

# **Improving the Interpretation of Small Molecule Diffusion Coefficients**

Robert Evans<sup>\*a</sup>, Guilherme Dal Poggetto<sup>b</sup>, Mathias Nilsson<sup>b</sup> and Gareth A. Morris<sup>b</sup>

a. Aston Institute of Materials Research, School of Engineering and Applied Science, Aston University, Birmingham, B4 7ET, UK.

b. School of Chemistry, University of Manchester, Manchester M13 9PL, UK.

## **Electronic Supporting Information**

### **Contents**

#### **SI.1 Software**

##### **SI.1.1 Matlab**

##### **SI.1.2 Excel**

#### **SI.2 Diffusion Data Resources**

#### **SI.3 Comparison of Viscosities Used**

#### **SI.4 Arrhenius Plots for All Solvents**

##### **SI.4.1 Deuteriated Solvents**

##### **SI.4.2 Protiated Solvents**

#### **SI.5 Summary of Outlying Diffusion Data**

#### **SI.6 Histogram Summary of Diffusion Data**

#### **SI.7 Subsets of Literature Diffusion Data, Organised by Solvent Used**

#### **SI.8 Subset of Literature Diffusion Data, from Reference 37**

#### **SI.9 Revised Approach, Fitting Solvent Hydrodynamic Radius**

#### **SI.10 Summary of Additional Material found at doi:10.17632/cxt99xf2d2.2**

## SI.1. Software

To simplify the D/MW estimation calculation, minimize the possibility for errors and make the methodology more accessible, the extended model has been implemented both as a Matlab package and as an Excel spreadsheet; both are available for free download from <http://nmr.chemistry.manchester.ac.uk/>.

### SI.1.1 Matlab

The Stokes-Einstein-Gierer-Wirtz Estimation (SEGWE) model has been implemented in MATLAB, and a General User Interface (GUI) was design for user-friendly interaction. Figure SI.2.1 is a screenshot of the “D/MW Calculator” GUI. Both estimations of expected diffusion coefficient from molecular weight and molecular weight from experimental diffusion coefficient are possible, by choosing different “Estimation Modes”. It is also possible to choose from either SEGWE or the classic “Stokes-Einstein” estimation in the “Method” popup list. The GUI can handle a large range (34) of deuterated and protiated solvents – acetone- $d_6$ , acetonitrile, acetonitrile- $d_3$ , benzene, benzene- $d_6$ , chloroform, chloroform- $d_1$ , cyclohexene, cyclohexene- $d_{12}$ , cyclopentane, dioxane, dioxane- $d_8$ , DMF, DMF- $d_7$ , DMSO, DMSO- $d_6$ , n-dodecane, ethanol, ethanol- $d_6$ , fluorobenzene, n-hexane, methanol methanol- $d_3$ , methanol- $d_4$ , nitromethane- $d_3$ , pentanol, pyridine- $d_5$ , n-tetradecane, THF, THF- $d_8$ , toluene, toluene- $d_8$ , water- $d_2$ , and water. The viscosity is estimated accordingly to the chosen solvent and temperature, and then used within any of the various options. Instructions for the addition of solvent, based on knowledge of molecular weight, bulk density and Arrhenius parameters  $a$  and  $b$ , are found within the GUI code. The GUI is also available in its compiled version for either Windows or Linux.

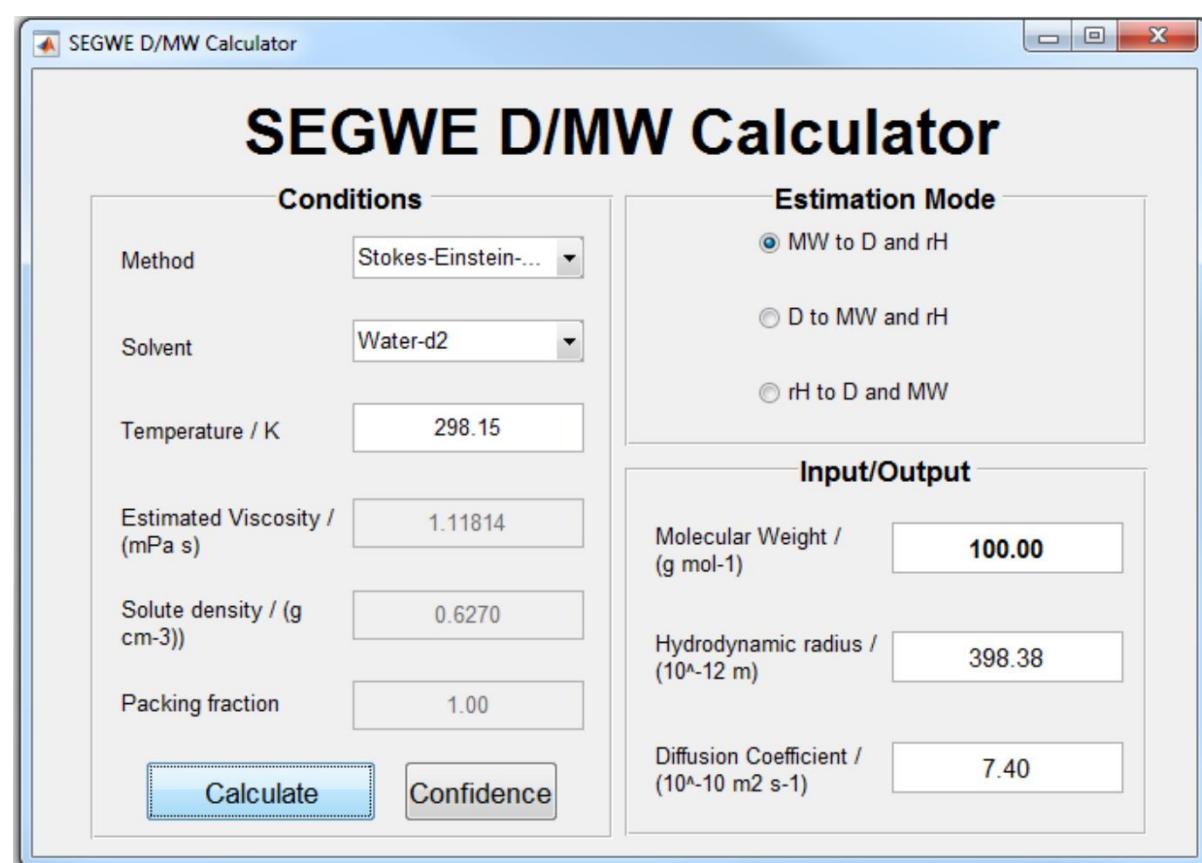


Figure SI.2.1 : Screenshot of D/MW Calculator for Matlab

### SI.1.2 Excel

The SEGWE model has been implemented in Excel. Figure SI.2.2 is an annotated screenshot of the Excel sheet. Both estimations of expected diffusion coefficient from molecular weight and molecular weight from experimental diffusion coefficient are possible. The spreadsheet can handle nine deuterated solvents – water-d<sub>2</sub>, methanol-d<sub>4</sub>, dimethylsulfoxide-d<sub>6</sub>, toluene-d<sub>8</sub>, chloroform-d, tetrahydrofuran-d<sub>8</sub>, benzene-d<sub>6</sub>, acetonitrile-d<sub>3</sub> and cyclohexane-d<sub>12</sub>. Additional solvents can be added manually, provided molecular weight, bulk density and Arrhenius parameters *a* and *b*, are known.

Diffusion Coefficient Calculator RE 2016		
Sample:	Test Sample	
Type of Calculation	M --> D	Top section predicts diffusion coefficient from molecular weight
Solvent:	methanol-d4	Set experimental conditions - solvent from dropdown menu (can add own data)
Experimental T	298	- sample temperature
MW:	46	- enter expected molecular weight
g mol-1		
Diffusion Coefficient:	2.22E-09	
m <sup>2</sup> s <sup>-1</sup>		
Type of Calculation	D --> M	Bottom section predicts diffusion coefficient from molecular weight
Solvent:	methanol-d4	Set experimental conditions - solvent from dropdown menu (can add own data)
Experimental T	298	- sample temperature
Diffusion Coefficient	2.22E-09	- enter experimental diffusion coefficient
m <sup>2</sup> s <sup>-1</sup>		
Predicted MW	46.0	
g mol-1		

Figure SI.2.1 : Annotated screenshot of D/MW Calculator for Excel

## SI.2. Diffusion Data Resources

Ref #	DOI	# values	solvents used
[3]	10.1002/anie.201207403	108‡	water-d2, chloroform-d <sub>1</sub> , dimethyl sulfoxide-d <sub>6</sub> , toluene-d <sub>8</sub> , methanol-d <sub>4</sub>
[14]	10.1039/C6PY00691D	3	benzene-d <sub>6</sub> tetrahydrofuran -d <sub>8</sub> toluene-d8
[18]	10.1021/ja810154u	4	toluene-d8
[21]	10.1021/acs.joc.5b01457	5	chloroform-d <sub>1</sub>
[22]	10.1039/c5sc00670h	32	toluene, tetrahydrofuran
[23]	10.1002/chem.201601145	122	dimethyl sulfoxide-d <sub>6</sub> , cyclohexane-d <sub>12</sub> , benzene-d <sub>6</sub> , chloroform-d <sub>1</sub>
[25]	10.1039/c4sc00848k	3	acetonitrile-d <sub>3</sub>
[26]	10.1039/c4cc02147a	1	water-d <sub>2</sub>
[28]	10.1039/c6sc01974a	1	benzene-d <sub>6</sub>
[37]	10.1016/j.jmr.2006.12.004	200*	chloroform-d <sub>1</sub> , water-d <sub>2</sub>
[38]	10.1006/jmre.2002.2536	4	water-d <sub>2</sub>
[39]	10.1006/jmre.1997.1332	3	methanol-d <sub>4</sub>
[40]	10.1021/ac049174f	3	water-d <sub>2</sub>
[41]	10.1002/anie.200300637	15*‡	water-d <sub>2</sub> , methanol-d <sub>4</sub> , dimethyl sulfoxide-d <sub>6</sub>
[42]	10.1039/c3cc46228e	2	water-d <sub>2</sub>
[43]	10.1039/b415099f	3	water-d <sub>2</sub>
[44]	10.1002/cmr.10026	8*	water-d <sub>2</sub> , toluene
[45]	10.1021/ed200471k	10*‡	water-d <sub>2</sub>
[46]	10.1021/jo961315t	5	benzene
[47]	10.1039/b005319h	35*‡	water, cyclohexane, dioxane, dimethyl sulfoxide, pentanol,

[48]	10.1039/c4ob01373e	14*‡	toluene-d <sub>8</sub> , dimethyl sulfoxide-d <sub>6</sub>
[49]	10.1016/0022-2364(91)90252-O	13†	benzene, water, water-d <sub>2</sub> , acetonitrile, n-hexane, cyclopentane, methanol, fluorobenzene, N, N – dimethyl formamide, cyclohexane, ethanol, dimethyl sulfoxide, tetradecane
[50]	10.1039/A800382C	13	Tetrahydrofuran-d <sub>8</sub>
[51]	10.1016/j.jelechem.2012.06.013	59†	acetonitrile, dimethyl sulfoxide, N, N – dimethyl formamide,
<b>Total</b>		666	

**Table SI.3 Summary of literature diffusion coefficients sources. \* indicates variable temperature study. † indicates reference includes data acquired by a method other than diffusion NMR. ‡ have data points removed, as described in experimental section.**

### SI.3. Comparison of Viscosities Used

solvent	Viscosity at 298 K using Arrhenius behaviour / kg m <sup>-1</sup> s <sup>-1</sup>	Viscosity at 298 K ref 3 / kg m <sup>-1</sup> s <sup>-1</sup>
D <sub>2</sub> O	1.12 × 10 <sup>-3</sup>	1.10 × 10 <sup>-3</sup>
CDCl <sub>3</sub>	5.43 × 10 <sup>-4</sup>	5.44 × 10 <sup>-4</sup>
DMSO-d <sub>6</sub>	2.18 × 10 <sup>-3</sup>	2.15 × 10 <sup>-3</sup>
Methanol-d <sub>4</sub>	6.06 × 10 <sup>-4</sup>	6.12 × 10 <sup>-4</sup>
Toluene-d <sub>8</sub>	6.01 × 10 <sup>-4</sup>	6.02 × 10 <sup>-4</sup>

**Table SI.4. Comparison of viscosities used in this work with those used in original work, Reference 3.**

### SI.4. Arrhenius Data for All Solvents

Solvent	MW / g mol <sup>-1</sup>	<i>a</i> /kg m <sup>-1</sup> s <sup>-1</sup>	<i>b</i> / K	# samples
<b>Protiated</b>				
acetonitrile	41.05	1.78 × 10 <sup>-5</sup>	899	30
benzene	78.11	1.78 × 10 <sup>-5</sup>	1252	1
chloroform	119.37	9.06 × 10 <sup>-6</sup>	878	0
cyclohexane	84.16	6.00 × 10 <sup>-6</sup>	1492	6
cyclopentane	70.10	6.17 × 10 <sup>-5</sup>	917	1
N, N – dimethyl formamide	73.09	1.93 × 10 <sup>-5</sup>	1114	11
dimethyl sulfoxide	78.13	8.35 × 10 <sup>-6</sup>	1631	25
dioxane	88.11	9.22 × 10 <sup>-6</sup>	1450	5
dodecane	170.33	1.02 × 10 <sup>-5</sup>	1458	6
ethanol	46.07	4.75 × 10 <sup>-6</sup>	1790	1
fluorobenzene	96.10	1.73 × 10 <sup>-5</sup>	1051	1
<i>n</i> -hexane	86.18	1.49 × 10 <sup>-5</sup>	890	1
methanol	32.04	9.57 × 10 <sup>-6</sup>	1203	1
pentanol	88.15	8.84 × 10 <sup>-7</sup>	2458	6
tetradecane	198.39	5.99 × 10 <sup>-6</sup>	1745	1
tetrahydrofuran	72.11	3.05 × 10 <sup>-5</sup>	812	0
toluene	92.14	1.55 × 10 <sup>-5</sup>	1068	0
water	18.02	8.12 × 10 <sup>-7</sup>	2093	10
<b>TOTAL protiated</b>				106

<b>Deuteriated</b>				
acetonitrile-d <sub>3</sub>	44.07	$1.78 \times 10^{-5}$	899	<b>3</b>
benzene-d <sub>6</sub>	84.15	$9.06 \times 10^{-6}$	1256	<b>39</b>
chloroform-d	120.38	$2.86 \times 10^{-5}$	878	<b>213</b>
cyclohexane-d <sub>12</sub>	96.23	$6.17 \times 10^{-6}$	1504	<b>30</b>
dimethyl sulfoxide-d <sub>6</sub>	84.17	$1.93 \times 10^{-5}$	1742	<b>60</b>
methanol-d <sub>4</sub>	36.07	$8.62 \times 10^{-6}$	1268	<b>24</b>
tetrahydrofuran-d <sub>8</sub>	80.16	$2.07 \times 10^{-5}$	930	<b>30</b>
toluene-d <sub>8</sub>	100.19	$1.50 \times 10^{-5}$	1217	<b>54</b>
water-d <sub>2</sub>	20.03	$4.52 \times 10^{-7}$	2330	<b>107</b>
<b>TOTAL</b> deuteriated				<b>560</b>
<b>TOTAL</b>				<b>666</b>

**Table SI.5: Molecular masses and Arrhenius viscosity parameters for all solvents used in this study. Parameters for chloroform-d are estimated from protiated solvent data.**

#### SI.4.1. Deuteriated

Diffusion measurements for compounds in nine different deuteriated solvents were obtained. Viscosity data for eight of the solvents is summarised here. The case of chloroform-d is discussed in the main text. Arrhenius viscosity data for a further four deuteriated solvents – dioxane-d<sub>8</sub>, dmso-d<sub>7</sub>, ethanol-d<sub>6</sub>, nitromethane-d<sub>3</sub> and pyridine-d<sub>5</sub> – are also available, but there are no reports of their use, so far, in diffusion NMR experiments.

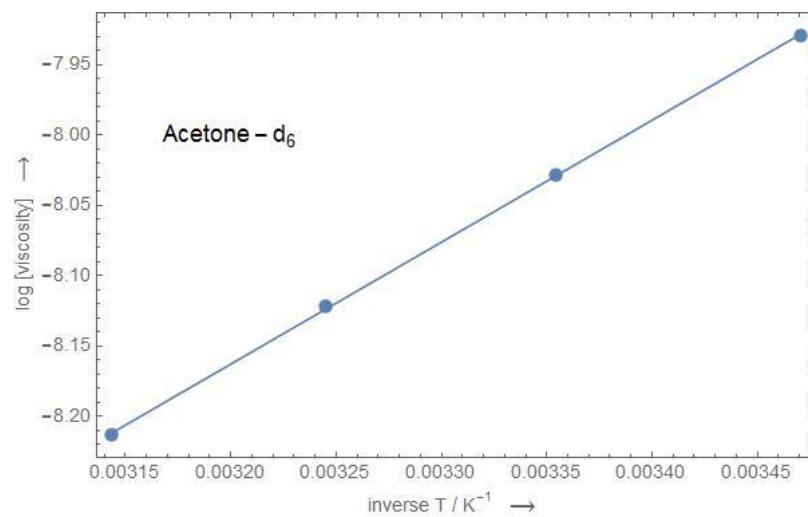
In every case, the data is fitted to an Arrhenius-like equation, with two variable parameters *a* and *b* produced for each solvent.

$$\eta = a e^{\frac{b}{T}} \quad (6)$$

acetonitrile-d <sub>3</sub>
benzene-d <sub>6</sub>
cyclohexane-d <sub>12</sub>
dimethyl sulfoxide-d <sub>6</sub>
methanol-d <sub>4</sub>
tetrahydrofuran-d <sub>8</sub>
toluene-d <sub>8</sub>
water-d <sub>2</sub>

#### SI.4.1.1: Acetone-d<sub>6</sub>

Source: *J. Chem. Phys.* 104, 669-679 (1996)

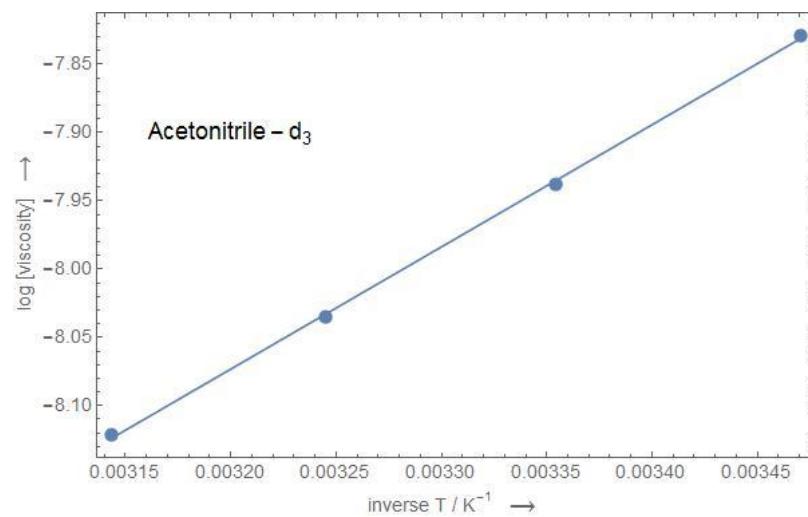


$$a = 1.782 \times 10^{-5} \text{ kg m}^{-1} \text{ s}^{-1}; b = 866.334 \text{ K}$$

Figure SI.5.1.1: Plot of natural logarithm of viscosity against reciprocal temperature for acetone-d<sub>6</sub>.

#### SI.4.1.2: Acetonitrile-d<sub>3</sub>

Source: *J. Chem. Phys.* 104, 669-679 (1996)

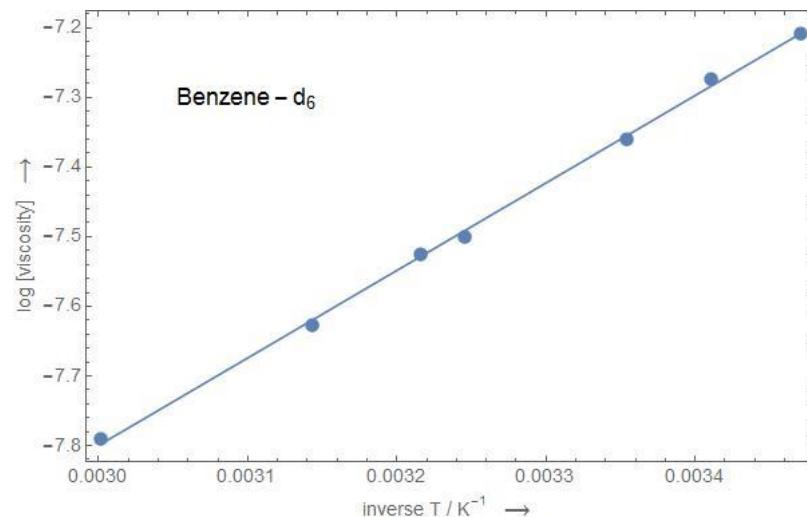


$$a = 1.780 \times 10^{-5} \text{ kg m}^{-1} \text{ s}^{-1}; b = 894.699 \text{ K}$$

Figure SI.5.1.2: Plot of natural logarithm of viscosity against reciprocal temperature for acetonitrile-d<sub>3</sub>.

#### SI.4.1.3: Benzene-d<sub>6</sub>

Source: *J. Chem. Phys.* 104, 669-679 (1996), *J. Phys. Chem.* 58, 430-432 (1954)

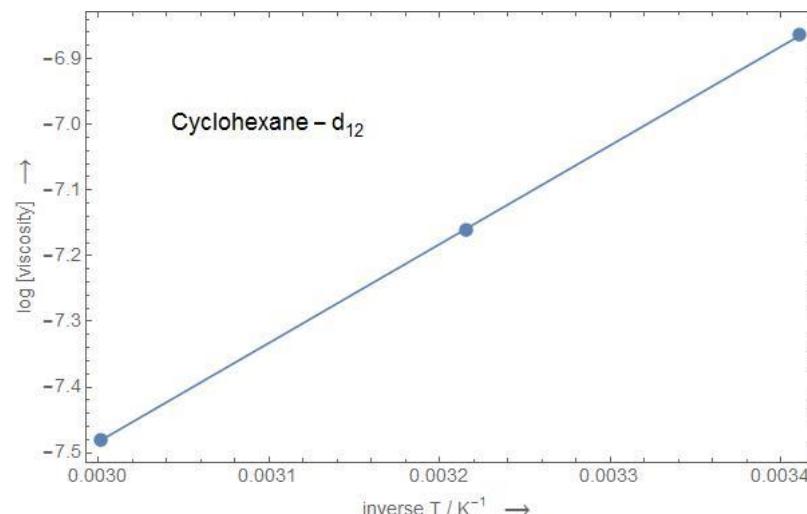


$$a = 9.456 \times 10^{-6} \text{ kg m}^{-1} \text{ s}^{-1}; b = 1256.41 \text{ K}$$

Figure SI.5.1.3: Plot of natural logarithm of viscosity against reciprocal temperature for benzene-d<sub>6</sub>.

#### SI.4.1.4: Cyclohexane-d<sub>12</sub>

Source: *J. Phys. Chem.* 58, 430-432 (1954)

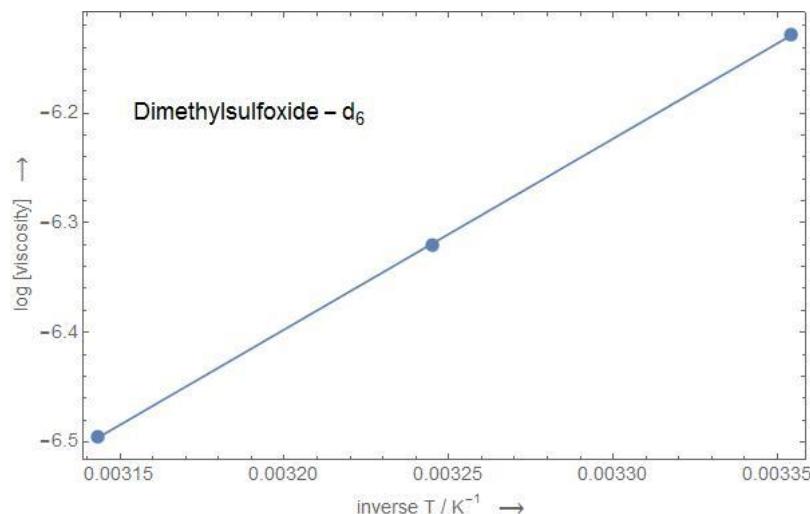


$$a = 6.165 \times 10^{-6} \text{ kg m}^{-1} \text{ s}^{-1}; b = 1504.43 \text{ K}$$

Figure SI.5.1.4: Plot of natural logarithm of viscosity against reciprocal temperature for cyclohexane-d<sub>12</sub>.

#### SI.4.1.5: Dimethylsulfoxide-d<sub>6</sub>

Source: *J. Chem. Phys.* 104, 669-679 (1996)

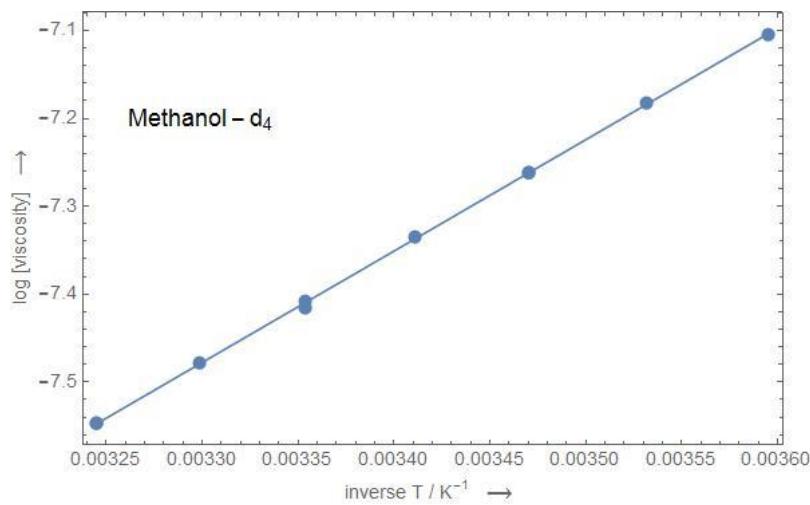


$$a = 6.324 \times 10^{-6} \text{ kg m}^{-1} \text{ s}^{-1}; b = 1741.84 \text{ K}$$

Figure SI.5.1.5: Plot of natural logarithm of viscosity against reciprocal temperature for dimethylsulfoxide-d<sub>6</sub>.

#### SI.4.1.6: Methanol-d<sub>4</sub>

Source: *J. Chem. Phys.* 104, 669-679 (1996), *J. Chem. Phys.* 91, 2568-2574 (1989)

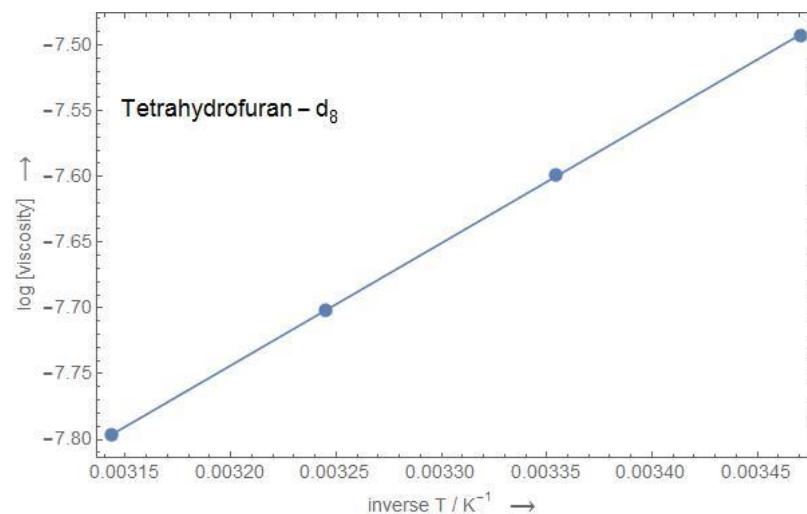


$$a = 8.623 \times 10^{-6} \text{ kg m}^{-1} \text{ s}^{-1}; b = 1267.68 \text{ K}$$

Figure SI.5.1.6: Plot of natural logarithm of viscosity against reciprocal temperature for methanol-d<sub>4</sub>.

#### SI.4.1.7: Tetrahydrofuran-d<sub>8</sub>

Source: *J. Chem. Phys.* 104, 669-679 (1996)

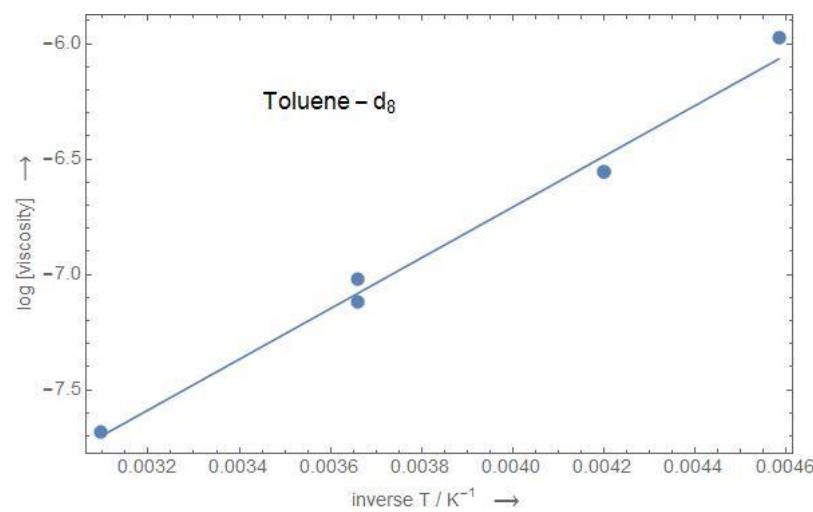


$$a = 2.208 \times 10^{-5} \text{ kg m}^{-1} \text{ s}^{-1}; b = 930.4 \text{ K}$$

Figure SI.5.1.7: Plot of natural logarithm of viscosity against reciprocal temperature for tetrahydrofuran-d<sub>8</sub>.

#### SI.4.1.8: Toluene-d<sub>8</sub>

Source: *J. Chem. Phys.* 82, 3360-3370 (1985), *J. Chem. Phys.* 62, 2800 (1975)

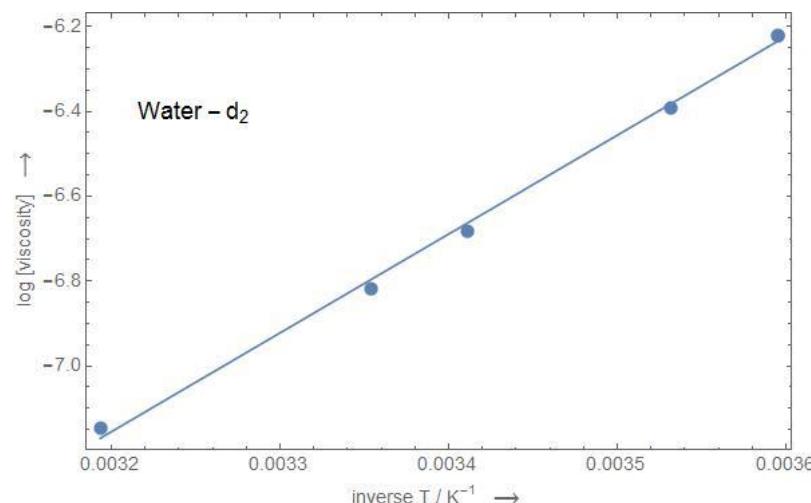


$$a = 1.504 \times 10^{-5} \text{ kg m}^{-1} \text{ s}^{-1}; b = 1099.36 \text{ K}$$

Figure SI.5.1.8: Plot of natural logarithm of viscosity against reciprocal temperature for toluene-d<sub>8</sub>.

#### SI.4.1.9: Water-d<sub>2</sub>

**Source:** *J. Chem. Eng. Data.* 49, 1064-1069 (2004), U.S. Department of Commerce, National Bureau of Standards, Research Paper RP1994, 42, June 1949



$$a = 4.521 \times 10^{-7} \text{ kg m}^{-1} \text{ s}^{-1}; b = 2329.54 \text{ K}$$

Figure SI.5.1.9: Plot of natural logarithm of viscosity against reciprocal temperature for water-d<sub>2</sub>.

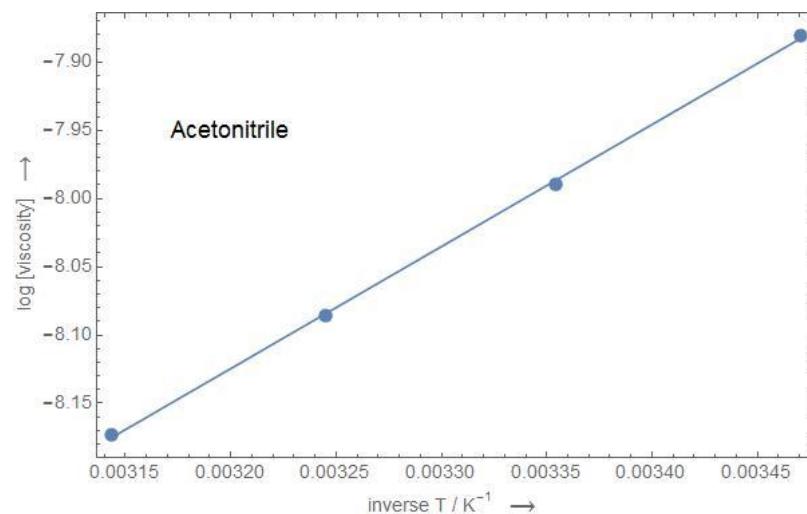
#### SI.4.2. Protiated Solvents

Diffusion measurements for compounds in eighteen different protiated solvents were obtained. Viscosity data for all liquids are summarised here. Two variable parameters *a* and *b* are produced for each solvent.

acetonitrile
benzene
chloroform
cyclohexane
cyclopentane
N, N - dimethyl formamide
dimethyl sulfoxide
dioxane
dodecane
ethanol
fluorobenzene
<i>n</i> -hexane
methanol
pentanol
tetradecane
tetrahydrofuran
toluene
water

#### SI.4.2.1: Acetonitrile

Source: *J. Chem. Phys.* 104, 669-679 (1996)



$$a = 1.780 \times 10^{-5} \text{ kg m}^{-1} \text{ s}^{-1}; b = 894.375 \text{ K}$$

Figure SI.5.2.1: Plot of natural logarithm of viscosity against reciprocal temperature for acetonitrile.

#### SI.4.2.2: Benzene

Source: *J. Chem. Eng. Data* 40, 567-571 (1995), *J. Chem. Eng. Data* 34, 37-43 (1989), Landolt - Bornstein, Transport Phenomena I; Springer-Verlag : Berlin, 1969; Band 11, Teil5, Bandteil a : (a) p 158, (b) p 157, (c) p 195, (d) p 161;

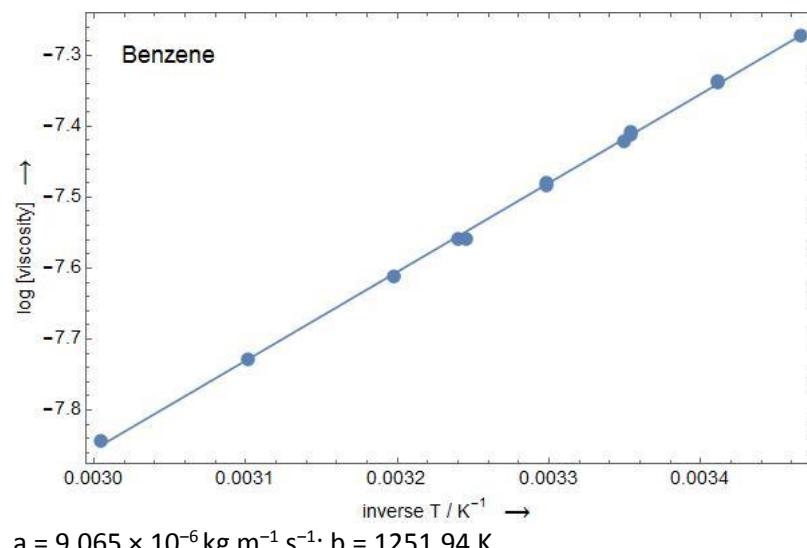


Figure SI.5.2.2: Plot of natural logarithm of viscosity against reciprocal temperature for benzene.

#### SI.4.2.3: Chloroform

Source: *J. Chem. Eng. Data*, 40, 1058-1061 (1995), *J. Chem. Eng. Data*, 40, 567-571 (1995).

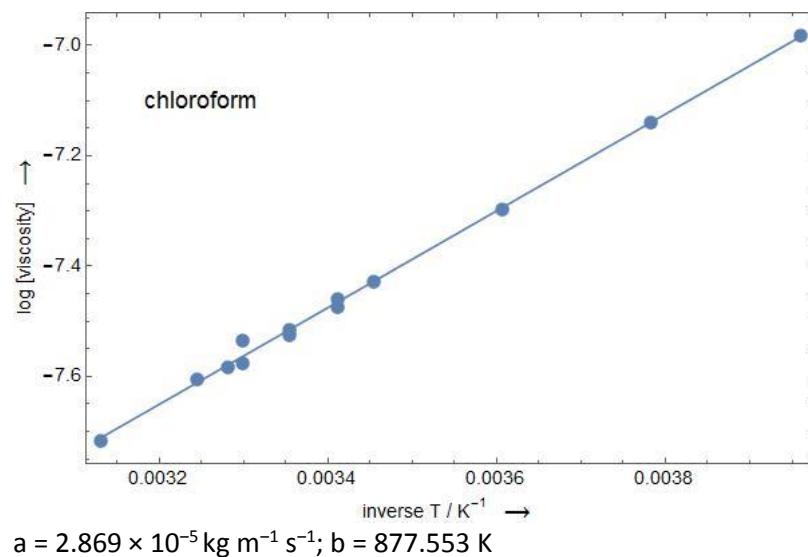
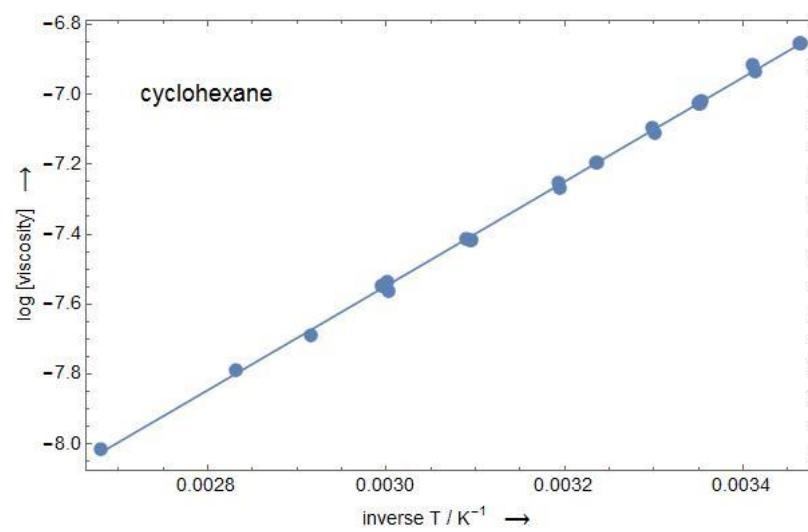


Figure SI.5.2.3: Plot of natural logarithm of viscosity against reciprocal temperature for chloroform.

#### SI.4.2.4: Cyclohexane

Source: *Izv. Vyssh. Uchebn. Zaved. Neft Gaz* 64-100, (1975), *Zh. Fiz. Khim.* 49, 548-549 (1975), *J. Chem. Eng. Data*, 34, 37-43 (1989).

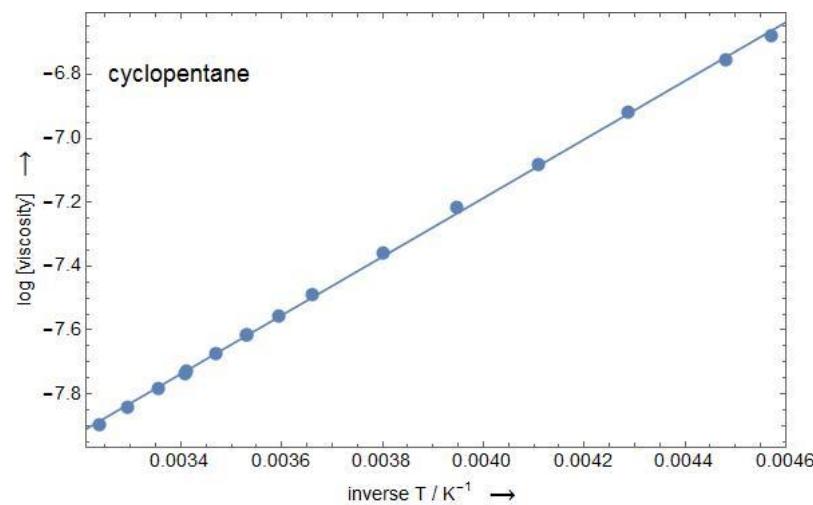


$a = 5.998 \times 10^{-6} \text{ kg m}^{-1} \text{ s}^{-1}$ ;  $b = 1492.1 \text{ K}$

Figure SI.5.2.4: Plot of natural logarithm of viscosity against reciprocal temperature for cyclohexane.

#### SI.4.2.5: Cyclopentane

Source: *High Temp. - High Press.* 32, 179-184 (2000), *J. Chem. Eng. Data*, 49, 138-142 (2004)

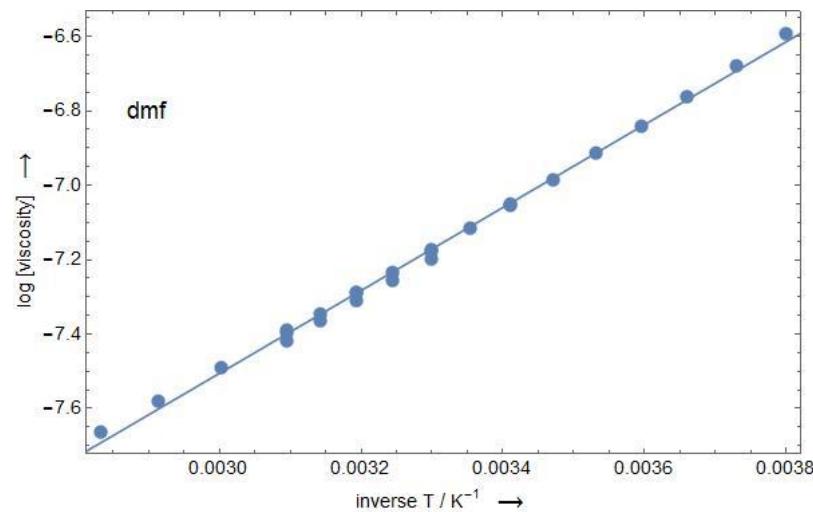


$$a = 1.932 \times 10^{-5} \text{ kg m}^{-1} \text{ s}^{-1}; b = 916.883 \text{ K}$$

Figure SI.5.2.5: Plot of natural logarithm of viscosity against reciprocal temperature for cyclopentane.

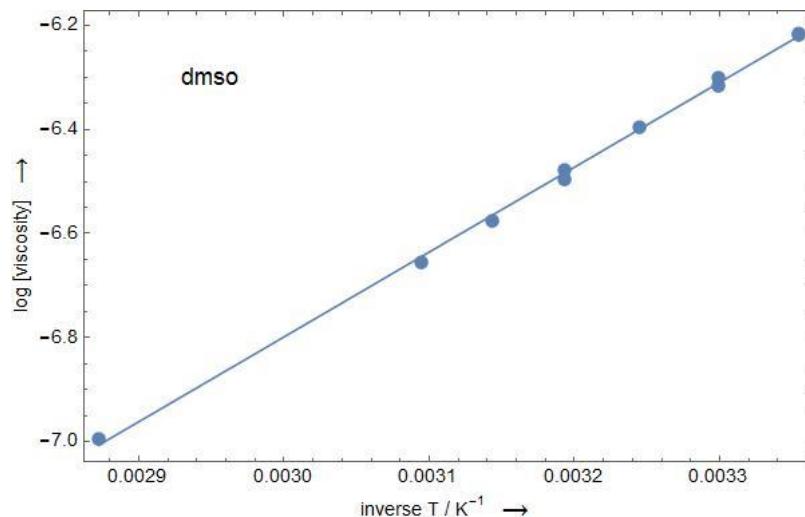
#### SI.4.2.6: N, N – dimethylformamide

Source: *J. Chem. Eng. Data* 53, 1024-1027 (2008), *J. Chem. Eng. Data* 36, 1442-1445 (1991), *Phys. Chem. Liq.* 39, 383-399 (2001).



#### SI.4.2.7: Dimethylsulfoxide

**Source:** *J. Chem. Thermodyn.* 39, 1049-1056 (2007), Gaylord Chemical Company. "Dimethyl Sulfoxide Physical Properties." June 2014. [www.gaylordchemical.com/wp-content/uploads/2015/09/GC-Literature-101B-REV.pdf](http://www.gaylordchemical.com/wp-content/uploads/2015/09/GC-Literature-101B-REV.pdf)

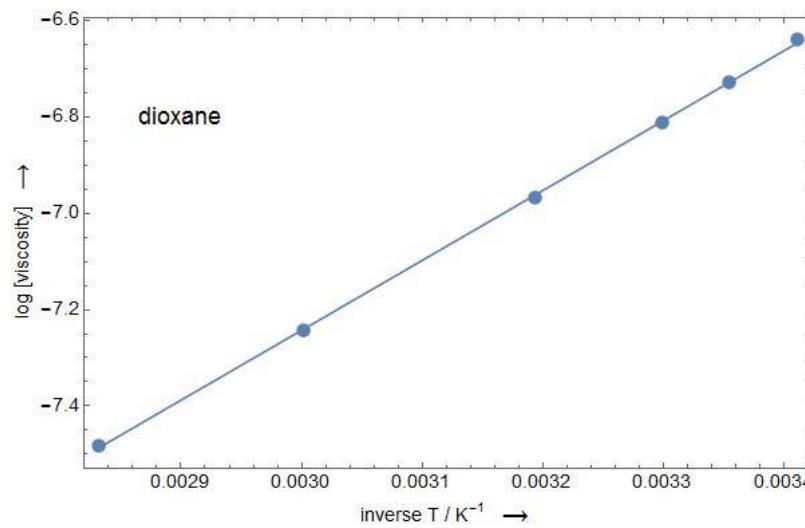


$$a = 8.353 \times 10^{-6} \text{ kg m}^{-1} \text{ s}^{-1}; b = 1631.34 \text{ K}$$

Figure SI.5.2.7: Plot of natural logarithm of viscosity against reciprocal temperature for dimethylsulfoxide.

#### SI.4.2.8: Dioxane

**Source:** *J. Am. Chem. Soc.* 55, 4832-4837 (1933)

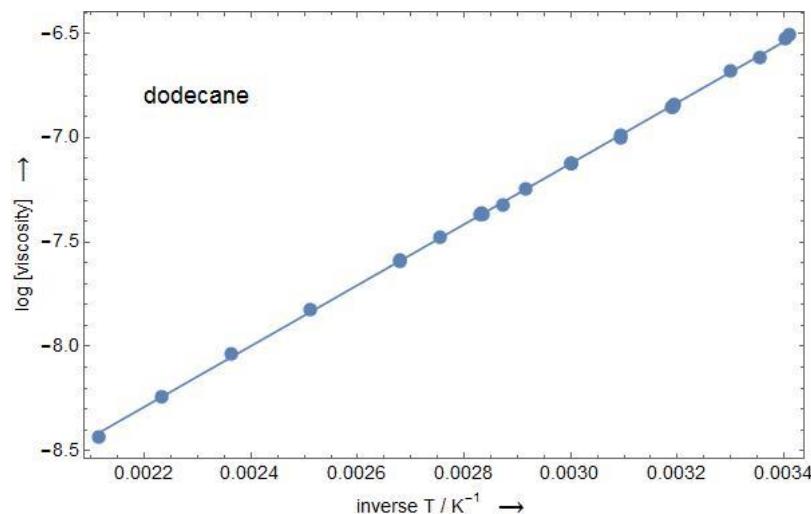


$$a = 9.221 \times 10^{-6} \text{ kg m}^{-1} \text{ s}^{-1}; b = 1450.38 \text{ K}$$

Figure SI.5.2.8: Plot of natural logarithm of viscosity against reciprocal temperature for dioxane.

#### SI.4.2.9: Dodecane

**Source:** *Int. J. Thermophys.* 25, 1339-1352 (2004), *J. Chem. Eng. Data*, 34, 37-43 (1989), 'Measurements of viscosity and density of quantitative n-alkane mixtures (C<sub>6</sub> – C<sub>60</sub>)' H. G. Yucel and A. Uysal, presented at 17th European Conference on Thermophysical Properties, Bratislava, 2005.

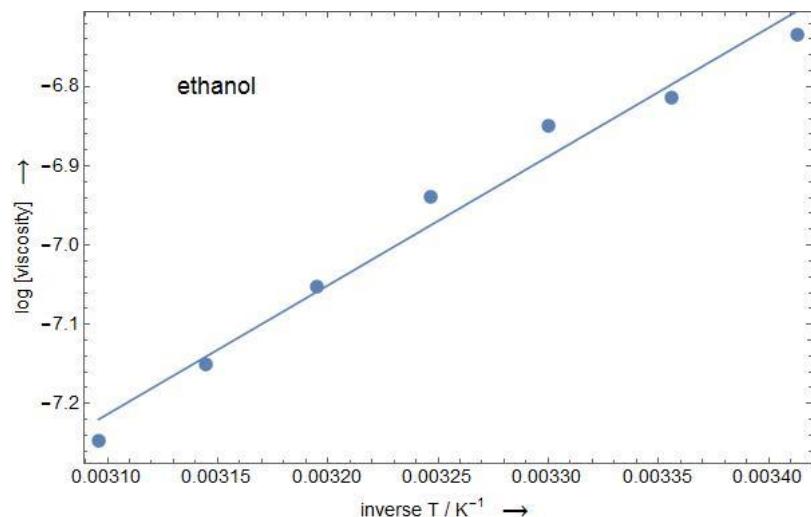


$$a = 1.016 \times 10^{-5} \text{ kg m}^{-1} \text{ s}^{-1}; b = 1458.05 \text{ K}$$

Figure SI.5.2.9: Plot of natural logarithm of viscosity against reciprocal temperature for dodecane.

#### SI.4.2.10: Ethanol

**Source:** *Korean J. Chem. Eng.* 29, 812-817 (2012).

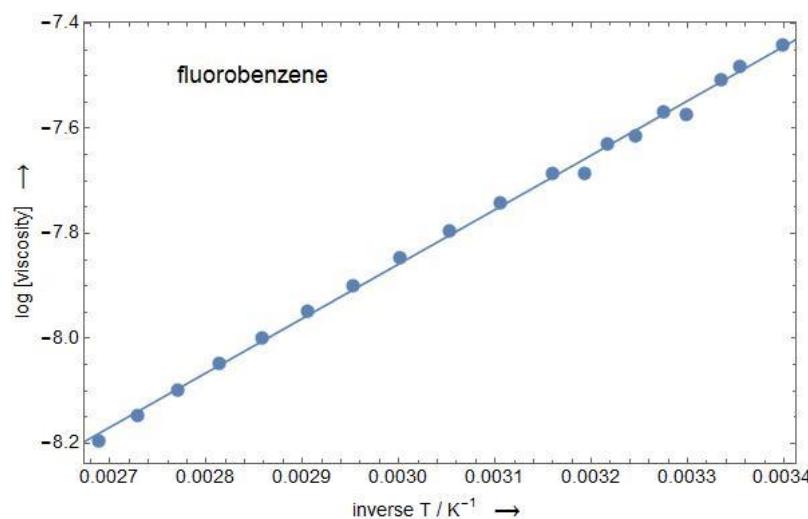


$$a = 4.754 \times 10^{-6} \text{ kg m}^{-1} \text{ s}^{-1}; b = 1626.83 \text{ K}$$

Figure SI.5.2.10: Plot of natural logarithm of viscosity against reciprocal temperature for ethanol.

#### SI.4.2.11: Fluorobenzene

**Source:** *Eur. Chem. Bull.* 2, 434-444 (2013), Cameo Chemicals. "Fluorobenzene." June 1999. <https://cameochemicals.noaa.gov/chris/FLB.pdf>

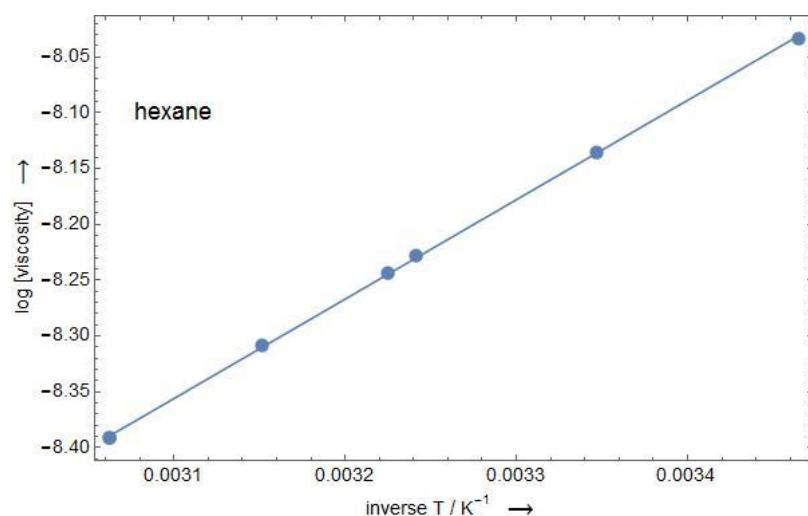


$$a = 1.726 \times 10^{-5} \text{ kg m}^{-1} \text{ s}^{-1}; b = 1036.43 \text{ K}$$

Figure SI.5.2.11: Plot of natural logarithm of viscosity against reciprocal temperature for fluorobenzene.

#### SI.4.2.12: *n*-Hexane

**Source:** *J. Chem. Eng. Data*, 34, 37-43 (1989).

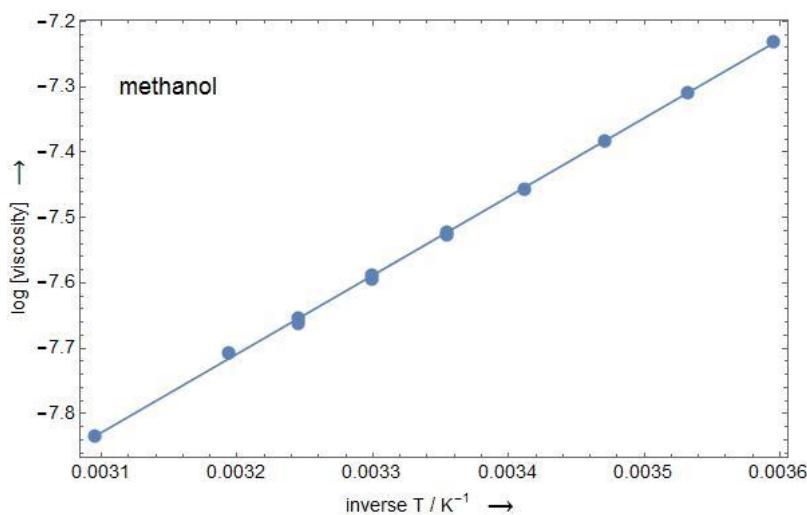


$$a = 1.487 \times 10^{-5} \text{ kg m}^{-1} \text{ s}^{-1}; b = 890.301 \text{ K}$$

Figure SI.5.2.12: Plot of natural logarithm of viscosity against reciprocal temperature for *n*-hexane.

#### SI.4.2.13: Methanol

Source: *J. Chem. Eng. Data* 6, 533-537 (1961), *J. Chem. Phys.* 91, 2568-2574 (1989).

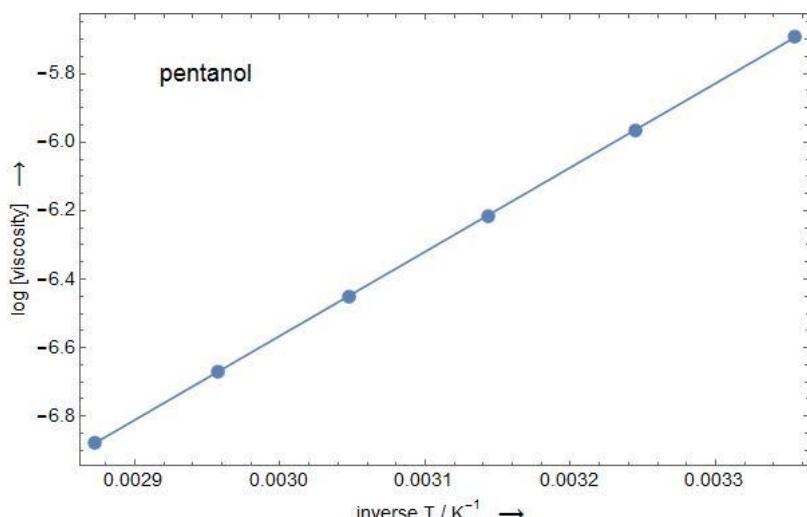


$$a = 9.573 \times 10^{-6} \text{ kg m}^{-1} \text{ s}^{-1}; b = 1202.6 \text{ K}$$

Figure SI.5.2.13: Plot of natural logarithm of viscosity against reciprocal temperature for methanol.

#### SI.4.2.14: Pentanol

Source: *J. Solution Chem.* 38, 779-799 (2009).

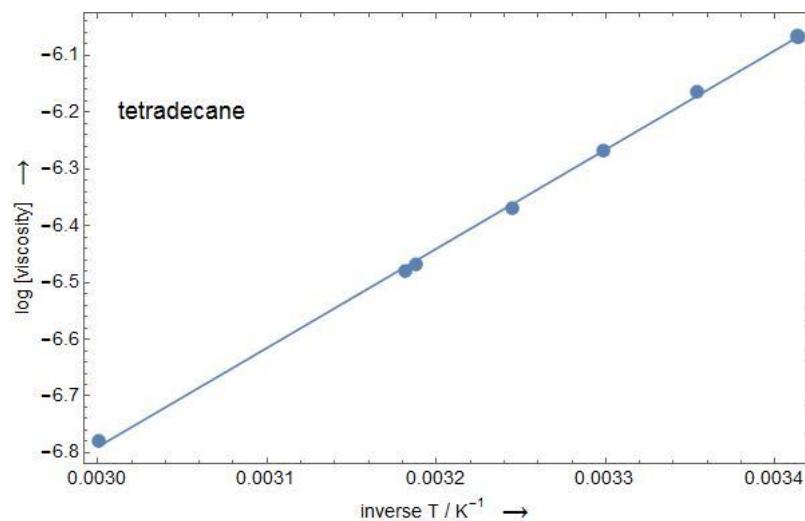


$$a = 8.838 \times 10^{-7} \text{ kg m}^{-1} \text{ s}^{-1}; b = 2457.95 \text{ K}$$

Figure SI.3.2.14: Plot of natural logarithm of viscosity against reciprocal temperature for pentanol.

#### SI.4.2.15: Tetradecane

**Source:** *J. Chem. Thermodyn.* 38, 1062-1071 (2006), *J. Chem. Eng. Data* 34, 37-43 (1989), *J. Phys. Chem. Ref. Data* 23, 41-53 (1994).

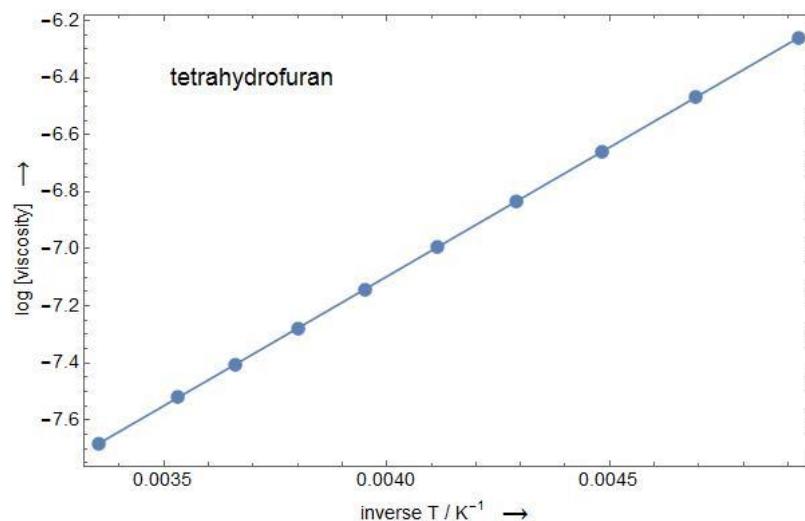


$$a = 5.989 \times 10^{-6} \text{ kg m}^{-1} \text{ s}^{-1}; b = 1745.34 \text{ K}$$

Figure SI.5.2.15: Plot of natural logarithm of viscosity against reciprocal temperature for tetradecane.

#### SI.4.2.16: Tetrahydrofuran

**Source:** *J. Phys. Chem.* 71, 1158 – 1158 (1967), *J. Am. Chem. Soc.* 87, 5548- 5553 (1965).

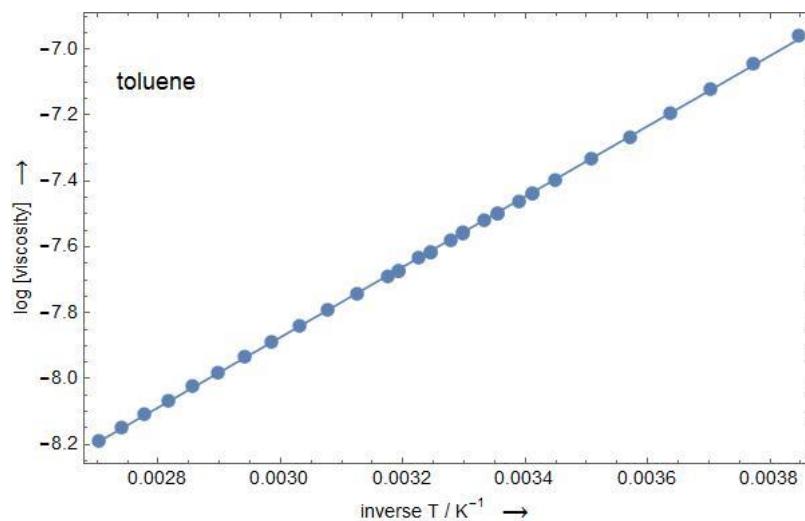


$$a = 3.045 \times 10^{-5} \text{ kg m}^{-1} \text{ s}^{-1}; b = 906.445 \text{ K}$$

Figure SI.5.2.16: Plot of natural logarithm of viscosity against reciprocal temperature for tetrahydrofuran.

#### SI.4.2.17: Toluene

**Source:** *J. Phys. Chem. Ref. Data*, 35, 1-8 (2006), *J. Chem. Eng. Data* 40, 567-571 (1995), *J. Chem. Eng. Data*, 32, 344-348 (1987).

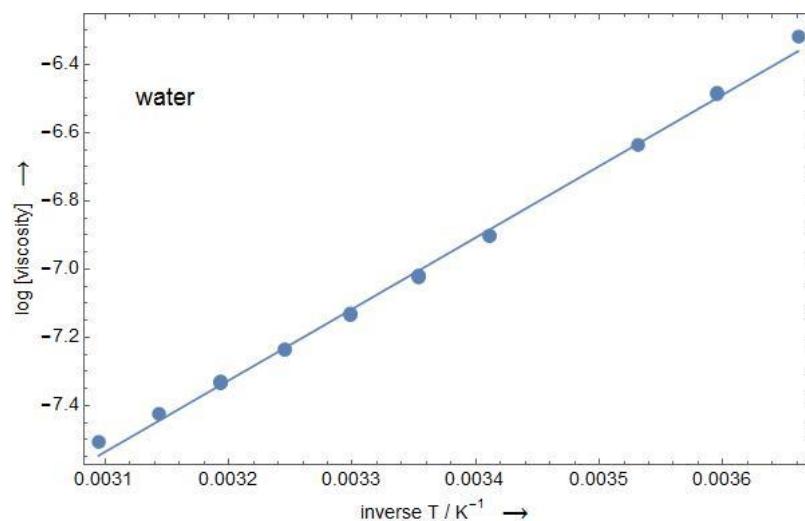


$$a = 1.546 \times 10^{-5} \text{ kg m}^{-1} \text{ s}^{-1}; b = 1067.67 \text{ K}$$

Figure SI.5.2.17: Plot of natural logarithm of viscosity against reciprocal temperature for toluene.

#### SI.4.2.18: Water

**Source:** *J. Chem. Thermodyn.* 39, 1049-1056 (2007), *J. Chem. Eng. Data* 49, 1064-1069 (2004), *J. Am. Chem. Soc.* 55, 4832-4837 (1933), U.S. Department of Commerce, National Bureau of Standards, Research Paper RP1994, 42, June 1949, *J. Chem. Eng. Data* 6, 533-536 (1961).



$$a = 8.121 \times 10^{-7} \text{ kg m}^{-1} \text{ s}^{-1}; b = 2092.95 \text{ K}$$

Figure SI.5.2.18: Plot of natural logarithm of viscosity against reciprocal temperature for water.

## SI.5. Summary of Outlying Data Points

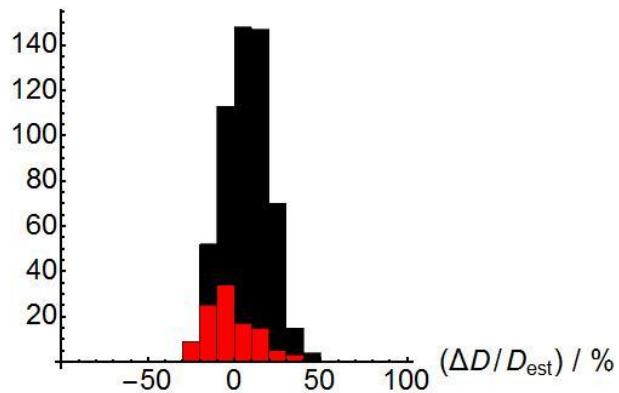
Of the 108 measurements in the original data set and 558 measurements in the literature data set, 43 have % errors greater than 25 %, corresponding to 6.5 % of all data used. These are listed in Table SI.5 in reference order, by solvent, then by largest % error.

Reference	Solute	Solvent	$D_{\text{exp}}$ / $10^{-10} \text{ m}^2 \text{ s}^{-1}$	$D_{\text{est}}$ / $10^{-10} \text{ m}^2 \text{ s}^{-1}$	% error
3	n-butylbenzene	DMSO-d <sub>6</sub>	5.5	4.2	31.3
	1,2-dimethoxybenzene	methanol-d <sub>4</sub>	16.9	12.7	32.6
	guaiacol	methanol-d <sub>4</sub>	17.8	13.4	32.5
	menthol	toluene-d <sub>8</sub>	18.0	14.4	25.0
21	chrysene-d <sub>12</sub>	CDCl <sub>3</sub>	16.8	13.2	27.3
22	anthracene	THF-d <sub>8</sub>	19.6	15.5	26.2
	anthracene	toluene-d <sub>8</sub>	17.3	13.5	28.3
	THF	toluene-d <sub>8</sub>	26.6	20.8	28.1
	9-methylanthracene	toluene-d <sub>8</sub>	16.5	12.9	27.7
23	toluene-d <sub>7</sub>	toluene-d <sub>8</sub>	23.3	18.3	26.8
	pyrene	toluene-d <sub>8</sub>	16.0	12.6	26.6
	indene	toluene-d <sub>8</sub>	21.2	16.9	25.9
	anthracene	C <sub>6</sub> D <sub>12</sub>	11.9	8.4	42.3
23	naphthalene	C <sub>6</sub> D <sub>12</sub>	14.0	9.9	41.0
	pyrene	C <sub>6</sub> D <sub>12</sub>	10.9	7.8	38.4
	2-phenylpyridine	C <sub>6</sub> D <sub>12</sub>	12.4	9.0	37.4
	toluene	C <sub>6</sub> D <sub>12</sub>	16.2	11.9	36.5
23	Si(OMe) <sub>4</sub>	C <sub>6</sub> D <sub>12</sub>	12.2	9.1	34.3
	acridine	C <sub>6</sub> D <sub>12</sub>	10.9	8.3	30.1
	indandione	C <sub>6</sub> D <sub>12</sub>	11.9	9.3	27.8
	9-methylanthracene	C <sub>6</sub> D <sub>12</sub>	10.3	8.1	27.4
23	indene	C <sub>6</sub> D <sub>12</sub>	13.4	10.5	27.4
	toluene	DMSO-d <sub>6</sub>	7.0	5.1	38.6
	tetrahydrofuran	DMSO-d <sub>6</sub>	7.9	5.8	36.1
	diisopropylether	DMSO-d <sub>6</sub>	6.3	4.8	31.8
23	cyclopentane	DMSO-d <sub>6</sub>	7.6	5.9	29.2
	naphthalene	DMSO-d <sub>6</sub>	5.4	4.3	27.7
	Si(OMe) <sub>4</sub>	DMSO-d <sub>6</sub>	5.0	3.9	27.5
	indene	DMSO-d <sub>6</sub>	5.6	4.5	25.3
37	tetradecane	tetradecane	5.6	4.2	34.0
	unknown	CDCl <sub>3</sub>	28.8	22.2	30.1
	unknown	CDCl <sub>3</sub>	26.7	20.6	29.6
	unknown	CDCl <sub>3</sub>	24.7	19.3	27.8
41	unknown	CDCl <sub>3</sub>	29.6	23.3	27.1
	unknown	CDCl <sub>3</sub>	27.4	21.8	25.4
	[18]crown-6	D <sub>2</sub> O	4.1	3.1	33.3
	[18]crown-6	methanol-d <sub>4</sub>	13.5	9.4	44.3
46	butyl acetate	benzene-d <sub>6</sub>	22.7	15.2	48.9
	isopropyl butyrate	benzene-d <sub>6</sub>	19.9	14.4	38.6
	butyl levulinate	benzene-d <sub>6</sub>	16.7	12.4	34.5

	propyl acetate	benzene-d6	21.5	16.3	31.7
51	chrysene	DMSO	4.5	3.4	31.1
	triphenylene	DMSO	4.4	3.4	27.0

## SI.6. Histogram Summary of Diffusion Data

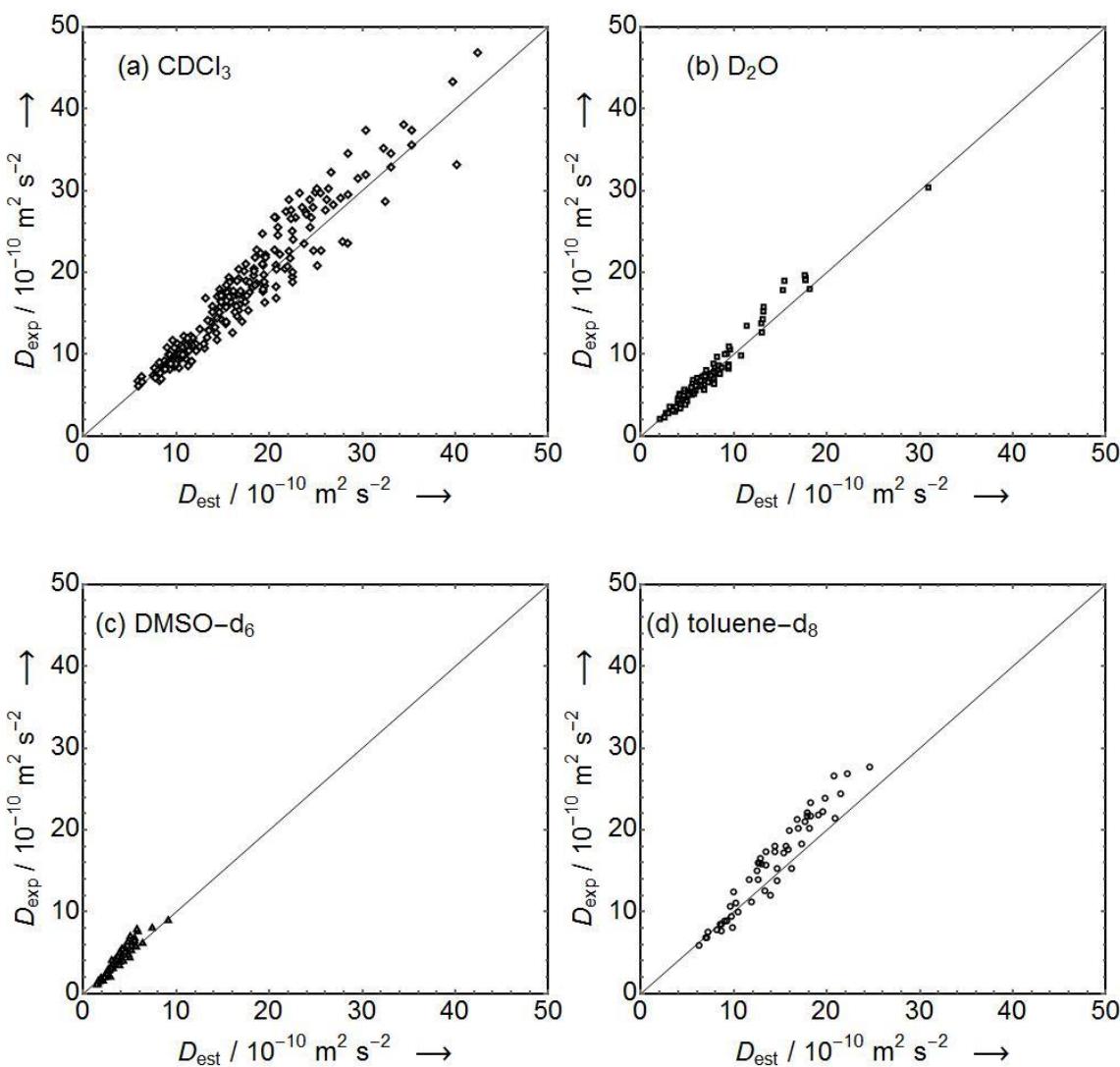
Histogram summarising all of the information in Figs. 1, 2.a and 2.b.

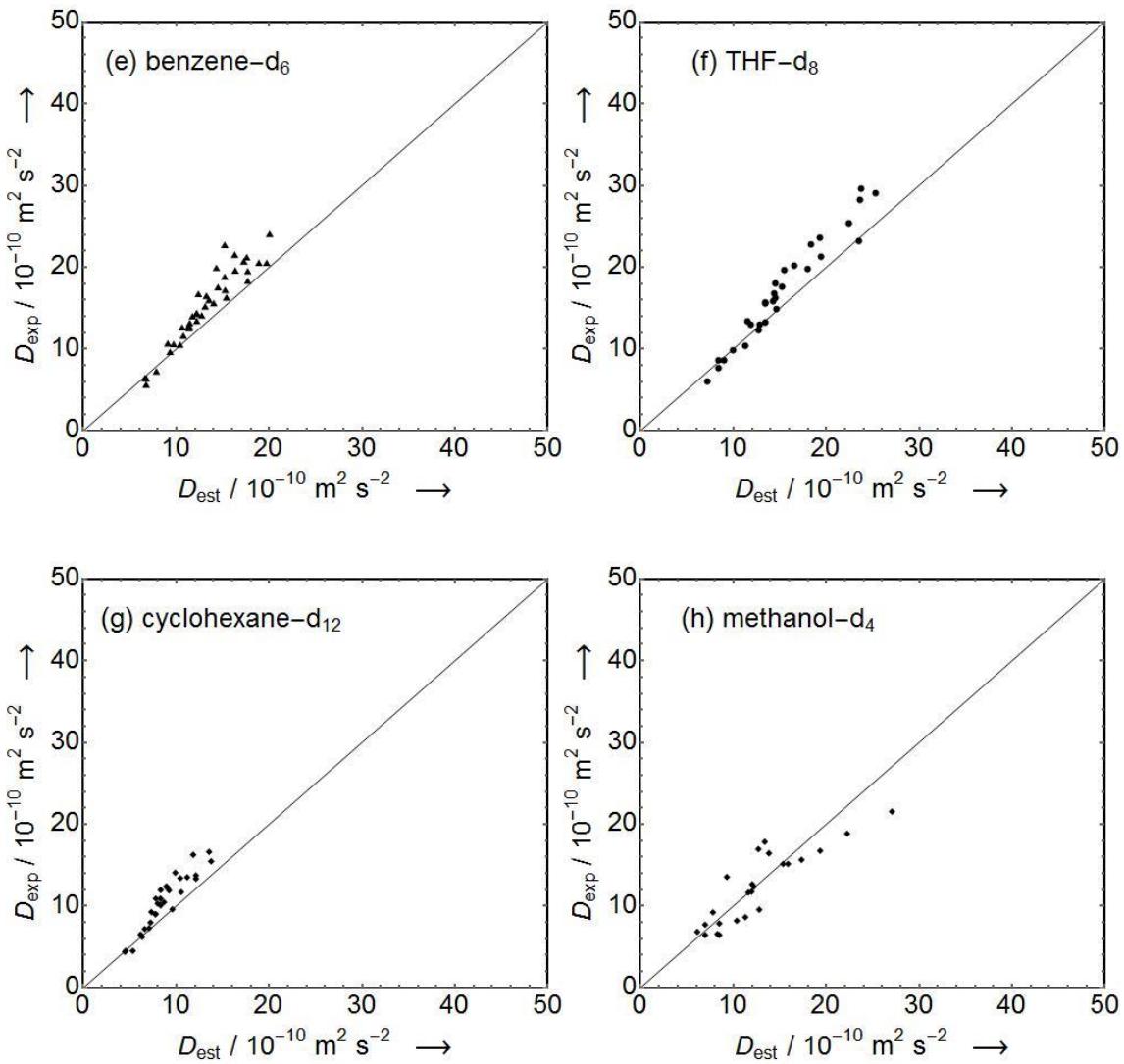


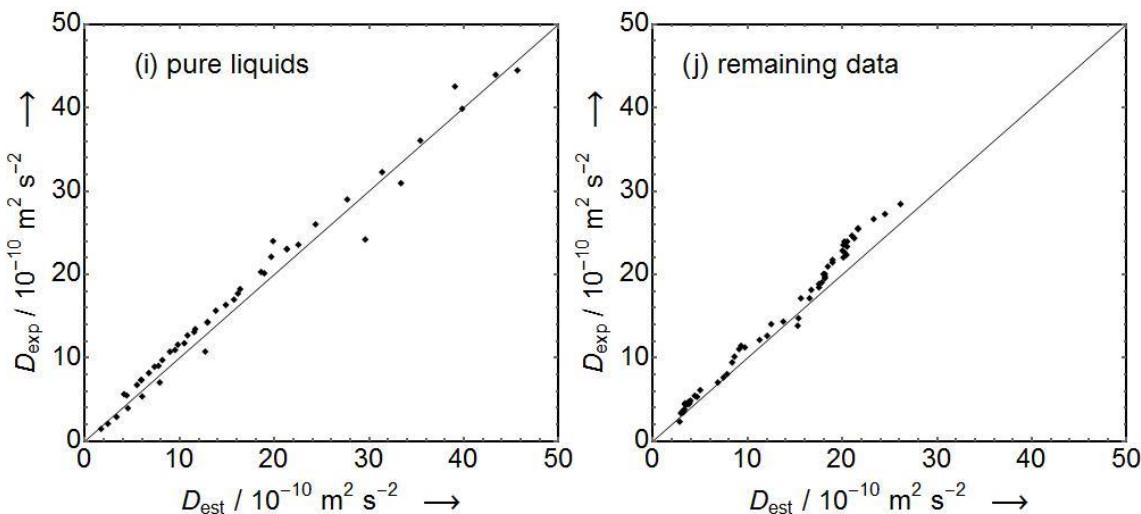
**Figure SI.6. Histogram of percentage differences between measured and estimated diffusion coefficient for original data set (red) and literature data set (black).**

## SI.7. Subsets of Literature Diffusion Data, Organised by Solvent Used

Analysis of the subsets of the complete data set can be studied to obtain further insight into the performance of the estimation method over a range of different experimental parameters. Of the nine deuteriated solvents used across the literature data, eight have sufficient diffusion coefficients reported to make plotting the individual fits worthwhile. Alongside these data for deuteriated solvents, there are several references that report measurements of diffusion in pure liquids. A tenth data set containing the remaining diffusion coefficients, 62 measurements in all, is also presented here. Figures SI.7.1 (a) through (j) plot the experimental diffusion coefficients for the ten subsets of diffusion data as functions of diffusion coefficients estimated using Eqn. 5.







**Figure SI.7.1. Measured diffusion coefficient plotted against diffusion coefficient estimated using Eqn. 5 for (a) 213 small molecules in  $\text{CDCl}_3$  (open diamonds) (b) 103 small molecules in  $\text{D}_2\text{O}$  (open squares), (c) 60 small molecules in  $\text{DMSO-d}_6$  (open triangles), (d) 54 small molecules in toluene- $\text{d}_8$  (open circles), (e) 39 small molecules in benzene- $\text{d}_6$  (closed triangles) (f) 30 small molecules in tetrahydrofuran- $\text{d}_8$  (filled circles) (g) 30 small molecules in cyclohexane- $\text{d}_{12}$  (h) 24 small molecules in methanol- $\text{d}_4$  (i) 47 measurements of pure liquids and (j) an additional 62 measurements of diffusion in a range of solvents, both protiated and deuteriated (all filled diamonds). Data are drawn from the combined data set. Data are drawn from the combined data set.**

Solvent	# diffusion coefficients	RMS error (%)
$\text{CDCl}_3$	213	12.4
$\text{D}_2\text{O}$	110	11.1
$\text{DMSO-d}_6$	60	17.8
Toluene- $\text{d}_8$	54	16.2
benzene- $\text{d}_6$	39	18.3
tetrahydrofuran- $\text{d}_8$	30	14.4
cyclohexane- $\text{d}_{12}$	30	22.8
methanol- $\text{d}_4$	24	19.3
pure liquids	47	14.3
additional measurements	62	14.2

**Table SI.7.1. Comparison of RMS error for subsets of the combined data set corresponding to the solvents  $\text{CDCl}_3$ ,  $\text{D}_2\text{O}$ , toluene- $\text{d}_8$ ,  $\text{DMSO-d}_6$ , benzene- $\text{d}_6$ , tetrahydrofuran- $\text{d}_8$ , cyclohexane- $\text{d}_{12}$  and methanol- $\text{d}_4$ . A further two subsets of the data corresponding to diffusion measurements of pure liquids and all remaining reported diffusion coefficients, respectively, are also included.**

## SI.8. Subset of Literature Diffusion Data, from Reference 37

Figure SI.8.1 is a reproduction of Figure 3 from the manuscript.

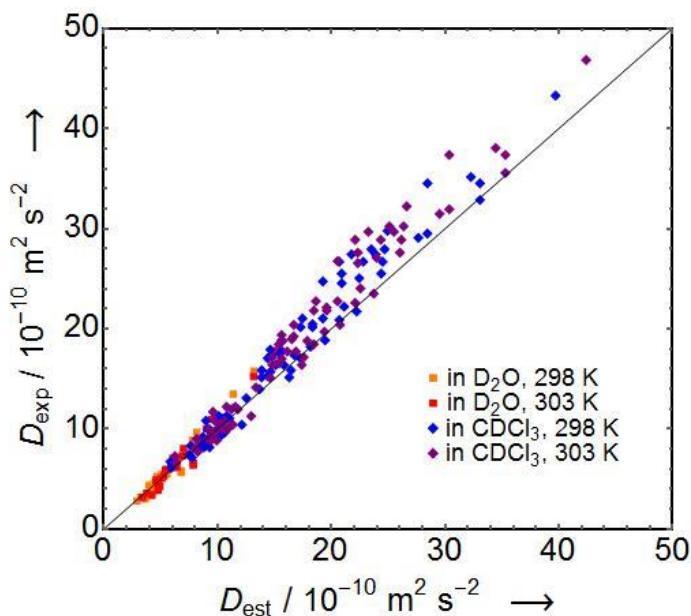
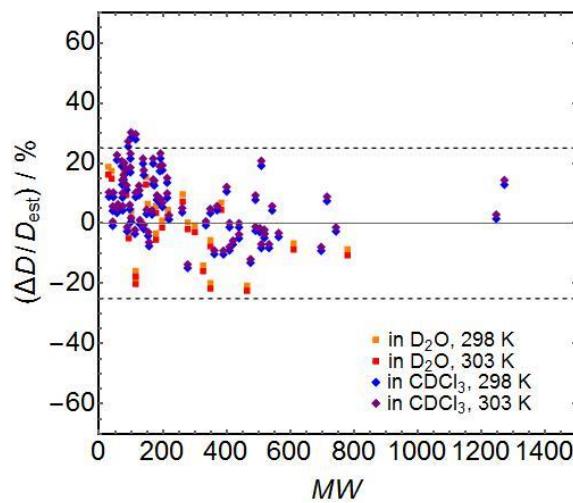
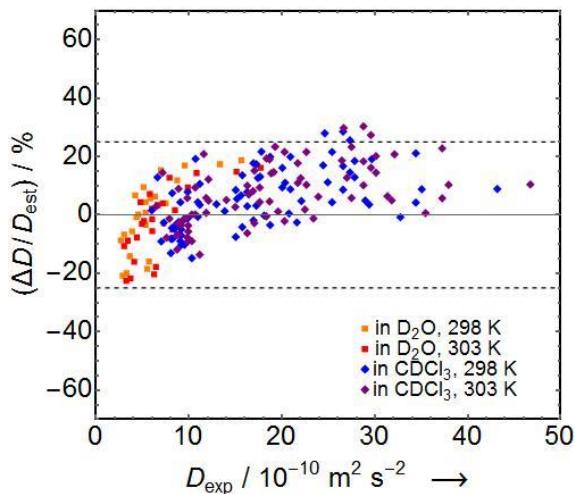


Figure SI.8.1. Measured diffusion coefficient plotted against diffusion coefficient calculated using Eqn. 5 for 24 small molecules in  $\text{D}_2\text{O}$  at 298 K (filled yellow squares), 24 small molecules in  $\text{D}_2\text{O}$  at 303 K (filled red squares), 76 small molecules in  $\text{CDCl}_3$  at 298 K (filled blue diamonds) and 76 small molecules in  $\text{CDCl}_3$  at 303 K (filled purple diamonds). Data are drawn from reference 37.

Figure SI.8.2 displays the diffusion data from Reference 37/Figure SI.8.1 as difference between calculated and measured diffusion coefficients, expressed as a percentage of experimental diffusion coefficient, plotted against (upper) molecular mass and (lower) experimental diffusion coefficient

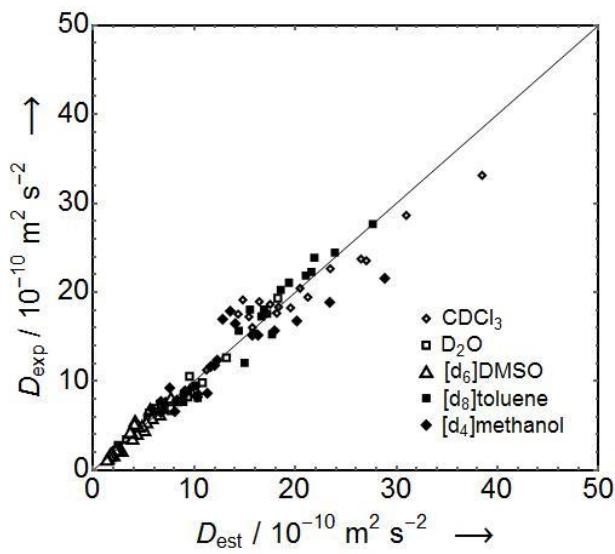




**Figure SI.8.2** Percentage difference between calculated and measured diffusion coefficients, expressed as a percentage of calculated diffusion coefficient, plotted against (upper) compound molecular mass and (lower) experimental diffusion coefficient for 24 small molecules in D<sub>2</sub>O at 298 K (filled yellow squares), 24 small molecules in D<sub>2</sub>O at 303 K (filled red squares), 76 small molecules in CDCl<sub>3</sub> at 298 K (filled blue diamonds) and 76 small molecules in CDCl<sub>3</sub> at 303 K (filled purple diamonds). Data drawn from reference 37.

### SI.9. Revised Approach, Fitting Solvent Hydrodynamic Radius

Based on the SEGWE method, here the hydrodynamic radii of the 5 original data set solvents are adjustable parameters while the radii of the samples are estimated from MW and an effective density. This approach improves the RMS fit to 12.9 % but at the expense of simplicity. The effective density is much reduced at 399 kg m<sup>-3</sup>. The fitted effective radii can be found in Table SI.9.



**Figure SI.9.** Measured diffusion coefficient plotted against diffusion coefficient calculated using SEGWE, modified to fit the variable radii of all five solvents used.

Solvent	Hydrodynamic radius / nm
D <sub>2</sub> O	1.64
CDCl <sub>3</sub>	2.63
DMSO-d <sub>6</sub>	2.63
CD <sub>3</sub> OD	2.09
C <sub>7</sub> D <sub>8</sub>	3.08

**Table SI.9 Summary of fitted hydrodynamic radii for the results depicted in Figure SI.9.**

### **SI.10. Summary of Additional Material found at doi:**

All data and software is available from Mendeley (<https://data.mendeley.com/datasets/cxt99xf2d2>, doi:10.17632/cxt99xf2d2.2). This resource contains:

- Excel sheet for calculation of diffusion coefficient from MW and *vice versa*.
- Matlab GUI (implemented for both Windows and Mac) for calculation of diffusion coefficient from MW and *vice versa*.
- Excel spreadsheet containing all diffusion coefficients used in this work.
- Mathematica notebook to reproduce all images and calculations in manuscript and in SI.
- Excel spreadsheet for original data set.
- Mathematica notebook to produce images and reproduce important calculations from original data set.
- Complete set of solvent data and calculations for Arrhenius parameters.
- Mathematica notebook to reproduce internal cross-validation of data set.