calculation of the molecular electronic structure (AM1) found in the Argus lab a freeware available on the web, in order to minimize these conformational energies.

**Table 3.** First XRD reflections peaks of air-dried toluene organogels 1-3, 5-6 and air-dried AcOEt organogel 4 ; predicated length of gelators 1-6.

| Gelators | d-spacing (nm) | first reflection peaks (nm) | Calculated length (nm) |
|----------|----------------|-----------------------------|------------------------|
| 1        | 3.76           | 2.38                        | 2.97                   |
| 2        | 3.47           | 2.66                        | 3.08                   |
| 3        | 5.06           | 3.47                        | 3.17                   |
| 4        | 4.08           | 3.94                        | 3.55                   |
| 5        | 3.0            | 3.54                        | 3.21                   |
| 6        | 3.4            | 2.80                        | 3.07                   |

For all gelators, in the wide angle region, the xerogels displayed also sharp diffraction peaks representative of highly ordered lateral chain packing. As reported elsewhere [5, 8], for lipopeptide-type gelators, the d-spacing ratios are practically 1 : 1/2 : 1/3. Indeed, gelators **1** and **3** have ratios of reciprocal of integers, which is characteristic of a lamellar structure. In case of gelator **4**, the ratio pattern evolved by the same manner. Interestingly, for gelator **5** the first d-spacing value of 3.0 nm was not compatible with the typical lamellar structure generally observed with lipopeptide-type organogelators as depicted on Fig. 10. Even the normal d-spacing ratio of a parallel interdigitated network was finally observed, the gelators **5** is suspected to show a real tendency to self-assemble into a parallel  $\beta$ -sheet system reinforced by a stronger intermolecular hydrogen networks between the sheets, due to the presence of urea terminal groups. Thus, the d-spacing is really lower compared to those expressed by other gelators. Hence, for all gelators except gelator **4** and **5**, regarding the computational simulation, the bilayer thickness appears to be smaller than twice the molecular length at full extension. This situation is representative of partial interpenetration of the fatty alkyl chain inside the gel networks.



**Fig. 10.** Molecular models of bilayers structures obtained from 1-3, 6.

Comparing gelator **2** and **6** is also interesting. Just like gelators **1** and **2** that are able to self-assemble in toluene into parallel lamellar structures with approximately the same d-spacing of 3.40 nm, gelator **6** displays almost the same d-spacing values but ultimately shows an apparent distortion of its network that can be seen in its d-spacing ratios from 1: 1/2 given by **1**, **2** and **3**, which is always characteristic of a cubic system. This situation can be explained by the presence of bulky lateral methyl groups on both sides of the molecule. Indeed, they are the main source of enhanced **steric** hindrance around the hydroxyl groups

of the molecule. Also, compared to other compounds, gelator **3** shows in its xerogel networks a degree of penetration of the alkyl chain that is lower than usual. Without a doubt, the additional inserted amide group acts as a wall preventing complete entry of alkyl chains inside the gel networks as depicted on Figure 10.

#### 4. Conclusion

The machine learning methods that were tested performed similarly well as predictive tools, but one stood out in particular (SVM) to provide useful visualizations to help interpreting the results. Gelators **1**, **2** and **6** thus showed two clear and distinct gelation regions. Independently of the gelation of water, glycerol led to a gel in many cases. Polyols of higher viscosity such as mixtures of di- and tri glycerol could be studied in the same way. Indeed, a remarkable disequilibrium was observed for gelator **4**, with hydrophobic bonds dominating the intermolecular hydrogen bonds in alcohols and aromatic solvents. Finally, the case of gelator **5** provided a good demonstration of the importance of a well-thought molecular design of gelator candidates. Even though all required elements were present in its structure, the sequence was indeed not the most effective to reduce its solubility. However, remained as a challenge for us, how to find a method for interpreting through a machine learning prediction the impact of carbon chirality, especially for the peptide-based supramolecular gelators, but also the characteristic values of logP or the rheology parameters contribution. In summary, this present work introduced a good alternative, but we are currently trying to improve our models.

#### Acknowledgment

The authors did not receive any financial support for this present work. All programs employed in this work remain available from the authors on demand.

#### APPENDIX A. Supplementary data

Supplementary datas (<sup>1</sup>H NMR charts of gelators 3-6 and some photographs of gel samples) associated with this article can be found in the online version, at...

#### References

1. F. Delbecq, Supramolecular gels from lipopeptide gelators: Template improvement and strategies for the in-situ preparation of inorganic nanomaterials and for the dispersion of carbon nanomaterials. *Adv. Colloid and Interface Sci.* 209 (2014) 98-108.
2. F. Delbecq, Y. Masuda, Y. Ogue, T. Kawai, Salt complexes of two-component N-acylamino acid diastereoisomers: self-assembly studies and modulation of gelation abilities. *Tetrahedron Lett.* 53 (2012) 6588-6593.
3. F. Delbecq, K. Tsujimoto, Y. Ogue, T. Kawai, Synthesis of green chiral organogelators with a sulfide linkage via Michael addition: soft templates for the preparation of size-controlled gold nanoparticles. *Tetrahedron Lett.* 54 (2013) 651-656.
4. C. K. Rouse, A. D. Martin, C. J. Easton, P. Thordarson, A Peptide Amphiphile Organogelator of Polar Organic Solvents. *Nature Scientific Reports* 7 (2017) 43668.

5. C. Morita, C. Kawai, A. Kikuchi, Y. Imura, T. Kawai, Effect of amide moieties for hydrogelators on gelation property and heating-free pH responsive gel-sol phase transition. *J. Oleo. Sci.* 61 (2012) 707-713.
6. C. Morita, H. Sugimoto, Y. Imura, T. Kawai, Double-stimuli Responsive O/W Emulsion Gel Based on a Novel Amidoamine Surfactant. *J. Oleo Sci.* 60 (2011) 557-562.
7. C. Morita, H. Sugimoto, K. Matsue, T. Kondo, Y. Imura, T. Kawai, Changes in viscosity behavior from a normal organogelator to a heat-induced gelator for a long-chain amidoamine derivative. *Chem. Commun.* 46 (2010) 7969-7971.
8. C. Morita, C. Kawai, K. Tsujimoto, K. Kasai, Y. Ogue, Y. Imura, T. Kawai, High organogelation ability and soft-templating for ultrathin Au nanowires of long-chain amidoamine derivatives. *J. Oleo Sci.* 62 (2013) 81-87.
9. C. Morita, H. Tanuma, C. Kawai, Y. Ito, Y. Imura, T. Kawai, Room-Temperature Synthesis of Two-Dimensional Ultrathin Gold Nanowire Parallel Array with Tunable Spacing. *Langmuir* 29 (2013) 1669-1675.
10. F. Delbecq, P. Delfosse, G. Laboureix, C. Paré, T. Kawai, Study of a gelated Deep Eutectic solvent metal salt solution as template for the production of size-controlled small noble metal nanoparticles. *Colloid and Surface A* 567 (2019) 55-62.
11. F. Delbecq, I. Yamano, T. Kawai, A pH controlled Reversible Phase Transfer and Electrolytic Size-Fractionation of Stable Silver Nanoparticles Capped with a Long Amino Amide Dicarboxylic Acid. *Bull. Chem. Soc. Jpn* 88 (2015) 1719-1725.
12. J. Bonnet, G. Suissa, M. Raynal, L. Bouteiller, Organogel formation rationalized by Hansen solubility parameters: influence of gelator structure. *Soft Matter* (2015) 2308-2313.
13. Y. Lan, M. G. Corradini, X. Liu, T. E. May, F. Borondics, R. G. Weiss, M. A. Rogers, Comparing and Correlating Solubility Parameters Governing the self-Assembly of Molecular Gels Using 1,3:2,4-Dibenzylidene Sorbitol as the Gelator. *Langmuir* 30 (2014) 14128-14142.
14. E. R. Draper, D. J. Adams, Low-Molecular-Weight Gels: The State of Art. *Chem* 3 (2017) 390-410.
15. M. Hansen, Hansen solubility parameters: a user's handbook, 2<sup>nd</sup> ed. Boca Raton: CRC Press, 2007.
16. K. K. Diehn, H. Oh, R. Hashemipour, R. G. Weiss, S. R. Raghavan, Insights into organogelation and its kinetics from Hansen solubility parameters. Toward a priori predictions of molecular gelation. *Soft Matter* 10 (2014) 2632-2640.
17. M. G. Corradini, M. A. Rogers, Molecular gels: improving selection and design through computational methods. *Curr. Opin. Food Sci.* 9 (2016) 84-92.
18. N. Yan, Z. Xu, K. K. Diehn, S. R. Raghavan, Y. Fang, R. G. Weiss, How Do Liquid Mixtures Solubilize Insoluble Gelators? Self-Assembly of Pyrenyl-Linker-Glucono Gelators in Tetrahydrofuran-Water Mixtures. *J. Am. Chem. Soc.* 135 (2013) 8989-8999.
19. J. K. Gupta, D. J. Adams, N. G. Berry, Will it Gel? Successful computational prediction of peptide gelators using physicochemical properties and molecular fingerprints. *Chem. Sci.* 7 (2016) 4713-4719.
20. J. Bonnet, G. Suissa, M. Raynal, L. Bouteiller, Organogel formation rationalized by Hansen solubility parameters: dos and don'ts. *Soft Matter* 10 (2014) 3154-3160.

21. Y. Lan, M. G. Corradini, R. G. Weiss, S. R. Raghavan, M. A. Rogers, To gel or not to gel: correlating molecular gelation with solvent parameters. *Chem Soc Rev* 44 (2015) 6035-6058.
22. M. A. Rogers, Hansen Solubility Parameters as a Tool in the Quest for New Edible Oleogels. *J Am Oil Chem Soc* 2018, <https://doi.org/10.1002/aocs.12050>.
23. B. S. Lengeling, L. M. Roch, J. D. Perea, S. Langner, C. J. Brabec, A. Aspuru-Guzik, A Bayesian approach to predict solubility parameters. *Theory and Simulations* 2 (2019), 1800069.
24. S. Boobier, A. Osbourn, J. B. O. Mitchell, Can human experts predict solubility better than computers? *J Cheminform* 9 (2017) 63-76.
25. Y. Lan, M. G. Corradini, M. A. Rogers, Do Molecular Gelator Cluster in Hansen Space? *Cryst. Growth Des* 14 (2014) 4811-4818.
26. C. Liu, M. Corradini, M. A. Rogers, Self-assembly of 12-hydroxystearic acid molecular gels in mixed solvent systems rationalized using Hansen solubility parameters. *Colloid Polym Sci* 293 (2015) 975-983.
27. B. Hossin, K. Rizi, S. Murdan, Application of Hansen Solubility Parameters to predict drug-nail interactions, which can assist the design of nail medicines. *Eur. J. Pharm. Biopharm* 102 (2016) 32-40.
28. F. Li, J. Han, T. Cao, W. Lam, B. Fan, W. Tang, S. Chen, K. Lamfok, L. Li, Design of self-assembly dipeptide hydrogel and machine learning via their chemical features. *PNAS* 116 (2019) 11259-11264.
29. B. M. Jamcel, A. Huynh, A. Chadha, S. Pamdey, J. Duncam, M. Chandler, G. Baki, *Int. J. Pharm* 569 (2019) 118549.
30. A. Yan, J. Gasteiger, Prediction of Aqueous Solubility of Organic Compounds by Topological Descriptors. *QSAR Comb. Sci.* 22 (2013) 821-829.
31. J. P. Janet, H. J. Kulik, Resolving Transition Metal Chemical Space: Feature Selection for Machine Learning and Structure-Property Relationships. *J Phys Chem A* 121 (2017) 8939-8954.
32. M. M. Batista, R. Guirardello, M. A. Krahenbuhl, Determination of the Hansen solubility Parameters of Vegetable Oils, Biodiesel, Diesel, and Biodiesel-Diesel Blends. *J Am Oil Chem Soc* 92 (2015) 95-109.
33. M. A. Rogers, A. G. Marangoni, Kinetics of 12-Hydroxyoctadecanoic Acid SAFiN Crystallization Rationalized Using Hansen Solubility Parameters. *Langmuir* 32 (2016) 12833-12841.

# Gelation Properties of Various Long Chain Amido Amines: Prediction of Solvent Gelation via Machine Learning using Hansen Solubility Parameters

Frederic Delbecq<sup>a\*</sup>, Guillaume Adenier<sup>b</sup>, Yuki Ogue<sup>c</sup>, Takeshi Kawai<sup>c</sup>

