Electronic Supplementary Material (ESI) for Sustainable Energy & Fuels. This journal is © The Royal Society of Chemistry 2018

Supporting Information for Computational Design of Biofuels from Terpenes and Terpenoids

Efrat Pahima , Shmaryahu Hoz, Moshe Ben-Tzion, Dan Thomas Major*

Department of Chemistry, Bar Ilan University, Ramat-Gan 52900, Israel

* Electronic mail: majort@biu.ac.il

Theoretical Methods:

All molecules were built using GaussView 5.0,¹ and all calculations were carried out using the Gaussian 09 software (version B.01).²

The enthalpies of combustion were calculated using the following density functional theory (DFT) methods and basis sets:

Functionals	Basis sets	
DFTBA ³		
M06-2X ⁴	6-31+G(d,p)	6-311+G(2d,p)
B3LYP-D3 ^{5,6}	6-31+G(d,p)	6-311+G(2d,p)
B2PLYP-D3 ^{7,8}		6-311+G(2d,p)

We note that the DFTBA functional (the spin-polarized self-consistent-charge density-functional tight-binding method) is semi-empirical, and hence is several orders of magnitude faster than the two other functionals, but typically less accurate. M06-2X and B3LYP-D3 are hybrid functionals that include dispersion corrections (M06-2X by construction and B3LYP via the D3 dispersion correction). B2PLYP-D3 is a dispersion corrected double hybrid functional, which is considerably slower than standard hybrid functionals, due to inclusion of second-order Møller-Plesset (MP2) perturbation theory.

The enthalpies of formation were calculated using Gaussian-4 with second-order MP2 perturbation theory (G4MP2) in order to get accurate results (explained below). This method is considerably more accurate than the DFT methods mentioned above, but costlier. Initial calculations of enthalpies of formation using DFT indicated that these methods are not accurate enough.

The enthalpies of vaporization were calculated using the SMD solvation model,⁹ in conjunction with the M06-2X/6-31+G(d,p) method.

An important approximation in the thermodynamic analysis of vibrations, rotations and translations is that all the equations assume non-interacting molecules and are therefore applicable only to an ideal gas. This limitation will introduce some error, depending on the system being studied.¹⁰ Nonetheless, we stress that the treatment for each individual molecule is rigorous and gives the electronic energy and zero-point energy and thermal corrections, *H*^{corr}.

All the thermodynamic properties calculated by Gaussian in this work are directly derived from the electronic properties of the molecule and standard equations of statistical mechanics.

Thermodynamic considerations:

For any molecule, the *total energy* is the energy required to convert the molecule into isolated atoms (breaking all bonds) and then remove all electrons from those atoms.

Enthalpy is defined as H = U + pV, the summation of the *internal thermal energy* and the pressure times the volume of the system.¹¹ The enthalpy change is the energy transferred as heat at constant pressure, $\Delta H = Q_p$. The standard Enthalpy of Combustion is the standard reaction enthalpy for the complete oxidation of an organic compound to CO₂ and H₂O if the compound contains C, H, and O. (The standard enthalpy change is the change in enthalpy for a process in which the initial and final substances are in their standard states, i.e. the pure substance at 1 atm). Hence, by definition the enthalpy (heat) of Combustion (i.e. ΔH_{comb} , also known as HHV-higher heating value) is the heat produced from burning 1 mol of fuel, and equals to the enthalpy change for the combustion reaction, i.e. $\Delta H_{reaction}$.

Hess's law states that the standard enthalpy of an overall reaction is the sum of the standard enthalpies of the individual reactions into which a reaction may be divided. The standard reaction enthalpy may be estimated by combining enthalpies of formation: $\Delta H_r = \Sigma_{Products} v \Delta H_f - \Sigma_{Reactants} v \Delta H_f = \Delta H_{comb}$ (v=stoichiometric coefficient).

The Enthalpy of Formation, ΔH_f , of a molecule is the difference in energy between a molecule and the atoms that make up that molecule in their naturally occurring elemental form, meaning the enthalpy of formation of a hydrocarbon is defined as the formation of the molecule from its constituent elements in their reference standard states (graphite for C; H₂ at 1 atm for H).

The Enthalpy of Vaporization, ΔH_{vap} , is defined as the amount of energy in the form of heat that must be added to the liquid substance, to transform a quantity of that substance into a gas.

The **Gibbs Free Energy** is defined by G = H - TS. The standard Gibbs free energy of reaction is given by: $\Delta G_r = \Delta H_r - T\Delta S_r = \Sigma_{Products} \vee G_m - \Sigma_{Reactants} \vee G_m$. (G_m stands for Gibbs free energy per mol). The standard Gibbs free energy of formation (ΔG_i) is the standard reaction Gibbs free energy for the formation of a compound from its elements in their reference states. The standard Gibbs free energy of reaction may be expressed in terms of ΔG_f : $\Delta G_r = \Sigma_{Products} \vee \Delta G_f - \Sigma_{Reactants} \vee \Delta G_f$.

Heat capacity of a substance is the amount of heat required to raise the temperature of a substance by 1°C at constant volume. It is defined as the change in the

$$Cv = \left(\frac{\partial U}{\partial T}\right)_v$$

energy of a substance as its temperature is changed under constant volume: dT^{*} . Translation, rotation, and vibration of atoms represent the degrees of freedom of motion, which classically contribute to the heat capacity of gases.

 \mathcal{E}_0 is the internal energy (i.e. electronic energy) of the molecule at 0 K within the Born-Oppenheimer approximation, and E_{zpe} the zero-point energy of the molecule at 0 K (a correction to the electronic energy). E_{trans} , E_{rot} and E_{vib} are the thermal energy corrections due to the effects of molecular translation, rotation and vibration at the specified temperature, respectively.

The internal total thermal energy $E_{tot} = U$ obtained from the partition function is calculated by: $E_{tot} = Nk_BT^2(\frac{\partial lnq}{\partial T})_V$ where $q = q_v q_r q_t q_e$, while each motion has its own contribution to the total energy: $E_{tot} = E_t + E_r + E_v + E_e$

$$Cv = \left(\frac{\partial L}{\partial T}\right)_{N,V}$$
 and the entropy can be computed by:

$$S_{tot} = R\left(\ln q + T\left(\frac{\partial \Pi q}{\partial T}\right)_V\right)$$
. The enthalpy and Gibbs free energy are then calculated by: $H = E_{tot} + k_B T$ and $G = H - T S_{tot}$

aln a

After the addition of the total electronic energy, ε_0 , to these thermal energies, the final thermodynamic property at a given temperature may be computed.

ar

Computation of the enthalpy of combustion:

All Gaussian calculations with a given DFT functional for each terpene molecule were performed using two basis sets: 6-31+G(d,p), and 6-311+G(2d,p), except for the double hybrid B2PLYP as mentioned above.

The input file included the commands "Opt" - that instructs a geometry optimization be performed, i.e. the geometry is adjusted until a stationary point on the potential energy surface is found, and the "Freq" keyword instructs that force constants are computed, as well as the resulting vibrational frequencies. No imaginary frequencies were identified in all calculations (except for transition state calculations), indicating that the molecules are at a local minimum.

In the B3LYP functional calculations we also added "EmpiricalDispersion=GD3", which adds the D3 version of Grimme's dispersion correction to the calculations. This accounts for the lack of dispersion interactions in the B3LYP and B2PLYP functionals.

We also employed a DFT double hybrid functional, which combines exact HF exchange with MP2 correlation to a DFT calculation.

We also note that Boltzmann-weighted averages of the energies were considered throughout this work for flexible molecules in their stable conformations. For the pinene-dimers, in an attempt to make sure the current geometry of the flexible molecules is at the global minimum and not a high-energy local minimum, we performed simulated annealing calculations using Materials Studio's Forcite module with the COMPASS force field.^{12,13} The conformers found were then optimized using DFT in Gaussian.

The ΔH_{comb} was calculated according to the following Gaussian white page¹ scheme:

(51)
$$\Delta H_{comb} = n \cdot (\varepsilon_0 + H^{corr})_{CO_2(g)}^{298K} + \frac{m}{2} \cdot (\varepsilon_0 + H^{corr})_{H_2O(g)}^{298K} - [(\varepsilon_0 + H^{corr})_{C_nH_m(g)}^{298K} + \left(n + \frac{m}{4}\right) \cdot (\varepsilon_0 + H^{corr})_{O_2(g)}^{298K}]$$

The equation above calculates the enthalpy of combustion at 298K when all reactants and products are in the **gas phase**. By definition, reaction enthalpies are measured experimentally when all reactants and the products are at 298K and 1 atm. During the oxidation reaction at room temperature, the terpene molecule's initial state is the **liquid phase**. The terpenes are oxidized and vaporized at the same time, creating the products CO₂ and water in the gas phase. The water then condenses upon reaching 298K at 1 atm.

 ΔH_{comb} values were calculated using the following equation, *after* the first correction to the ΔH_{vap} :

$$(52) \qquad \Delta H_{comb} = n \cdot (\varepsilon_0 + H^{corr})_{\mathcal{C}0_2(g)}^{298K} + \frac{m}{2} \cdot [(\varepsilon_0 + H^{corr})_{H_2^{0}(g)}^{298K} - \Delta H_{vap, H_2^{0}}] - \left[(\varepsilon_0 + H^{corr})_{\mathcal{C}_n^{H_m}(g)}^{298K} - \Delta H_{vap, \mathcal{C}_n^{H_m}} + \left(n + \frac{m}{4}\right) \cdot (\varepsilon_0 + H^{corr})_{\mathcal{O}_2(g)}^{298K}\right]$$

Further explanations regarding the calculation of ΔH_{vap} appear in the section below.

Correction to the ΔH_{comb} due to the O₂ overbinding effect in DFT calculations:

The second correction was added in order to account for a deviation possible caused by electron self interaction¹⁴. DFT's tendency to overbind the electrons in the O_2 molecule^{15–17} results in O_2 molecule's enthalpy being too negative (using so-called GGA/LDA functionals). The binding energies in sp-bonded molecules are usually overestimated by both LDA, GGA, and hybrid functionals (such as M06-2X); this overestimation is greater for bonds involving atoms with more than half-filled shells. The O_2 molecule's computed enthalpy is present in all the combustion energy calculations, thus causes a constant deviation from the experimental values. This problem can be overcome by using linear regression that correlates between the **computed** ΔH_{comb} (without the O_2 correction) and the **experimental** ΔH_{comb} . There is a linear correlation between these two values because of the constant deviation caused by the O_2 overbinding. These figures are presented in Fig. S1a-f.

Initial calculations of the enthalpy of combustion, using each functional and basis set, were carried out for the **monoterpenes** only, while applying the first correction of the *enthalpy of vaporization* by using the **experimental values of the** ΔH_{vap} . Then, the second correction of the O₂ overbinding effect was added, and each method's overall deviations from the experimental ΔH_{comb} values were calculated, in order to determine which method gives the most accurate results. The most accurate method is, as shown in Table S7, the *M06-2X/6-31+G(d,p)*. The remaining calculations were carried out using this method.

(Note that all calculations were performed at room temperature- 298K)

M06-2X/6-31+G	(d,p)	Electronic Energy [Hartree/particle]	Electronic and Thermal	ΔH _{comb}	Experimental ΔH _{comb}	ΔH_{comb} after O ₂ overbinding	Deviation [kcal/mol]
	Molecules		Enthalpy [kcal/mol]	correction, before O ₂ correction [kcal/mol]	[kcal/mol]	correction [kcal/mol]	
C ₁₀ H ₁₆	α-Pinene	-390.49681	-244884.82	-1404.15	-1483.00 ¹⁸	-1481.86	-1.1
	β-Pinene	-390.49385	-244882.76	-1406.11	-1485.10 ¹⁸	-1483.81	-1.3
	Camphene	-390.51584	-244896.13	-1393.02	-1468.92 ¹⁹	-1470.84	1.9
	Car-3-ene	-390.50431	-244889.45	-1398.80	-1479.97 ²⁰	-1476.57	-3.4
	Limonene	-390.50683	-244890.97	-1397.07	-1473.95 ¹⁸	-1474.85	0.9
	Phellandrene	-390.50784	-244891.46	-1396.81		-1474.59	
	Sabinene	-390.49646	-244884.67	-1403.94	-1477.50 ¹⁹	-1481.66	4.2
	γ-Terpinene	-390.51135	-244894.27	-1393.27	-1472.04 ²¹	-1471.08	-1.0
	Terpinolene	-390.50772	-244891.56	-1396.19		-1473.98	
C ₁₀ H ₁₈ O	Fenchol	-466.94185	-244884.82	-1385.15		-1463.04	
	Terpinol	-466.93136	-244882.76	-1390.74	-1468.78 ²¹	-1468.58	-0.2
	O ₂	-150.26395	-94287.44				
	CO ₂	-188.51605	-118285.89				
	H ₂ O	-76.39496	-47922.62				

Table S1: M06-2X computed data and O_2 correction, using the 6-31+G(d,p) basis set, and using the experimental ΔH_{vap} values.

M06-2X/	6-311+G(2d,p)	Electronic Energy [Hartree/particle]	Electronic and Thermal	ΔH _{comb} after ΔH _{vap}	Experimental ΔH _{comb}	ΔH_{comb} after O ₂	Deviation [kcal/mol]
	Molecules		Enthalpy [kcal/mol]	correction, before O ₂ correction [kcal/mol]	[kcal/mol]	overbinding correction [kcal/mol]	
$C_{10}H_{16}$	α-Pinene	-390.58275	-244938.88	-1441.29	-1483.00 ¹⁸	-1483.59	0.6
	β-Pinene	-390.58547	-244940.88	-1439.20	-1485.10 ¹⁸	-1481.73	-3.4
	Camphene	-390.60490	-244952.32	-1428.03	-1468.92 ¹⁹	-1471.78	2.9
	Car-3-ene	-390.59320	-244945.49	-1433.96	-1479.97 ²⁰	-1477.06	-2.9
	Limonene	-390.59791	-244948.40	-1430.85	-1473.95 ¹⁸	-1474.29	0.3
	Phellandrene	-390.59851	-244948.70	-1430.78		-1474.22	
	Sabinene	-390.58568	-244940.92	-1438.89	-1477.50 ¹⁹	-1481.46	4.0
	γ-Terpinene	-390.60255	-244951.39	-1427.36	-1472.04 ²¹	-1471.17	-0.9
	Terpinolene	-390.59856	-244948.87	-1430.08		-1473.60	
C ₁₀ H ₁₈ O	Fenchol	-467.05503	-292906.89	-1422.03		-1466.42	
	Terpinol	-467.04977	-292904.25	-1423.99	-1468.78 ²¹	-1468.17	-0.6
	O ₂	-150.25411	-94317.35				
	CO ₂	-188.57889	-118325.31			1	
	H ₂ O	-76.42204	-47939.60				

Table S2: M06-2X computed data and O₂ correction, using the 6-311+G(2d,p) basis set and using the experimental ΔH_{vap} values.

DFTBA		Electronic Energy [Hartree/particle]	Electronic and Thermal	ΔH_{comb} after ΔH_{vap} correction, before O_2	Experimental ΔH _{comb} [kcal/mol]	ΔH_{comb} after O_2 overbinding	Deviation [kcal/mol]
	Molecules		Enthalpy [kcal/mol]	correction [kcal/mol]		correction [kcal/mol]	
$C_{10}H_{16}$	α-Pinene	-23.16783	-14385.77	-1469.49	-1483.00 ¹⁸	-1478.72	-4.3
	β-Pinene	-23.17194	-14388.61	-1466.55	-1485.10 ¹⁸	-1477.23	-7.9
	Camphene	-23.19206	-14400.35	-1455.08	-1468.92 ¹⁹	-1471.40	2.5
	Car-3-ene	-23.16098	-14381.79	-1472.75	-1479.97 ²⁰	-1480.38	0.4
	Limonene	-23.18092	-14394.44	-1459.90	-1473.95 ¹⁸	-1473.85	-0.1
	Phellandrene	-23.18020	-14393.79	-1460.77		-1474.29	
	Sabinene	-23.15211	-14376.42	-1478.48	-1477.50 ¹⁹	-1483.30	5.8
	γ-Terpinene	-23.18677	-14398.14	-1455.68	-1472.04 ²¹	-1471.71	-0.3
	Terpinolene	-23.18739	-14398.58	-1455.46		-1471.59	
C ₁₀ H ₁₈ O	Fenchol	-27.29109	-16954.84	-1452.91		-1470.30	
	Terpinol	-27.28089	-16949.51	-1457.56	-1468.78 ²¹	-1472.66	3.9
	O ₂	-6.50496	-4078.38				
	CO ₂	-8.38423	-5251.99			-	
	H ₂ O	-4.07779	-2543.35				

Table S3: DFTBA computed data and O_2 correction, using the experimental ΔH_{vap} values.

B3LYP-D3/	/6-31+G(d,p)	Electronic Energy	Electronic and Thermal	ΔH_{comb} after ΔH_{van} correction.	Experimental ΔH _{comb} [kcal/mol]	∆H _{comb} after O₂	Deviation [kcal/mol]
	Molecules		Enthalpy [kcal/mol]	before O ₂ correction [kcal/mol]		overbinding correction [kcal/mol]	
$C_{10}H_{16}$	α-Pinene	-390.71315	-245021.24	-1373.90	-1483.00 ¹⁸	-1483.28	0.3
	β-Pinene	-390.71612	-245023.47	-1371.59	-1485.10 ¹⁸	-1481.49	-3.6
	Camphene	-390.73512	-245034.65	-1360.68	-1468.92 ¹⁹	-1473.07	4.1
	Car-3-ene	-390.72286	-245027.68	-1366.75	-1479.97 ²⁰	-1477.76	-2.2
	Limonene	-390.73255	-245033.69	-1360.54	-1473.95 ¹⁸	-1472.96	-1.0
	Phellandrene	-390.73508	-245034.84	-1359.61		-1472.24	
	Sabinene	-390.71580	-245023.21	-1371.58	-1477.50 ¹⁹	-1481.49	4.0
	γ-Terpinene	-390.73723	-245036.77	-1356.95	-1472.04 ²¹	-1470.19	-1.8
	Terpinolene	-390.73409	-245034.83	-1359.10		-1471.85	
C ₁₀ H ₁₈ O	Fenchol	-467.19478	-292995.37	-1356.29		-1469.68	
	Terpinol	-467.19398	-292995.53	-1355.45	-1468.78 ²¹	-1469.03	0.2
	0 ₂	-150.26608	-94321.32				
	CO ₂	-188.59058	-118332.87				
	H ₂ O	-76.43406	-47947.37				

Table S4: B3LYP-D3 computed data and O_2 correction, using the 6-31+G(d,p) basis set and using the experimental ΔH_{vap} values.

B3LYP-D3/	/6-311+(2d,p)	Electronic Energy	Electronic and	ΔH_{comb} after AH_{un} correction	Experimental	ΔH_{comb} after	Deviation
	Molecules		Enthalpy [kcal/mol]	before O ₂ correction [kcal/mol]		overbinding correction [kcal/mol]	[keal/mol]
$C_{10}H_{16}$	α-Pinene	-390.79131	-245070.68	-1416.03	-1483.00 ¹⁸	-1483.17	0.2
	β-Pinene	-390.79420	-245072.87	-1413.74	-1485.10 ¹⁸	-1481.52	-3.6
	Camphene	-390.81354	-245084.26	-1402.64	-1468.92 ¹⁹	-1473.53	4.6
	Car-3-ene	-390.80187	-245077.64	-1408.35	-1479.97 ²⁰	-1477.65	-2.3
	Limonene	-390.81353	-245084.87	-1400.93	-1473.95 ¹⁸	-1472.30	-1.6
	Phellandrene	-390.81543	-245085.66	-1400.35		-1471.89	
	Sabinene	-390.79434	-245072.92	-1413.43	-1477.50 ¹⁹	-1481.30	3.8
	γ-Terpinene	-390.81851	-245088.14	-1397.14	-1472.04 ²¹	-1469.58	-2.5
	Terpinolene	-390.81510	-245086.05	-1399.45		-1471.24	
C ₁₀ H ₁₈ O	Fenchol	-467.29555	-293059.03	-1400.18		-1471.77	
	Terpinol	-467.29690	-293060.53	-1398.00	-1468.78 ²¹	-1470.20	1.4
	O ₂	-150.31319	-94353.86				
	CO ₂	-188.65082	-118370.63			_	
	H ₂ O	-76.45953	-47963.35				

Table S5: B3LYP-D3 computed data and O_2 correction, using the 6-311+G(2d,p) basis set and using the experimental ΔH_{vap} values.

B2PLYP-D3	3/6-311+G(2d,p)	Electronic Energy	Electronic	ΔH_{comb} after	Experimental	ΔH _{comb} after	Deviation
	Molecules	[naitiee/particle]	Thermal Enthalpy [kcal/mol]	before O ₂ correction [kcal/mol]		overbinding correction [kcal/mol]	[KCal/IIIOI]
C ₁₀ H ₁₆	α-Pinene	-390.36982	-244960.76	-1516.50	-1483.00 ¹⁸	-1480.61	-2.4
	β-Pinene	-390.36631	-244958.56	-1518.90	-1485.10 ¹⁸	-1482.56	-2.5
	Camphene	-390.38977	-244973.28	-1504.83	-1468.92 ¹⁹	-1471.11	2.2
	Car-3-ene	-390.37366	-244963.17	-1513.44	-1479.97 ²⁰	-1478.12	-1.8
	Limonene	-390.38453	-244969.99	-1506.54	-1473.95 ¹⁸	-1472.51	-1.4
	Phellandrene	-390.38606	-244970.95	-1506.00		-1472.07	
	Sabinene	-390.36574	-244958.20	-1518.68	-1477.50 ¹⁹	-1482.39	4.9
	γ-Terpinene	-390.38988	-244973.35	-1502.51	-1472.04 ²¹	-1469.23	-2.8
	Terpinolene	-390.38648	-244971.21	-1504.90		-1471.17	
C ₁₀ H ₁₈ O	Fenchol	-466.81070	-292928.13	-1502.42		-1469.16	
	Terpinol	-466.80350	-292923.62	-1506.81	-1468.78 ²¹	-1472.72	3.9
	O ₂	-150.26394	-94292.04				
	CO ₂	-188.51788	-118296.76			-	
	H ₂ O	-76.39854	-47940.81				

Table S6: B2PLYP-D3 computed data and O₂ correction, using the 6-311+G(2d,p) basis set and using the experimental ΔH_{vap} values.



Figures S1,a-f: Regression analysis for calculating the ΔH_{comb} for each method and basis set (R² values: a: 0.86, b: 0.85, c: 0.42, d: 0.83, e: 0.79, f: 0.75)

By calculating the absolute average deviation, standard deviation and RMS as shown in Table S7, the most accurate computational method for calculating the enthalpy of combustion turned out to be M06-2X, where both basis sets gave similarly good accuracy. In this study, we decided to use the smaller 6-31+G(d,p) basis set, considering that its computational cost is significantly lower.

	Experimental ΔH _{comb}	Exp. error	M06-2X		B3LYP-D3		B2PLYP-D3	DFTBA
Molecules	[kcal/mol]		6-31+G(d,p)	6-311+G(2d,p)	6-31+G(d,p)	6-311+G(2d,p)	6-311+G(2d,p)	
α-Pinene	-1483.00 ¹⁸	± 0.7	-1.1	0.6	0.3	0.2	-2.4	-4.3
β-Pinene	-1485.10 ¹⁸	± 0.7	-1.0	-3.4	-3.6	-3.6	-2.5	-7.9
Camphene	-1468.92 ¹⁹	± 0.5	1.9	2.9	4.1	4.6	2.2	2.5
Car-3-ene	-1479.97 ²⁰	± 0.4	-3.5	-2.9	-2.2	-2.3	-1.8	0.4
Limonene	-1473.95 ¹⁸		1.1	0.3	-1.0	-1.6	-1.4	-0.1
Sabinene	-1477.50 ¹⁹		3.8	4.0	4.0	3.8	4.9	5.8
γ-Terpinene	-1472.04 ²¹		-0.8	-0.9	-1.8	-2.5	-2.8	-0.3
Terpinol	-1468.78 ²¹	± 1.5	-0.4	-0.6	0.2	1.4	3.9	3.9
Average absolute deviation	n		1.7	1.9	2.2	2.5	2.8	3.1
Standard deviation			1.3	1.5	1.6	1.4	1.1	2.8
RMS			2.1	2.4	2.6	2.8	3.0	4.1
Maximum absolute deviati	ion		3.8	4.0	4.1	4.6	4.9	7.9

Table S7: Deviations between computed and experimental enthalpies of combustion (kcal/mol).

Note: all the above calculations were performed using the experimental ΔH_{vap} .

Computation of the enthalpy of vaporization:

The enthalpies of vaporization were calculated using the SMD solvation model⁹, with the M06-2X/6-31+G(d,p) method.

Using this model, the Gibbs free energy of solvation, which is defined as the free energy required to insert a molecule from the gas phase into the liquid phase, is computed using:

$$\Delta G_{sol} \approx E^*_{Liquid} - \varepsilon_{0Gas} \tag{S3}$$

 ε_0 is the electronic energy of the molecule in the gas phase, and E* is the total energy calculated by the SMD model, consisting of the electrostatic and nonelectrostatic contributions to the energy of the molecule inside the continuum medium and is affected by the reaction field (the electric potential derived from the polarized medium and solute that mutually polarize each other), the solvent cavitation, dispersion interaction, solvent structure. More information about this model can be found in the original SMD paper.⁹

In order to compute the enthalpy of combustion (see above), we need to convert ΔG_{sol} to the enthalpy of vaporization ΔH_{vap} (the SMD and related models only compute the free energy of solvation). We found that the computed ΔG_{sol} is linearly correlated to the experimental ΔH_{vap} . To convert ΔG_{sol} into ΔH_{vap} without explicitly calculating ΔS_{sol} , we employed multiple linear regression:

According to Ref. ^{22,23} the entropic contribution to the solvation free energy is due to the rearrangement of the solvent, induced by the solute's presence. The *solvent* accessible surface area (SASA) can therefore be used as a measure of the entropic effect. This can be summarized as the following:

 $\Delta G_{solv} + T \cdot \Delta S_{solv} = \Delta H_{solv}, \ \Delta S_{solv} \propto SASA \Rightarrow \Delta G_{solv} + T \cdot SASA \propto \Delta H_{solv}$

By applying multiple linear regression (MLR), where ΔG_{sol} (computed as explained above) and T-SASA (computed as part of the SMD model calculation, where T=298.15 K) are the variables, the values of the ΔH_{vap} of the terpene molecules were predicted with less than 1 kcal/mol deviation from the experimentally known values.

The terpenes' solvation free energies were calculated using *n*-hexane as a solvent and for the terpenoids *1*-hexanol was used as a solvent, resulting in slightly greater deviations (due to the inability to express hydrogen bonds in the continuum model).

To validate the model, we performed cross validation for the set of molecules used in the MLR. The Q² correlation coefficients calculated from the cross validation (i.e. the correlation between the experimental ΔH_{vap} values and the predicted values for each molecule that was left out in each step of the cross validation) were 0.92 for the terpenes and 0.46 for the terpenoids. These correlation figures are presented under Figures S2,a-b. Despite the poor correlation for the terpenoids, the predicted values are within 1.5 kcal/mol of the experimental values. Tables S8,a-b present the terpene / terpenoid molecules' data that were used in the MLR. Figures 1,a-b (in the main paper) show the R² correlation between the experimental ΔH_{vap} and the predicted ΔH_{vap} that were calculated using the MLR *after* the cross validation test.

An important note is that despite the small coefficient for the x_2 variable ($T \cdot SASA$), the various statistic data shown in Table S9 indicated that the x_2 property is important for explaining the variance of y and for predicting the true ΔH_{vap} values in this model. As shown in Table S9, the F statistic value (calculated at 5% significance level) for the MLR data is higher (186.2) than the value calculated using linear regression with ΔG_{sol} being x and the experimental ΔH_{vap} being y (178.4). This indicates the high R² value given for the MLR is not coincidental. Indeed, it is expected that MLR performs better than LR due to the added variables.



Figures S2,a-b: Multiple linear regression analysis results for calculating the ΔH_{vap} using cross validation. The plotted values are of the predicted ΔH_{vap} from the cross-validation calculations against the experimental ΔH_{vap} . (Q² values: (a) Terpenes – 0.92, (b) Terpenoids – 0.46)

	Molecules	ΔG _{solv} Solvation Free Energy comp. [kcal/mol] -x ₁	Cavity Surface Area (=SASA) [Å ²]	SASA·T [Ų·K] x ₂	Experimental ΔH _{vap} [kcal/mol] Y	Predicted ∆H _{vap} [kcal/mol]	Deviation	
Terpene	s used in the multiple	e linear regres	sion:					
C ₁₀ H ₁₆	α-pinene	-5.24	184.36	54968.13	10.84 ²⁴	10.88	0.0	Monoterpenes
C ₁₀ H ₁₆	β-pinene	-5.09	184.04	54872.42	10.94 ²⁵	10.73	0.2	
C ₁₀ H ₁₆	Camphene	-4.96	184.14	54901.34	10.66 ±0.2 ^{26,27}	10.60	0.1	
C ₁₀ H ₁₆	Limonene	-5.59	192.48	57389.10	11.77 ²⁷	11.48	0.3	
$C_{10}H_{16}$	Sabinene	-5.69	192.12	57280.88	11.20 ²⁷	11.57	-0.4	
C ₁₀ H ₁₆	γ-Terpinene	-6.02	196.05	58452.01	12.27 ²⁷	12.02	0.3	
$C_{10}H_{16}$	Car-3-ene	-5.85	188.07	56073.37	11.56 ²⁰	11.61	0.0	
C ₁₀ H ₁₆	β-Phellandrene	-5.48	194.56	58008.36	11.54 ²⁶	11.43	0.1	
$C_{10}H_{16}$	Terpinolene	-6.13	194.43	57970.20	12.06 ²⁶	12.08	0.0	
C ₁₀ H ₁₆	Myrcene	-4.60	212.30	63296.65	12.09 ²⁶	11.06	1.0	
C ₁₀ H ₁₈	Pinane	-4.97	186.24	55527.46	9.98 ²⁶	10.67	-0.7	Saturated Terpenes
C ₁₅ H ₂₄	β-Farnesene	-8.51	306.06	91252.68	17.32 ²⁷	17.73	-0.4	Sesquiterpenes
$C_{15}H_{24}$	β-Caryophyllene	-7.07	264.42	78835.93	15.65 ²⁷	15.07	0.6	
Predicte	d values:							
$C_{10}H_{16}$	Alloocimene	-6.34	211.48	63052.17		12.78		Monoterpenes
$C_{10}H_{16}$	Sylvestrene	-5.63	200.34	59731.67		11.74		
$C_{15}H_{24}$	α-Bisabolene	-8.83	287.29	85656.41		17.50		Sesquiterpenes
C ₁₅ H ₂₄	β-Bisabolene	-8.29	287.91	85841.26		16.98		
$C_{15}H_{24}$	α-Farnesene	-8.85	304.12	90671.89		18.02		
C ₁₅ H ₃₀	Bisabolane	-7.64	310.18	92480.47		16.98		
$C_{15}H_{30}$	Farnesane	-8.54	329.46	98229.39		18.44		
C ₂₀ H ₃₂	Pinenedimer1	-9.15	340.46	101509.34		19.38		
$C_{20}H_{32}$	Pinenedimer2	-7.87	319.33	95207.94		17.47		
C ₂₀ H ₃₂	Pinenedimer3	-8.05	323.12	96337.63		17.76		

Table S8,a: Calculated ΔH_{vap} values for terpenes (solvent-n-hexane). All values were computed at room temperature (298K).

C ₂₀ H ₃₂	Pinenedimer4	-8.66	338.50	100924.37		18.83	
C ₂₀ H ₃₂	Pinenedimer5	-4.92	186.33	55554.59		10.62	
$C_{20}H_{32}$	Pinenedimer6	-5.55	193.06	57561.73		11.45	
C ₁₀ H ₁₈	Camphane	-5.41	197.31	58828.57		11.43	Saturated Terpenes
$C_{10}H_{18}$	Hydrogenated Car-3-ene	-5.09	204.83	61069.47		11.33	
$C_{10}H_{18}$	Sabinane	-6.66	276.25	82362.45		14.99	
$C_{10}H_{20}$	Limonane	-6.61	271.64	80989.17		14.81	
C ₁₅ H ₂₈	β-Caryophyllane	-6.55	274.93	81970.68		14.85	
$C_{15}H_{28}$	Valencane	-7.43	305.11	90967.35		16.62	
C ₁₅ H ₂₈	Hydrogenated Premnaspirodiene	-7.79	326.82	97440.19		17.61	
Average	absolute deviation				0.3		
Standard	l deviation				0.3		
RMS					0.4		
Maximu	m absolute deviation				1.0		

Regression equation: $y = 1.170093 \cdot x_1 + 0.000084 \cdot x_2 + 0.059668$

Table S8,b: Calculated ΔH_{vap} values for terpenoids (solvent- 1-hexanol). All values were computed at room temperature (298 K).

	Molecules	ΔG _{solv} Solvation Free Energy comp. [kcal/mol] x ₁	Cavity Surface Area (=SASA) [Å ²]	SASA·T [Ų · K] x ₂	Experimental ΔH _{vap} [kcal/mol] γ	Predicted ∆H _{vap} [kcal/mol]	Deviation	
Terpenes us	ed in the multiple	e linear regress	ion:					
C ₁₀ H ₁₈ O	Linalool	-9.45	227.01	67682.14	13.20 ²⁷	14.17	-1.0	Terpenoids
$C_{10}H_{18}O$	Fenchol	-7.90	196.93	58714.08	12.21 ²⁶	12.19	0.0	
$C_{10}H_{18}O$	Terpineol	-11.48	209.31	62404.29	12.90 ²⁶	12.56	0.3	
$C_{10}H_{20}O$	Menthol	-8.07	213.32	63601.66	13.51 ²⁷	13.37	0.1	
$C_{10}H_{20}O$	Citronellol	-10.09	234.43	69895.60	15.17 ²⁷	14.62	0.5	
C ₁₀ H ₁₈ O	Geraniol	-10.83	229.57	68446.00	15.02 ²⁶	14.15	0.9	

C ₁₀ H ₁₈ O	Borneol	-7.93	195.56	58305.62		12.16 ²⁶	12.08	0.1	
C ₁₀ H ₁₈ O	Nerol	-10.94	228.44	68110.58		13.23 ²⁶	14.05	-0.8	
Predicted va	alues:								
C ₁₅ H ₂₆ O	Farnesol	-13.37	323.12	96338.53			20.65		Terpenoids
$C_{10}H_{20}O$	Hydrogenated Terpineol	-8.63	213.58	63677.39			13.31		Saturated Terpenoids
C ₁₀ H ₂₂ O	Hydrogenated Linalool	-8.50	234.81	70007.71			14.89		
C ₁₀ H ₂₂ O	Hydrogenated Citronellol	-9.20	237.48	70805.26			14.98		
C ₁₅ H ₃₂ O	Hydrogenated Farnesol	-10.98	333.95	99566.89			21.82		
Average abs	solute deviation				0.5				
Standard de	eviation				0.4				
RMS					0.6				
Maximum a	bsolute deviation	ı			1.0				

Regression equation: y= -0.15303·x₁ + 0.00025·x₂ - 1.11581

Table S9: Statistic parameters at 5% significance level.

Multiple linear regression					
$(\mathbf{x}_1 = -\Delta \mathbf{G}_{sol}, \mathbf{x}_2 = T \cdot SAS)$	^А , у=ехр.				
ΔH _{vap})					
R ²	0.97				
Standard Error	0.36				
F statistic value	186.24				
F test critical value	4.10				
Significance F	1.22·10 ⁻⁸				
T statistic value-					
variable 1	3.19				
T statistic value-					
variable 2	3.49				
T test critical value	2.23				

Linear regression
(x=-ΔG _{sol} , y=exp. ΔH _{vap})

R ²	0.94
Standard Error	0.52
F statistic value	178.40
F test critical value	4.84
Significance F	3.84·10 ⁻⁸
T statistic value- variable	13.36
T test critical value	2.20

Calculation of all enthalpies of combustion using the computed ΔH_{vap} :

Regular linear regression was applied to calculate ΔH_{comb} , using 12 terpenes in the training set and 7 in the test set. Since the ΔH_{comb} values of the mono- and sesquiterpenes are so far from each other, the R² values has limited statistical meaning since they are biased towards 1. Thus, we created 3 correlation figures using different molecules as training and test sets for each curve, to ensure the correlation remains stable for the different sets. Figure 2 presents one of these curves, Figures S3,a-b presents two additional ones.





Figures S3,a-b: Regression analysis for calculating the ΔH_{comb} with different molecules for training and test sets. These figures present the training set molecules only. Regression equations: (a) y = 1.077697 x + 33.966097 , (b) y = 1.071963 x + 26.879368, R²=0.99.

Table S10: Calculated values of all ΔH_{comb} using ΔH_{vap} correction (Tables S8,a-b) and O₂ correlation curve presented in Fig. 2 (all molecules marked with * were Boltzmann averaged).

	Terpenes	Electronic and Thermal Enthalpy [Ha/particle]	ΔH _{comb} after ΔH _{vap} correction [kcal/mol]	Predicted ΔH _{comb} values, using the correlation fig. equation [kcal/mol]	ΔH _{comb} experimental [kcal/mol]	Deviation [kcal/mol]	
Terpenes ι	used in correlation curve (training set):						
$C_{10}H_{16}$	α-Pinene	-390.24881	-1404.11	-1478.69	-1483.00 ±0.2 ¹⁸	-4.3	Monoterpenes
$C_{10}H_{16}$	β-Pinene	-390.24553	-1406.32	-1481.06	-1485.10 ±0.3 ¹⁸	-4.0	
$C_{10}H_{16}$	Camphene	-390.26682	-1393.09	-1466.89	-1468.92 ±0.5 ¹⁹	-2.0	
$C_{10}H_{16}$	Car-3-ene	-390.25618	-1398.76	-1472.96	-1479.97 ±0.4 ²⁰	-7.0	
$C_{10}H_{16}$	Limonene*	-390.25888	-1397.36	-1471.47	-1473.95 ±0.3 ¹⁸	-2.5	
$C_{10}H_{16}$	Alloocimene*	-390.24163	-1409.40	-1484.36	-1480.32 ± 1.2^{18}	4.0	
$C_{10}H_{16}$	Myrcene*	-390.21999	-1419.70	-1495.39	-1490.50 ±0.2 ¹⁸	4.9	
$C_{10}H_{16}$	Sylvestyrene	-390.25924	-1396.70	-1470.76	-1465.63 ¹⁹	5.1	
$C_{10}H_{16}O$	Terpineol	-466.65733	-1390.74	-1464.38	-1468.78 ± 1.5^{21}	-4.4	Terpenoids
$C_{10}H_{18}$	Pinane	-391.44049	-1445.94	-1523.49	-1512.59 ±1.0 ²⁸	10.9	Hydrogenated Terpenes
$C_{15}H_{28}$	β-Caryophyllane	-587.77650	-2178.23	-2307.72	-2314.31 ¹⁵	-6.6	
C ₁₅ H ₂₈	Valencane	-587.80413	-2161.08	-2289.35	-2283.45 ¹⁵	5.9	
Test set:							
C ₁₀ H ₁₆	Sabinene	-390.24900	-1403.57	-1478.12	-1477.50 ¹⁹	0.6	Monoterpenes
C ₁₀ H ₁₆	γ-Terpinene*	-390.26308	-1393.52	-1467.36	-1472.04 ²¹	-4.7	
C ₁₀ H ₁₈ O	Borneol	-466.66182	-1386.15	-1459.46	-1465.18 ²¹	-5.7	
C ₁₀ H ₂₀ O	Menthol	-467.84885	-1429.40	-1505.78	-1507.07 ²⁹	-1.3	
C ₁₀ H ₁₈	Sabinane	-391.44993	-1439.25	-1516.33	-1507.60 ³⁰	8.7	
C ₁₅ H ₂₈	Hydrogenated-Premanaspirodiene	-587.81227	-2155.93	-2283.83	-2278.97 ³¹	4.9	
C ₁₅ H ₃₂	Farnasane*	-590.15646	-2262.03	-2397.46	-2403.65 ³²	-6.2	
Predicted	values:						
C ₁₀ H ₁₆	β-Phellandrene*	-390.26006	-1396.93	-1471.00			
C ₁₀ H ₁₆	Terpinolene	-390.25954	-1396.17	-1470.20			
C ₁₀ H ₁₈ O	Geraniol*	-466.61164	-1415.46	-1490.85			Terpenoids
C ₁₀ H ₁₈ O	Linalool*	-466.61957	-1411.01	-1486.08			

C ₁₀ H ₁₈ O	Fenchol	-466.66324	-1385.15	-1458.39		
C ₁₀ H ₁₈ O	Nerol	-466.61189	-1415.51	-1490.91		
C ₁₀ H ₂₀ O	Citronellol*	-467.81331	-1450.58	-1528.46		
C ₁₅ H ₂₆ O	Farnesol*	-661.74080	-2113.56	-2238.46		
C ₁₅ H ₂₄	α-Bisabolene*	-585.39014	-2096.44	-2220.13		Sesquiterpenes
C ₁₅ H ₂₄	β-Bisabolene*	-585.38800	-2096.22	-2219.90		
C ₁₅ H ₂₄	α-Farnasene*	-585.35804	-2115.06	-2240.07		
C ₁₅ H ₂₄	β-Farnasene*	-585.35182	-2120.54	-2245.93		
C ₁₅ H ₂₄	β-Caryophyllene	-585.37854	-2104.07	-2228.30	-2206.80 2366.84 ³³	
C ₂₀ H ₃₂	Pinene-dimer 1*	-781.68439	-2068.29	-2189.98		
C ₂₀ H ₃₂	Pinene-dimer 2*	-781.70591	-2053.32	-2173.95		
C ₂₀ H ₃₂	Pinene-dimer 3*	-781.70698	-2051.71	-2172.22		
C ₂₀ H ₃₂	Pinene-dimer 4*	-781.69959	-2058.25	-2179.23		
C ₂₀ H ₃₂	Pinene-dimer 5*	-781.70133	-2056.87	-2177.75		
C ₂₀ H ₃₂	Pinene-dimer 6*	-781.71343	-2048.22	-2168.48		
C ₁₀ H ₁₈	Camphane	-391.46541	-1430.35	-1506.80		Hydrogenated Terpenes
C ₁₀ H ₁₈	Hydrogenated Car-3-ene	-391.44930	-1439.63	-1516.74		
C ₁₀ H ₂₀	Limonane*	-392.65588	-1472.02	-1551.42		
C ₁₅ H ₃₀	Bisabolane*	-588.98725	-2206.26	-2337.73		
C ₁₀ H ₂₀ O	Hydrogenated-Terpineol	-467.84736	-1430.40	-1506.85		
C ₁₀ H ₂₂ O	Hydrogenated-Linalool	-469.02185	-1481.23	-1561.28		
C ₁₀ H ₂₂ O	Hydrogenated-Citronellol	-469.00950	-1488.88	-1569.48		
C ₁₅ H ₃₂ O	Hydrogenated-Farnesol	-665.33485	-2225.35	-2358.17		
Average absolute deviation				4.9		
Standard deviation			2.5			
RMS				5.5		
Maximum	absolute deviation			10.9		

Table S11: Computed specific energy.

	Terpenes	Specific Energy- energy density per kg [MJ/kg]
C ₁₅ H ₃₂	Farnasane	-47.25
C ₁₅ H ₃₀	Bisabolane	-46.52
C ₁₅ H ₂₈	β-Caryophyllane	-46.37
C ₁₀ H ₂₀	Limonane	-46.31
C ₁₀ H ₁₈	Pinane	-46.14
C ₁₅ H ₂₄	β-Farnasene	-46.01
C ₁₅ H ₂₈	Valencane	-46.00
C ₁₀ H ₁₆	Myrcene	-45.96
C ₁₀ H ₁₈	Hydrogenated Car-3-ene	-45.93
C ₁₀ H ₁₈	Sabinane	-45.92
C ₁₅ H ₂₄	α-Farnasene	-45.89
C ₁₅ H ₂₈	Hydrogenated Premnaspirodiene	-45.89
C ₁₅ H ₂₄	β-Caryophyllene	-45.65
C ₁₀ H ₁₈	Camphane	-45.63
C ₁₀ H ₁₆	Alloocimene	-45.62
C ₁₀ H ₁₆	β-Pinene	-45.52
C ₁₅ H ₂₄	α-Bisabolene	-45.49
C ₁₅ H ₂₄	β-Bisabolene	-45.48
C ₁₀ H ₁₆	α-Pinene	-45.44
C ₁₀ H ₁₆	Sabinene	-45.43
C ₁₀ H ₁₆	Car-3-ene	-45.27
C ₁₀ H ₁₆	Limonene	-45.22
C ₁₀ H ₁₆	Phellandrene	-45.21
C ₁₀ H ₁₆	Sylvestyrene	-45.20
C ₁₀ H ₁₆	Terpinolene	-45.18
C ₁₀ H ₁₆	γ-Terpinene	-45.09
C ₁₀ H ₁₆	Camphene	-45.08
C ₁₅ H ₃₂ O	Hydrogenated Farnesol	-43.22

C ₁₅ H ₂₆ O	Farnesol	-42.15
C ₁₀ H ₂₂ O	Hydrogenated Citronellol	-41.51
C ₁₀ H ₂₂ O	Hydrogenated Linalool	-41.30
C ₁₀ H ₂₀ O	Citronellol	-40.95
C ₁₀ H ₁₈ O	Nerol	-40.47
C ₁₀ H ₁₈ O	Geraniol	-40.47
C ₁₀ H ₂₀ O	Hydrogenated Terpineol	-40.37
C ₁₀ H ₂₀ O	Menthol	-40.34
C ₁₀ H ₁₈ O	Linalool	-40.34
C ₁₀ H ₁₈ O	Terpineol	-39.75
C ₁₀ H ₁₈ O	Borneol	-39.61
C ₁₀ H ₁₈ O	Fenchol	-39.58
C ₂₀ H ₃₂	Pinene-dimer 1	-33.65
C ₂₀ H ₃₂	Pinene-dimer 4	-33.49
C ₂₀ H ₃₂	Pinene-dimer 5	-33.46
C ₂₀ H ₃₂	Pinene-dimer 2	-33.40
C ₂₀ H ₃₂	Pinene-dimer 3	-33.38
C ₂₀ H ₃₂	Pinene-dimer 6	-33.32

Computation of the enthalpy of formation:

The enthalpy of formation should be calculated as described in the following equations:¹

(S5)
$$\Delta H_f^{0K}(M) = \sum_{atoms} x \Delta H_f^{0K}(X) - \sum D_0(M) = \sum_{atoms} x \Delta H_f^{0K}(X) - \left[\sum_{atoms} x \varepsilon_0(X) - \varepsilon_0(M) - E_{ZPE}(M)\right]$$

(S6)
$$\Delta H_{f}^{298K}(M) = \Delta H_{f}^{0K}(M) + \left(H^{298K}(M) - E_{ZPE}(M)\right) - \sum_{atoms} x(H^{298K}(X) - H^{0K}(X))$$

 $\Delta H_{f}^{0K}(X) \text{ is the experimental enthalpy of formation of the atoms at 0K (from JANAF³⁴);}$ $\sum_{\sigma_{0}(M)} D_{0}(M) \text{ is the computed atomization energy;}$ $\varepsilon_{0}(X) \text{ is the atomic electronic energy;}$ $(\varepsilon_{0}(M) + E_{ZPE}(M)) \text{ are the electronic energy and ZPE correction of the terpene;}$ $(H^{298K}(M) - E_{ZPE}(M)) \text{ are the calculated enthalpy correction of the terpene minus its ZPE correction;}$ $(H^{298K}(X) - H^{0K}(X)) \text{ is the experimental enthalpy correction of the atomic elements, taken from the Gaussian white page.}$

However, predicting enthalpy of formation directly from the DFT computed data can be difficult for two reasons:

1. The elemental forms are often in the solid state.

2. Atomization energies can be calculated, but it is hard to obtain high accuracy as atoms are often open-shell, and large basis sets are required to maintain a proper balance between singlets (for most molecules) and high spin wave functions (for most atoms, following Hund's rule).^{35,36, 37}

(In atoms and ions, for example, the local-density approximations underestimate the energy required to transfer an s electron to a d shell, and sp transfer energies if the p shell is more than half-full.)

Also, size dependence errors in quantum-mechanical methods should, in general, grow with the size of the system. Since the enthalpies of formation are computed from the calculated atomization energies, a roughly linear dependence on the size of the system can be expected, particularly in a sequence of molecules containing similar kinds of bonds e.g., alkane chains.

Thus, in order to compute atomization energy as accurately as possible, the calculations were performed using G4MP2 theory.³⁷ The results of the enthalpy of formation calculations are presented in Table S12.

Table S12: Computed values of ΔH_f .

	Electronic Energy	Electronic Energy	Electronic and Thermal	Electronic and Zero-Point	ΔH _f (0K) [kcal/mol]	ΔH _f (298K) [kcal/mol]	ΔH _f experimental [kcal/mol]
	Terpenes	[Ha/particle]	[Ha/particle]	[Ha/particle]			
C ₁₀ H ₁₆	α-Pinene	-390.20226	-389.96003	-389.96980	19.47	7.53	7.22 ³⁸

$C_{10}H_{16}$	β-Pinene	-390.20567	-389.96401	-389.97401	16.83	5.04	
C ₁₀ H ₁₆	Camphene	-390.22622	-389.98348	-389.99299	4.92	-7.18	-6.69±0.5 ³⁹
C ₁₀ H ₁₆	Car-3-ene	-390.20633	-389.96460	-389.97522	16.07	4.67	4.54±1.0 ²⁰
C ₁₀ H ₁₆	Limonene	-390.21523	-389.97343	-389.98436	10.33	-0.88	0.74±0.2 ⁴⁰
$C_{10}H_{16}$	Phellandrene	-390.21720	-389.97483	-389.98539	9.69	-1.76	
C ₁₀ H ₁₆	Sabinene	-390.20054	-389.95885	-389.96928	19.80	8.28	
C ₁₀ H ₁₆	γ-Terpinene	-390.21943	-389.97791	-389.98906	7.39	-3.69	-4.90 ⁴¹
C ₁₀ H ₁₆	Terpinolene	-390.21552	-389.97396	-389.98531	9.74	-1.21	
$C_{10}H_{16}$	Alloocimene	-390.19754	-389.95854	-389.97076	18.87	8.47	
$C_{10}H_{16}$	Myrcene	-390.17376	-389.93688	-389.94729	33.59	22.06	
$C_{10}H_{16}$	Sylvestyrene	-390.21587	-389.97401	-389.98490	10.00	-1.24	
C ₁₀ H ₁₈ O	Geraniol	-466.57190	-466.30261	-466.31726	-10.15	-22.08	
C ₁₀ H ₁₈ O	Linalool	-466.57673	-466.30824	-466.32247	-13.42	-25.62	
C ₁₀ H ₁₈ O	Fenchol	-466.62542	-466.35388	-466.36511	-40.17	-54.25	
C ₁₀ H ₁₈ O	Terpineol	-466.61625	-466.34584	-466.35828	-35.89	-49.21	
C ₁₀ H ₁₈ O	Borneol	-466.62472	-466.35311	-466.36445	-39.76	-53.77	
C ₁₀ H ₁₈ O	Nerol	-466.57375	-466.30402	-466.31819	-10.73	-22.97	
$C_{10}H_{20}O$	Menthol	-467.84030	-467.54555	-467.55821	-55.45	-70.66	
$C_{10}H_{20}O$	Citronellol	-467.80533	-467.51222	-467.52706	-35.91	-49.74	
$C_{15}H_{26}O$	Farnesol	-661.68042	-661.28831	-661.30952	-5.03	-22.18	
C ₁₅ H ₂₄	α-Bisabolene	-585.32518	-584.96097	-584.97876	14.11	-2.12	
C ₁₅ H ₂₄	β-Bisabolene	-585.32397	-584.95942	-584.97684	15.31	-1.15	
C ₁₅ H ₂₄	α-Farnasene	-585.29501	-584.93209	-584.95127	31.36	16.00	
C ₁₅ H ₂₄	β-Farnasene	-585.28822	-584.92555	-584.94468	35.49	20.10	
C ₁₅ H ₂₄	β-Caryophyllene	-585.31424	-584.94901	-584.96477	22.89	5.38	
$C_{10}H_{18}$	Pinane	-391.42916	-391.16323	-391.17356	-2.49	-16.10	
C ₁₀ H ₁₈	Camphane	-391.45667	-391.19028	-391.20012	-19.16	-33.07	
C ₁₀ H ₁₈	Dihydro-Car-3-ene	-391.43184	-391.16609	-391.17691	-4.60	-17.89	
C ₁₀ H ₁₈	Sabinane	-391.43350	-391.16819	-391.17914	-6.00	-19.21	
C ₁₀ H ₂₀	Limonane	-392.67974	-392.39070	-392.40228	-40.12	-54.96	
C ₁₅ H ₂₈	β-caryophyllane	-587.77247	-587.35998	-587.37651	-23.69	-44.75	
C ₁₅ H ₂₈	Valencane	-587.80618	-587.39241	-587.40872	-43.90	-65.11	

C ₁₅ H ₂₈	Hydrogenated- Premnaspirodiene	-587.81395	-587.40070	-587.41703	-49.12	-70.31
C ₁₅ H ₃₀	Bisabolane	-589.01980	-588.58427	-588.60273	-59.75	-81.62
C ₁₅ H ₃₂	Farnasane	-590.21942	-589.76235	-589.78326	-67.14	-89.50
C ₁₀ H ₂₀ O	Hydrogenated - Terpineol	-467.83847	-467.54406	-467.55681	-54.57	-69.72
C ₁₀ H ₂₂ O	Hydrogenated - Linalool	-469.04386	-468.72782	-468.74262	-65.28	-81.15
C ₁₀ H ₂₂ O	Hydrogenated - Citronellol	-469.03366	-468.71676	-468.73183	-58.51	-74.22
C ₁₅ H ₃₂ O	Hydrogenated - Farnesol	-665.36954	-664.90604	-664.92780	-75.32	-98.19
	02	-150.19472	-150.18862	-150.19099		
	CO2	-188.42041	-188.40614	-188.40875		
	H2O	-76.37692	-76.35301	-76.35585		

In Table S13 we present the detonation properties of the most energetic terpenes (the ones with the most negative enthalpy of combustion), calculated by the EXPLO5 program.^{42,43} The input data required for these calculations are the computed enthalpy of formation, the experimental density and the molecular formula of the required terpenes.

The EXPLO5 computer program is based on the chemical equilibrium steady-state model of detonation. The state of gaseous detonation products is described by the Becker-Kistiakowsky- Wilson equation of state.

The equations' system describing chemical equilibrium in detonation products is solved by a modified Newton-Raphson method. Further theoretical details can be found in Ref. ^{42,43}.

These calculations present another computational approach for calculating the inherent properties of terpenes, and might be useful for determining whether terpene molecules can serve as explosives.

	Terpenes	Density [g/cm³] ª	Heat of detonation [KJ/kg]	Detonation temperature [K]	Detonation pressure [GPa]	Detonation velocity [m/s]	Particle velocity [m/s]	Density of products [g/cm³]	Volume of products [cm ³ /g]
C ₁₅ H ₃₂	Farnasane	0.768	-128.5	506.0	1.61	3327.3	628.1	0.9	1.1
C ₁₅ H ₃₀	Bisabolane	0.820	-74.8	481.0	1.69	3339.6	618.9	1.0	1.0
C ₁₅ H ₂₈	β-caryophyllane	0.850	-593.0	716.9	2.47	3867.9	752.4	1.1	0.9
$C_{10}H_{20}$	Limonane	0.841	-49.8	466.4	1.75	3415.6	607.9	1.0	1.0
C ₁₀ H ₁₈	Pinane	0.857	-856.7	832.2	2.85	3993.4	833.6	1.1	0.9
$C_{10}H_{16}$	Myrcene	0.794	-1619.4	1112.6	2.39	3695.3	815.9	1.0	1.0
C ₁₅ H ₂₄	β-Farnasene	0.813	-1382.9	1028.7	2.38	3689.1	792.7	1.0	1.0
C ₁₅ H ₂₄	α-Farnasene	0.813	-1293.5	999.1	2.35	3646.8	793.0	1.0	1.0
C ₁₅ H ₂₈	Valencane	0.879	-206.5	546.5	2.13	3634.5	665.3	1.1	0.9
C ₁₀ H ₁₈	Dihydro-Car-3-ene	0.867	-786.8	797.8	2.75	4000.2	792.9	1.1	0.9
C ₁₀ H ₁₈	Sabinane	0.844	-751.3	784.3	2.51	3842.7	774.7	1.1	0.9

Table S13: Computed detonation properties using EXPLO5.42,43

^aSigma-Aldrich

The detonation data was computed for the terpenes that had the highest specific energy.

Estimation of Cetane number:

Fuels for compression ignition engines must ignite readily through autoignition alone. If ignition does not occur promptly when the fuel is injected into the cylinder, unburned fuel will accumulate as the injection process proceeds, and when ignition does occur, the rate of burning will be too rapid, resulting in engine knock- which decreases efficiency while increasing engine noise and wear. Thus, the ability to rate the ignition quality of compression-ignition fuels is important for diesel fuel formulation. Without adequate fuel ignition quality the engine will start with difficulty and run poorly.^{44–46}

The cetane number is an indication of the ignition delay, measured as the time delay between the decided beginning of the injector needle lift and the chamber pressure recovery point.⁴⁴ The factors estimated to influence this delay can be divided into physical processes and chemical processes. The physical processes influence the time required for a droplet of fuel to heat, vaporize, and mix with hot air in the cylinder, causing the "physical delay" in ignition. This delay is affected by the fuel's enthalpy of vaporization, it's heat capacity, viscosity, density, vapor pressure and more. The chemical processes influence the radical forming oxidation cascade occurring in the engine. In this case the structure of the fuel's molecule plays an important role, in terms of bond strength and the ability to stabilize the formed radicals. Weaker bonds will be broken to form radicals, which initiates other bond breaking reactions to forms more radicals, and once a sufficient concentration of free radicals is reached, rapid oxidation occurs.⁴⁴

For all but very heavy fuels, chemical ignition delay is a dominant factor over the physical ignition delay for determining the measured ignition delay time.⁴⁴

In this work, we combine three calculated properties in a multiple linear regression calculation to estimate the cetane number values, based on experimental cetane number data: the calculated enthalpy of vaporization, the computed heat capacity (given by the electronic structure calculations) and the global radical formation rate, log k, (a logarithmic expression of the k rate constant, weighted by the number of equivalent hydrogens in each transition state structure). The values used are presented in Table S14.

The kinetic model used for calculating the rate constants is based on conventional transition state theory (TST).^{47–50} The following equation was used to calculate the rate constant for each calculated electronic energy barrier:

S7)
$$k\left[\frac{cm^{3}}{mol s}\right] = \frac{k_{B}T_{RT}}{h P_{0}}e^{\frac{-\Delta E_{e}^{\dagger}}{RT}}$$

The temperature used was 1000 K, as this temperature is considered representative of the temperature in the combustion chamber of the Diesel engine.⁵⁰

The global log k is calculated as explained in the following scheme:

$$Globallog k = \sum_{i}^{N} (r_i / \sum_{j}^{N} r_j) \cdot \log k_i$$
(S8)

The molecule M has N possible transition state structures. The sum runs over all calculated structures, as the transition state structure i will have its log k calculated using the above (S7) equation, and then multiplied by r_i - the number of equivalent hydrogens for the transition state structure i, divided by the total number of equivalent hydrogens in all N transition state structures of the molecule M. This creates a weight of the probability for the oxidation reaction to occur from a specific transition state. (the global log k equation (S8) used in this paper is derived from the global log k equation presented in ref. ³⁶)

For farnesene, the experimental data did not specify if it is for α - or β -farnesene, thus the training set includes a global log k that includes both. For most molecules, the TS structure is calculated for one specific conformer, each structure describing a different C-H bond break, the molecules that has more than one conformer in their TS calculations are specified. The structures of all TS conformers are presented in Table S17.

Table S14,a: Computed values used to estimate the Cetane number values.

Chemical formula	Terpenes	Electronic energy barrier ΔE _e [†] (at 1000K) [kcal/mol]	k (E _e) [cm³/mol·s]	log k (ΔE _e †)	Number of equivalent hydrogens
C ₁₀ H ₁₆	β-pinene TS 1	35.93	2.40E+10	10.38	1
C ₁₀ H ₁₆	β-pinene TS 2	44.50	3.21E+08	8.51	1
C ₁₀ H ₁₆	α-pinene TS 1	36.85	1.51E+10	10.18	3
C ₁₀ H ₁₆	α-pinene TS 2	37.90	8.91E+09	9.95	1
C ₁₀ H ₁₆	3-carene TS 1	33.63	7.64E+10	10.88	1
C ₁₀ H ₁₆	3-carene TS 2	41.93	1.17E+09	9.07	1
C ₁₀ H ₁₆	3-carene TS 3	45.09	2.39E+08	8.38	3
C ₁₀ H ₁₆	3-carene TS 4	46.16	1.40E+08	8.15	3
C ₁₀ H ₁₆	Limonene- axial conformer TS 1	36.33	1.97E+10	10.29	1
C ₁₀ H ₁₆	Limonene-axial conformer TS 2	38.82	5.62E+09	9.75	1
C ₁₀ H ₁₆	Limonene-axial conformer TS 3	39.02	5.08E+09	9.71	1
C ₁₀ H ₁₆	Limonene- equatorial conformer TS 1	36.40	1.90E+10	10.28	3
C ₁₀ H ₁₆	Limonene- equatorial conformer TS 2	36.85	1.51E+10	10.18	3
C ₁₀ H ₁₆	Limonene- equatorial conformer TS 3	37.43	1.13E+10	10.05	1
C ₁₀ H ₁₆	Limonene- equatorial conformer TS 4	41.59	1.39E+09	9.14	1
C ₁₀ H ₁₆	Limonene- equatorial conformer TS 5	39.64	3.72E+09	9.57	1
C ₁₀ H ₁₆	Limonene- equatorial conformer TS 6	46.68	1.07E+08	8.03	1
C ₁₀ H ₁₆	γ-Terpinene conformer 1 TS1	29.45	6.25E+11	11.80	1
C ₁₀ H ₁₆	γ-Terpinene conformer 1 TS2	32.49	1.35E+11	11.13	1
C ₁₀ H ₁₆	γ-Terpinene conformer 1 TS 3	32.99	1.05E+11	11.02	1
C ₁₀ H ₁₆	γ-Terpinene conformer 2 TS 1	37.02	1.39E+10	10.14	3
C ₁₀ H ₁₆	γ-Terpinene conformer 2 TS 2	38.97	5.21E+09	9.72	3
C ₁₀ H ₁₆	γ-Terpinene conformer 2 TS 3	49.57	2.51E+07	7.40	3
C ₁₀ H ₁₆	γ-Terpinene conformer 2 TS 4	49.82	2.21E+07	7.35	1
C ₁₀ H ₁₆	γ-Terpinene conformer 2 TS 5	37.28	1.21E+10	10.08	1
C ₁₀ H ₁₆	γ-Terpinene conformer 2 TS 6	29.45	6.25E+11	11.80	1

C ₁₀ H ₁₈	Pinane TS	38.69	5.98E+09	9.78	1
C ₁₀ H ₁₈	Sabinane TS	39.55	3.88E+09	9.59	1
C ₁₀ H ₂₀	Limonane- axial conformer TS 1	39.10	4.87E+09	9.69	1
C ₁₀ H ₂₀	Limonaneaxial conformer TS 2	42.29	9.78E+08	8.99	1
C ₁₀ H ₂₀	Limonane- equatorial conformer TS 1	42.28	9.83E+08	8.99	1
C ₁₀ H ₂₀	Limonane- equatorial conformer TS 2	42.50	8.80E+08	8.94	1
C ₁₀ H ₂₀	Limonane- equatorial conformer TS 3	42.50	8.80E+08	8.94	1
C ₁₅ H ₂₄	β-Bisabolene TS 1	35.55	2.90E+10	10.46	1
C ₁₅ H ₂₄	β-Bisabolene TS 2	37.57	1.05E+10	10.02	1
C ₁₅ H ₂₄	β-Bisabolene TS 3	38.40	6.93E+09	9.84	1
C ₁₅ H ₂₄	β-Bisabolene TS 4	37.64	1.02E+10	10.01	3
C ₁₅ H ₂₄	β-Bisabolene TS 5	38.23	7.56E+09	9.88	1
C ₁₅ H ₂₄	β-Caryophyllene	32.17	1.59E+11	11.20	1
C ₁₅ H ₂₄	α-farnesene TS 1	29.91	4.96E+11	11.70	1
C ₁₅ H ₂₄	α-farnesene TS 2	31.06	2.79E+11	11.45	1
C ₁₅ H ₂₄	α-farnesene TS 3	36.98	1.41E+10	10.15	1
C ₁₅ H ₂₄	β-farnesene TS 1	35.26	3.37E+10	10.53	1
C ₁₅ H ₂₄	β-farnesene TS 2	37.27	1.22E+10	10.09	1
C ₁₅ H ₂₈	β-Caryophyllane	42.12	1.02E+09	9.01	1
C ₁₅ H ₃₂	Farnesane TS 1	42.23	1.01E+09	9.00	1
C ₁₅ H ₃₂	Farnesane TS 2	42.29	9.78E+08	8.99	1
C ₁₅ H ₃₂	Farnesane TS 3	42.61	8.33E+08	8.92	1
C ₁₅ H ₃₂	Farnesane TS 4	46.31	1.29E+08	8.11	1

Another MLR calculation was performed to estimate whether the heat capacity value had any influence on the data, and the following data was obtained:

Table S14,b: Statistic parameters at 5% significance level

Multiple linear regression – 3 variables:		Multiple linear regression – 2 variables:		
$x_1=\Delta H_{vap}$, $x_2=C_v$, $x_3=$ global log k , y=exp. CN		$x_1=\Delta H_{vap}$, x_2 = global log k , y=exp. CN		
y = 1.033869 - 1.348807 x ₂ - 6.	603378 x ₃ + 60.143412	y = 3.497646 x ₁ - 7.9812	81 x ₂ + 61.70000	
R ²	0.91	R ²	0.86	
Standard Error	4.96	Standard Error	5.49	

F statistic value	16.51	F statistic value
F test critical value	5.41	F test critical value
Significance F	0.005	Significance F
T statistic value- variable 1	1.52	T statistic value- variable 1
T statistic value- variable 2	-0.42	T statistic value- variable 2
T statistic value- variable 3	-2.76	T test critical value
T test critical value	2.57	

19.32 5.14 0.002 5.48 -3.26 2.45

The values obtained from both tests (the 3 variable MLR and 2 variable MLR) are not very different from one other: the R₂ correlation value of the 3-variable MLR and also the standard error is better, but the F test critical value of the 2-variable MLR is slightly higher than the 3 variable MLR's. This indicates that both models give about the same level of accuracy.

Another way to estimate the CN is by correlating it with the computed specific energy values. A better fuel will have a low (negative) specific energy, or in other words, will release more energy upon combustion, and a better fuel will also have higher CN, indication low ignition delay. Therefore, good correlation suggests that terpenes may serve as a suitable potential biofuel alternative.

The 3-variable MLR model has better correlation with the experimental Cetane number, but this model might be easier to calculate.

Terpenes	Experimental Cetane number	Specific energy [MJ/mol]
α-Pinene	17.1	-45.44
β-Pinene	19	-45.52
Car-3-ene	27	-45.27
Limonene	17.1	-45.22
γ-Terpinene	18.7	-45.09
β-Bisabolene	32.6	-46.52
β-Caryophyllene	29	-45.65
α-Farnesene	32	-45.89
Farnesane	58	-47.25

Tables S15,a: Correlation between the specific energy and the experimental cetane number, correlation equation: y =-17.9016·x – 790.036 , R²=0.88.

Terpenes	Predicted CN:
α-Pinene	24.07
β-Pinene	25.41
Camphene	17.38
Car-3-ene	20.82
Limonene	19.97
Phellandrene	19.71
Sabinene	23.74
γ-Terpinene	17.64
Terpinolene	19.25
Alloocimene	27.28
Myrcene	33.54
Sylvestyrene	19.57
α-Bisabolene	24.86
β-Bisabolene	24.77
α-Farnasene	32.39
β-Farnasene	34.61
β-Caryophyllene	27.94
Pinane	36.87
Camphane	27.55
Hydrogenated Car-3-ene	33.10
Sabinane	32.87
Limonane	40.02
β-Caryophyllane	41.08
Valencane	34.28
Hydrogenated Premnaspirodiene	32.23
Bisabolane	43.91
Farnasane	57.48

Tables S15,b: Cetane numbers derived from the linear regression equation correlating the specific energy and experimental CN.

Computation of the Boiling Point and Vapor Pressure:

Both the boiling point and the vapor pressure were calculated using the ADF COSMO-RS (COnductor like Screening MOdel for Realistic Solvents) program.⁵¹ This program can be used for calculating thermodynamic properties of fluids (and mixtures) by using the COSMO-RS method, developed by Klamt and coworkers,^{52–54} and the implementation of COSMO-RS in ADF is described in Ref. ⁵⁵ (based on the COSMO-RS method developed by Klamt et al⁵⁴). With COSMO-RS it is possible to use a thermodynamically consistent combinatorial contribution to the chemical potential (μ).⁵³

The COSMO-RS method allows calculation of the (pseudo-)chemical potential of a compound in the liquid phase, as well as in the gas phase. The liquid phase is described as an effective continuum surrounding a single 'solute' molecule which is treated explicitly using QM methods. The difference of the COSMO model from other continuum solvation models is that instead of considering the fluids as ensembles of molecules which are interacting via static and induced electric fields as well as via vdW interactions, they are looked at as ensembles of ideally screened molecules contacting on the vdW surface with **pairwise** interactions of adjacent surface charge densities and pairwise dispersive interactions through the contact areas. All interactions are local and pairwise, and the only ingredients to the new description are the ideal, i.e., conductor-like, screening charge densities on the surfaces of the molecules, which can be derived from a single molecule's molecular orbital calculation with a continuum model.⁵⁴ Using these charge densities, the chemical potentials of the compounds in the liquid phase are derived. Further details are presented in ref. ⁵⁵.

Vapor Pressures: The logarithm of the vapor pressure of a liquid X is proportional to the difference of the chemical potentials of a molecule of X in the gas phase and the liquid phase X.^{53,54} The COSMO-RS model calculates the chemical potential in the gas phase (perfect gas with a reference state of 1 bar) and the chemical potential of the liquid phase, and gives the vapor pressure (in units of bar) using the following equation:

(S9)
$$p_{i}^{vapor} = exp(\mu_{i}^{pure} - \mu_{i}^{gas})/RT$$

Boiling Temperature: The calculation of the boiling temperature of a solvent is performed with an iterative method. The temperature is varied until the calculated vapor pressure is within a certain threshold of the desired pressure.

The model also features general parameters not specific to chemical groups or functionalities.

All calculated values show good correlation with the experimental values, as shown is Figures S4,a-b and Figures 5,a-b.

In addition to these two properties, the enthalpy of vaporization can be calculated using the COSMO-RS method as well, with the following equation:

(S10)
$$\Delta H_{vap} = RT^2 / p^{vapor} \partial \{p^{vapor}\} / \partial T$$

The calculated values for the ΔH_{vap} are given below in Table S16, and its correlation to the experimental data is shown in Figure S4,c. However, we find that the model presented in this project for converting the computed free energy of solvation (calculated using the SMD solvation model) to the enthalpies of vaporization show better correlation with the experimental enthalpies of vaporization values. We note that employing experimental vapor pressure in Eq. S10 can yield more accurate values.



Figures S4,a-b: Regression analysis for calculating (a) the Boiling point ($R^2=0.72$, $y=0.51824 \cdot x + 204.32862$); (b) the Enthalpy of vaporization ($R^2=0.80$, $y=0.32743 \cdot x + 8.59204$). The presented data is for the training set molecules only. The vapor pressure calculations had no need for a correlation figure.

In Figures S5,a-d we compare the experimental vapor pressure of the following terpenes with their computed values over a wide range of temperatures, which were calculated using the COSMO-RS method:



Figures S5,a-d: Comparison between the computed vapor pressure values and experimental vapor pressure data^{56,57} of the terpenes α -pinene (a), β -pinene (b), limonene (c) and verbenone (d).

Table S16a: Computed vapor pressures using the COSMO-RS method (no correlation figure was used). The values marked in red are values that show large deviation from the experimental values, suggesting the model didn't predict their vapor pressures as well as the other (these values were emitted from the RMS and average deviation calculation). All values are computed at 298K.

	Computed Vapor pressure [bar]	Experimental VP [bar] (at 298K)	Exp. error	Deviation (computed from experimental)
α-Pinene	4.60E-03	5.77E-03 ^{27,56,58-60}	± 2.2E-04	1.17E-03
β-Pinene	4.65E-03	3.95E-03 ^{27,56,58,60,61}	± 4.6E-05	6.98E-04
Camphene	5.63E-03	5.04E-03 ⁶²	± 1.2E-06	5.93E-04
Car-3-ene	4.26E-02	2.60E-03 ⁵⁷	± 1.7E-07	4.00E-02
Limonene	2.59E-03	1.94E-03 ^{27,57-60}	± 8.7E-05	6.44E-04
β-Phellandrene	2.06E-02	3.33E-03 ⁶³		1.73E-02
Sabinene	2.35E-02	3.30E-03 ²⁷	± 2.0E-04	2.00E-02
γ-Terpinene	2.14E-03	1.24E-03 ^{27,59}	± 2.1E-04	9.00E-04
Terpinolene	2.05E-03	9.90E-04 ⁵⁹		1.06E-03
α-Phellandrene	2.21E-03	1.91E-03 ^{60,64}	± 4.3E-05	2.99E-04
Alloocimene	3.40E-03			
Myrcene	4.03E-03	2.61E-03 ^{58,64}	± 9.9E-05	1.42E-03
Sylvestrene	2.05E-02			
Geraniol	5.21E-06	4.00E-05 ⁶⁵	± 4.3E-08	3.48E-05
Linalool	7.60E-05	2.35E-04 ^{27,58,59,61}	± 2.7E-05	1.59E-04
Fenchol	8.24E-04	1.50E-04 ⁵⁸	± 3.0E-05	6.74E-04
Terpineol	1.98E-04	5.69E-05 ⁵⁹		1.41E-04
Borneol	5.52E-04	3.65E-05 ^{58,62,65}	± 1.2E-05	5.16E-04
Nerol	7.09E-06			
Menthol	2.73E-04	1.49E-04 ^{27,29}	± 3.3E-05	1.24E-04
Citronellol	9.39E-06	3.82E-05 ^{27,29,65}	± 1.9E-05	2.88E-05
Farnesol	2.00E-08	5.25E-08 ²⁹		3.25E-08
α-Bisabolene	6.67E-05	2.11E-05 ^{<i>a</i>}		4.56E-05
β-Bisabolene	6.30E-05	2.93E-05 ^{<i>a</i>}		3.37E-05
α-Farnasene	1.66E-05			
β-Farnasene	1.42E-05	1.33E-05 ²⁷		9.08E-07
β-Caryophyllene	6.80E-04	4.61E-05 ²⁷		6.33E-04
Pinane	7.16E-02	3.77E-03 ⁵⁷	± 1.0E-07	6.78E-02
Camphane	9.31E-02			

Hydrogenated Car-3-ene	4.81E-02
Sabinane	4.46E-02
Limonane	2.78E-02
β-Caryophyllane	7.10E-04
Valencane	1.03E-03
Hydrogenated Premnaspirodiene	9.72E-04
Bisabolane	9.82E-05
Farnasane	1.70E-05
Average absolute deviation	4.59E-04
Standard deviation	4.37E-04
RMS	6.26E-04
Maximum absolute deviation	1.42E-03

^awww.chemspider.com (predicted using EPISuite software)

Table S16b: Predicted boiling points using the linear correlation equation presented in Figures S4,a, and the computed values using the COSMO-RS method. All values are computed at 298K.

	Computed Boiling Point [K]	Predicted Boiling Point [K]	Experimental Boiling Point [K]	Exp. error	Deviation (predicted - experimental)
Training set:					
α-Pinene	476.38	451.21	429.35 ²⁷	± 2.5	22.1
β-Pinene	474.72	450.35	438.15 ²⁷		12.2
Camphene	468.37	447.06	433.65 ⁶⁴	± 3.2	13.4
Car-3-ene	402.01	412.67	446.33 ^{27,66}	± 0.2	33.7
Limonene	492.28	459.45	449.65 ⁶⁷	± 1.5	9.8
γ-Terpinene	499.28	463.08	455.65 ^{68,69}	± 0.5	7.4
Terpinolene	500.64	463.78	459.20 ⁶⁸		4.6
α-Phellandrene	498.16	462.50	448.15 ⁶⁴		14.3
Myrcene	474.80	450.39	444.65 ⁶⁴		5.7
Geraniol	577.13	503.42	503.15 ⁶⁴		0.3
Nerol	558.33	493.68	499.15 ⁶⁴		5.5
Citronellol	564.31	496.78	497.65 ²⁷	± 0.3	0.4
β-Farnasene	658.25	545.46	541.80 ²⁷		3.7

β-Caryophyllene	538.02	483.15	537.15 ²⁷	± 3.8	54.0
Test set:					
β-Phellandrene	424.12	424.13	444.65 ^a	± 0.5	20.5
Sabinene	419.59	421.78	436.70 ²⁷	± 2.5	14.9
Alloocimene	479.53	452.84	461.15 ⁶⁷	± 1.0	8.3
Linalool	534.05	481.09	470.15 ²⁷	±0.1	10.9
Fenchol	459.88	442.66	474.15 ⁶⁴	± 1.0	31.5
Terpineol	496.18	461.47	490.70 ⁶⁷	± 0.5	29.2
Borneol	462.64	444.09	486.4070	± 0.5	42.3
Menthol	497.48	462.14	485.15 ⁶⁴	± 0.7	23.0
α-Bisabolene	612.05	521.52	549.15 ^b	± 8.0	27.6
β-Bisabolene	613.27	522.15	547.65 ^b	± 6.8	25.5
α-Farnasene	652.77	542.62	534.15 ^b	± 1.0	8.5
Pinane	386.01	404.37	436.4071	± 3.0	32.0
Bisabolane	606.82	518.81	538.45 ^b	± 7.0	19.6
Farnasane	665.88	549.42	522.25 ^b	± 1.0	27.2
Predicted values:					
Sylvestrene	424.46	424.30			
Farnesol	752.21	594.15			
Camphane	377.17	399.79			
Hydrogenated Car-3-ene	399.01	411.11			
Sabinane	401.28	412.29			
Limonane	417.64	420.77			
β-Caryophyllane	540.15	484.26			
Valencane	527.81	477.86			
Hydrogenated Premnaspirodiene	529.78	478.88			
Average absolute deviation		18.15			
Standard deviation		13.23			
RMS		22.32			
Maximum absolute deviation		54.00			

^ahttp://www.hmdb.ca/metabolites/HMDB0041634 ; ^bwww.chemspider.com (predicted using ACD/Labs software)

Table S16c: Predicted data of enthalpy of vaporization, using the linear correlation equation presented in figures S4,b, and the computed values of the COSMO-RS method. All values are computed at 298K

	Computed ΔH _{vap} [kcal/mol]	Predicted ΔH _{vap} [kcal/mol]	Experimental ΔH _{vap} [kcal/mol]	Deviation (predicted - experimental)
Training set:				
α-Pinene	8.57	11.40	10.84 ²⁴	0.6
β-Pinene	8.62	11.41	10.94 ²⁵	0.5
Camphene	8.51	11.38	10.66 ±0.2 ^{26,27}	0.7
Limonene	9.04	11.55	11.77 ²⁷	0.2
γ-Terpinene	9.13	11.58	12.27 ²⁷	0.7
Terpinolene	9.15	11.59	12.06 ²⁶	0.5
α-Phellandrene	9.11	11.58	11.39 ²⁹	0.2
Myrcene	8.93	11.51	12.09 ²⁶	0.6
Geraniol	18.09	14.51	15.02 ²⁶	0.5
Nerol	18.29	14.58	13.23 ²⁶	1.3
Citronellol	17.58	14.35	15.17 ²⁷	0.8
Test set:				
Car-3-ene	7.29	10.98	11.56 ²⁰	0.6
β-Phellandrene	7.83	11.16	11.54 ²⁶	0.4
Sabinene	7.76	11.13	11.20 ²⁷	0.1
Linalool	14.94	13.48	13.20 ²⁷	0.3
Fenchol	13.74	13.09	12.21 ²⁶	0.9
Terpineol	14.77	13.43	12.90 ²⁶	0.5
Borneol	0.00	8.59	12.16 ²⁶	3.6
Menthol	14.15	13.22	13.51 ²⁷	0.3
β-Farnasene	12.33	12.63	17.32 ²⁷	4.7
β-Caryophyllene	9.78	11.80	15.65 ²⁷	3.9
Pinane	6.91	10.85	9.98 ²⁶	0.9
Predicted values:				
Alloocimene	9.06	11.56		
Sylvestrene	7.83	11.15		
Farnesol	21.72	15.70		
α-Bisabolene	11.26	12.28		

β-Bisabolene	11.32	12.30
α-Farnasene	12.25	12.60
Camphane	6.76	10.80
Hydrogenated Car-3-ene	7.16	10.94
Sabinane	7.22	10.96
Limonane	7.45	11.03
β-Caryophyllane	9.62	11.74
Valencane	9.40	11.67
Hydrogenated Premnaspirodiene	9.43	11.68
Bisabolane	10.78	12.12
Farnasane	11.82	12.46
Average absolute deviation		1.02
Standard deviation		1.27
RMS		1.61
Maximum absolute deviation		4.69

REFERENCES

- 1 R. Dennington, T. Keith and J. Millam, Semichem Inc. , Shawnee Mission. KS, 2009.
- M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. M. Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Gaussian 09, Revision E.01*, 2009.
- 3 G. Zheng, H. A. Witek, P. Bobadova-Parvanova, S. Irle, D. G. Musaev, R. Prabhakar, K. Morokuma, M. Lundberg, M. Elstner, C. Köhler and T. Frauenheim, J. Chem. Theory Comput., 2007, 3, 1349–1367.
- 4 Y. Zhao and D. G. Truhlar, Acc. Chem. Res., 2008, 41, 157–167.
- 5 A. D. Becke, J. Chem. Phys., 1993, **98**, 5648–5652.
- 6 S. Grimme, J. Antony, S. Ehrlich and H. Krieg, J. Chem. Phys., 2010, **132**, 1–19.
- 7 L. Goerigk and S. Grimme, J. Chem. Theory Comput., 2011, **7**, 291–309.
- 8 S. Grimme, S. Ehrlich and L. Goerigk, *J. Comput. Chem.*, 2011, **32**, 1456–1465.
- 9 A. V. Marenich, C. J. Cramer and D. G. Truhlar, J. Phys. Chem. B., 2009, **113**, 6378–6396.
- 10 J. W. Ochterski, *Thermochemistry in Gaussian*, 2000.
- 11 P. Atkins and J. De Paula, *Atkins' Physical chemistry 8th edition*, 2009.
- 12 H. Sun, J. Phys. Chem. B, 1998, B102, 7338–7364.
- 13 H. Sun, P. Ren and J. R. Fried, *Comput. Theor. Polym. Sci.*, 1998, **8**, 229–246.
- 14 P. Mori-sánchez, A. J. Cohen and W. Yang, J. Chem. Phys., 2006, 125, 1–4.
- 15 O. Gunnarsson and R. O. Jones, *Phys. Rev. B*, 1985, **31**, 7588–7602.
- 16 L. Wang, T. Maxisch and G. Ceder, *Phys. Rev. B*, 2006, **73**, 1–6.
- 17 V. L. Chevrier, S. P. Ong, R. Armiento, M. K. Y. Chan and G. Ceder, *Phys. Rev. B*, 2010, **82**, 1–11.
- 18 J. E. Hawkins and W. T. Eriksen, J. Am. Chem. Soc., 1954, 76, 2669–2671.
- 19 K. Auwers, W. A. Roth and F. Eisenlohr, Justus Liebigs Ann. Chem., 1910, 373, 267–290.
- 20 M. P. Kozina, V. A. Aleshina, G. L. Gal'chenko, E. F. Buinova and I. I. Bardyshev, Vestsi Akad. Nauk. BSSR, Ser. Khim. Nauk., 1975, 14–17.
- 21 M. S. Kharasch, in Bureau of Standards Journal of Research (vol.2), 1929, pp. 359–430.
- 22 K. A. Sharp, A. Nicholls, R. Fine and B. Honig, *Science*, 1991, **252**, 106–109.
- 23 B. Lee and F. M. Richards, J. Mol. Biol., 1971, 55, 379–490.
- 24 R. A. Clará, A. C. G. Marigliano and H. N. Sólimo, J. Chem. Eng. Data, 2009, 54, 1087–1090.
- 25 X. An, R. Hu, H. Wang, M. Wu and Y. Zou, Acta Phys. Chim. Sin., 1987, 3, 668–671.
- 26 R. M. Stephenson and S. Malanowski, Handbook of the Thermodynamics of Organic Compounds, 1987.
- 27 M. Hoskovec, D. Grygarová, J. Cvačka, L. Streinz, J. Zima, S. P. Verevkin and B. Koutek, J. Chromatogr. A, 2005, 1083, 161–172.
- 28 M. H. Gollis, L. I. Belenyessy, B. J. Gudzinowicz, S. D. Koch, J. O. Smith and R. J. Wineman, J. Chem. Eng. Data, 1962, 7, 311–316.
- 29 J. W.E. Acree and J. S. Chickos, in NIST Chemistry WebBook, NIST Standard Reference Database Number 69, Eds. P.J. Linstrom and W.G. Mallard, National Institute of Standards and Technology, Gaithersburg MD, 20899.
- 30 W. A. Roth and G. J. Östling, Berichte der Dtsch. Chem. Gesellschaft, 2018, 46, 309–327.
- 31 B. G. Harvey, H. A. Meylemans, R. V. Gough, R. L. Quintana, M. D. Garrison and T. J. Bruno, *Phys. Chem. Chem. Phys.*, 2014, 16, 9448–9457.
- 32 F. Millo, S. Bensaid, D. Fino, S. J. C. Marcano, T. Vlachos and B. K. Debnath, *Fuel*, 2014, **138**, 134–142.
- 33 J. Lane, 9 Advanced Molecules That Could Revolutionize Jet and Missile Fuel, 2014.
- 34 M. W. Chase, NIST-JANAF Thermochemical tables (Journal of Physical and Chemical Reference Data Monographs), American Institute of Physics, 1983.
- 35 P. C. Redfern, P. Zapol, L. A. Curtiss and K. Raghavachari, J. Phys. Chem. A, 2000, 104, 5850–5854.
- 36 L. A. Curtiss, P. C. Redfern and K. Raghavachari, J. Chem. Phys., 2007, 127, 1–8.
- 37 D. Doron, D. T. Major, A. Kohen, W. Thiel and X. Wu, J. Chem. Theory Comput., 2011, 7, 3420–3437.
- 38 B. G. Harvey, M. E. Wright and R. L. Quintana, *Energy & Fuels*, 2010, **24**, 267–273.
- 39 M. P. Kozina, L. V. Bychikhina and G. L. Gal'chenko, Russ. J. Phys. Chem. (Engl. Transl.), 1977, 51, 1258–1259.
- 40 W. V Steele, R. D. Chirico, A. B. Cowell, S. E. Knipmeyer and A. Nguyen, J. Chem. Eng. Data, 2002, 47, 667–688.
- 41 D. J. G. Speight, *Lange's Handb. Chem. Seventeenth Ed.*, 2017.
- 42 M. Sućeska, in Proc. of 32nd Int. Annual Conference of ICT, Karlsruhe, 2001, p. 110/1-110/13.

- 43 M. Sućeska, in Proc. of 30th Int. Annual Conference of ICT, Karlsruhe, 1999, p. 50/1-50/14.
- 44 J. Yanowitz, M. A. Ratcliff, R. L. Mccormick, J. D. Taylor and M. J. Murphy, NREL Compendium of Experimental Cetane Numbers Compendium of Experimental Cetane Numbers, 2017.
- 45 W. Fan, M. Jia, Y. Chang and M. Xie, *Energy & Fuels*, 2015, **29**, 3413–3427.
- 46 W. Chen, S. Shuai and J. Wang, *Energy Fuels*, 2010, **24**, 856–862.
- 47 S. W. Benson, *Thermochemical Kinetics, 2nd ed.* (*Wiley, New York*), 1966.
- 48 K. J. Laidler, Theories of Chemical Reaction Rates, McGraw-Hill, New York, 1969, .
- 49 H. S. Johnston, Gas phase reaction rate theory. The Ronald Press Company, New York 1966, .
- 50 H. Abou-Rachida, K. El Marrounia and S. Kaliaguinea, J. Mol. Struct., 2003, 631, 241–250.
- 51 C. C. Pye and T. Ziegler, *Theor. Chem. Acc.*, 1999, **101**, 396–408.
- 52 A. Klamt, From Quantum Chemistry to Fluid Phase Thermodynamics and Drug Design, 2005.
- 53 A. Klamt, V. Jonas, T. Bu and J. C. W. Lohrenz, 1998, **5639**, 5074–5085.
- 54 A. Klamt, J. Phys. Chem., 1995, **99**, 2224–2235.
- 55 C. C. Pye, T. Ziegler, E. van Lenthe and J. N. Louwen, *Can. J. Chem.*, 2009, **87**, 790–797.
- 56 V. Štejfa, M. Fulem, K. Růžička, C. Červinka, M. A. A. Rocha, L. M. N. B. F. Santos and B. Schröder, J. Chem. Thermodyn., 2013, 60, 117–125.
- 57 V. Štejfa, M. Fulem, K. Růžička and C. Červinka, J. Chem. Thermodyn., 2014, **79**, 280–289.
- 58 I. Fichan, C. Larroche and J. B. Gros, J. Chem. Eng. Data, 1999, 44, 56–62.
- J. Li, E. M. Perdue, S. G. Pavlostathis and R. Araujo, *Environ. Int.*, 1998, **24**, 353–358.
- 60 J. Dykyj and M. Repáš, Tlak nas ýtenej pary organick ých zlúcenín (Saturation Vapor Pressure of Organic Compounds), 1979.
- 61 C. Bicchi, C. Cordero, E. Liberto, P. Rubiolo and B. Sgorbini, J. Chromatogr. A, 2004, 1024, 217–226.
- 62 V. Štejfa, M. Fulem, K. Růžička and C. Červinka, J. Chem. Thermodyn., 2014, **79**, 272–279.
- D. T. Tingey, M. Manning, L. C. Grothaus and W. F. Burns, *Plant Physiol.*, 1980, **65**, 797–801.
- 64 D. R. Stull, Ind. Eng. Chem., 1947, **39**, 517–540.
- 65 V. Štejfa, F. Dergal, I. Mokbel, M. Fulem, J. Jose and K. Růžička, Fluid Phase Equilib., 2015, 406, 124–133.
- 66 B. Arbuzov and B. M. Mikhailov, J. Prakt. Chem.
- 67 Aldrich Chemical Company Inc., Catalog Handbook of Fine Chemicals, Aldrich Chemical Company, Inc., Milwaukee WI, 1990.
- 68 R. . Weast and J. G. Grasselli, CRC Handbook of Data on Organic Compounds, 2nd Editon, CRC Press Inc., Boca Raton, Florida, 1989.
- 69 M. Lecat, C. R. Hebd. Seances Acad. Sci., 1943, 217, 273.
- 70 M. Lecat, Ann. Soc. Sci. Bruxelles, Ser. B, 1926, 45, 284–94.
- 71 G. Chiurdoglu, J. Decot and M. Van Lancker-Francotte, Bull. des Sociétés Chim. Belges, 1954, 63, 70–81.