

Solubility Challenge: Can You Predict Solubilities of 32 Molecules Using a Database of 100 Reliable Measurements?

Antonio Llinàs,* Robert C. Glen, and Jonathan M. Goodman*

Pfizer Institute for Pharmaceutical Materials Science & Unilever Centre for Molecular Informatics,
Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, United Kingdom

Received February 15, 2008

Solubility is a key physicochemical property of molecules. Serious deficiencies exist in the consistency and reliability of solubility data in the literature. The accurate prediction of solubility would be very useful. However, systematic errors and lack of metadata associated with measurements greatly reduce the confidence in current models. To address this, we are accurately measuring intrinsic solubility values, and here we report results for a diverse set of 100 druglike molecules at 25 °C and an ionic strength of 0.15 M using the CheqSol approach. This is a highly reproducible potentiometric technique that ensures the thermodynamic equilibrium is reached rapidly. Results with a coefficient of variation higher than 4% were rejected. In addition, the Potentiometric Cycling for Polymorph Creation method, [PC]², was used to obtain multiple polymorph forms from aqueous solution. We now challenge researchers to predict the intrinsic solubility of 32 other druglike molecules that have been measured but are yet to be published.

INTRODUCTION

Solubility is a difficult property to predict, and one reason for this is the absence of a high-quality data set of reliable and reproducible solubility measurements. We present a database of 100 intrinsic solubilities measured for druglike molecules, and these data should be an ideal basis for the development of models for solubility (within this chemical space). We also present a list of 32 molecules for which we have measured, but not reported, the intrinsic solubility. We invite researchers to calculate solubilities for these compounds and publish their methods, results, and analysis in the *Journal of Chemical Information and Modeling*.

The discovery of new pharmaceuticals requires the identification of high quality drug candidates with appropriate physicochemical properties. Orally administered drugs generally require sufficient water solubility to be bioavailable. Putative drug molecules that are insoluble in water cannot be tested in biological assays, have poor pharmacological profiles, and may come out of solution under storage conditions. For example, at concentrations of more than 2 mg mL⁻¹, haloperidol precipitates after 24 h in mixtures with diamorphine.¹ Clinical studies are often hampered by poor solubility characteristics, and precipitation may lead to several side effects, such as hemolysis and phlebitis,² and toxicity, such as crystalluria.³

Poor solubility and associated poor pharmacokinetics under physiological conditions can often result in expensive late-stage failure in drug discovery. It is estimated that up to 40% of drug discovery programs are abandoned due to problems with the pharmacokinetics of lead compounds.^{4,5} With the

advent of high-throughput synthesis and testing, the average molecular weight and lipophilicity of test compounds have increased over the years, with a concomitant decrease in solubility.⁶ The introduction of empirical rules derived from the experience of successful drug development projects, such as Lipinski's Rule of Five,^{7,8} has helped the situation but has not solved the problem. In order to allow in vitro testing of low solubility drugs, a cosolvent, such as DMSO, is often used. This facilitates the dissolution of the compound but can severely affect biochemical assays and also may lead to compounds of low solubility entering development. This can result in problems in formulation and delivery where inconsistent methods of solubility measurement can result in unexpected results when carried forward to dosage forms. For example, the salt of a compound may have a large aqueous solubility and yet may precipitate when injected because of a lower solubility of the free acid or the free base that forms at physiological pH. An example is loperamide⁹ which has a reported literature value of aqueous solubility of log *S* = -4.38, a value that was measured for the HCl salt. The intrinsic solubility (reported below) is only log *S* = -7.07. Similar issues arise for methotrexate. The renal toxicity of methotrexate is related to its precipitation in the renal tubules and collecting ducts. Levels in excess of 1 mM⁻¹ exceed the solubility of methotrexate at pH 5.0.³ If a compound can exist in different polymorphs, it may have different solubilities for each form. Norvir, a protease inhibitor from Abbott Laboratories, had to be reformulated in 1998 after the emergence of a polymorph, which halved the bioavailability. If new substances are known to have low solubility or to show polymorphism, the FDA will require expensive additional stability and bioequivalence studies.

There are many computational models in the literature for predicting the aqueous solubility of druglike molecules, but these usually have an uncertainty of about an order of magnitude. Any aqueous solubility prediction method pro-

* Corresponding author phone: (+44) (0)1223 336434; fax: (+44) (0)1223 763076; e-mail: J.M.Goodman@ch.cam.ac.uk (J.M.G.) and phone: (+46) 31 706 4132; fax: (+46) 31 776 3748; e-mail: Antonio.Llinas@astrazeneca.com (A.L.). Current address: AstraZeneca R&D Mölndal, SE-431 83 Mölndal, Sweden.

ducing standard errors of estimate 0.4–0.5 log unit for a diverse data set is probably overfitting the data.¹⁰ One reason for this is that the data used to generate the model are often of poor quality. The data are generally combined from values obtained under varied experimental conditions. There are many definitions of solubility, and it is not always clear which has been reported. For example, kinetic solubilities may have been mistaken for thermodynamic values. Measurements may be made in distilled water or in pH-buffered water, and these can lead to dramatically different results. The size and diversity of the training set can severely limit the scope of models that may only have application to a restricted region of chemical space.¹¹

All thermodynamic solubility measurements are difficult because it is hard to ensure equilibrium conditions have been reached, as supersaturated solutions often form readily. The conditions used for the experiments must be precisely reported, as small changes in pH, dissolved salts, and temperature can substantially affect solubility. Crystal form, polymorphism, and crystal hydrate formation can also be important. The lack of reliable fully characterized data has held back both the understanding of the phenomenon and the derivation of useful models to facilitate solubility prediction, a situation we are attempting to address, and challenge.

In this study we focus on intrinsic solubility, S_0 , which is defined as the solubility of a compound in its free acid or free base form.¹² This measure of solubility does not depend on pH and so is rather more reproducible than other measures. For example, we have measured the aqueous solubility of diclofenac, at 25 °C and 0.15 M ionic strength, as $6.06 \mu\text{g mL}^{-1}$ ($20.5 \mu\text{M}$), which is around one-third of the value reported by Chiarini¹³ at neutral pH, and about three-hundredth of the value reported by Fini¹⁴ using the shake flask method. Values for the aqueous solubility of diclofenac range from 17.8 to $1771 \mu\text{g mL}^{-1}$.¹⁴ It is not clear whether some of the measurements were performed in buffers, at a specific pH, or in pure water, which may lead to the system reaching an undetermined pH, because carbon-dioxide dissolved in the solution can significantly affect the pH and hence the reported solubility value. Our value for the intrinsic solubility, $1.03 \pm 0.07 \mu\text{g mL}^{-1}$, agrees well with some previously reported,^{14–16} measurements of the intrinsic solubility of diclofenac ($0.9 \pm 0.1 \mu\text{g mL}^{-1}$, $0.8 \pm 0.2 \mu\text{g mL}^{-1}$, and $1 \mu\text{g mL}^{-1}$) but disagrees with others: $2.36 \mu\text{g mL}^{-1}$.¹⁷

As a result of the difficulties in measuring solubility, data in the literature cannot easily be combined to give a database sufficiently large and diverse to cover the chemical space of druglike molecules, and this is probably one reason why the predictive power of models derived from such data is generally poor.¹⁰ According to Jorgensen and Duffy the average uncertainty in measured aqueous solubility values is probably no better than 0.6 log unit.¹⁸ Measured values can differ significantly when changing very small details of the experimental method, such as reseeding the original supersaturated solution with the original solid material. This is especially true at solubilities close to or below 0.3 mM (about 0.1 mg/mL for a typical druglike molecule), which is frequently referred to as the minimum solubility for avoiding dissolution-controlled absorption of orally administered drugs.¹⁹

Most solubility measurements are reported with no reference to the solid state of the material being studied. It may be crystalline, or amorphous, or even both, if the results from two or more studies are combined to give an average solubility. Differences between crystalline and amorphous solubility can be large.²⁰ Different polymorphs can show also differences in solubility.²¹ These differences can of course have significant effects on the observed pharmacokinetics of the formulation.²²

Early attempts to include consideration of the solid phase in predictive models include the General Solubility Equation (GSE) of Yalkowski, which includes the melting point (i.e. a solid state property) of the material.²³ Recent work also attempts to derive melting points for inclusion in modified GSEs.²⁴ However, it is clear that until an adequate description of the lattice energy²⁵ (or the crystalline state) of the material is available, progress on predicting solubilities *ab initio* will be limited.^{26–28}

A comprehensive data set of diverse compounds and associated characterization data is required to underpin a more fundamental understanding of solubility. To accomplish this, we have adopted a potentiometric method for the measurement of intrinsic solubility that is very accurate and allows rapid equilibration of the experiment and collection of the precipitate during the experiment for characterization. The ‘Chasing Equilibrium’ technique (CheqSol)^{15,21,29} produces a precipitate after several cycles, switching back and forth between a supersaturated and a subsaturated solution. The final precipitate obtained is thermodynamically driven (and not kinetically driven), and the solubility data are highly reproducible with an associated error of about 0.05 log units.

EXPERIMENTAL SECTION

Reagents and Solutions. The HCl titrant solution (0.5071 N) was purchased ready-made (Sigma Aldrich, Gillingham, U.K.). The KOH titrant solution (0.5 M) was prepared by diluting an ampule of concentrate supplied by Fisher Scientific (Loughborough, Leicestershire, U.K.). The KOH was standardized by titrating against a weighed quantity of potassium phthalate (KHP) dissolved in 15 mL of 0.15 M KCl solution. The HCl was standardized by titrating a measured volume against the standardized KOH. KHP, KCL, and methanol (analytical grade) were used without further purification. Deionized water of resistivity $> 1014 \Omega \text{ cm}$ was used to prepare all solutions.

Solubility Measurements. The apparatus used to perform the solubility determinations was a Sirius GLpKa titrator fitted with a pH electrode (combination Ag–Ag–Cl) and a Sirius D-PAS spectrometer fitted with a bifurcated fiber-optic dip probe (Hellma, Southend-on-Sea, Essex, U.K.) with a path length of 1 cm. The instrument was controlled from a computer running RefinementPro and CheqSol software (Sirius Analytical Instruments Ltd., Forest Row, East Sussex, U.K.). All experiments were performed in 0.15 M KCl solution under nitrogen atmosphere, at 25 ± 0.2 °C, using standardized 0.5 M HCl and 0.5 M KOH solutions. The pH electrode was calibrated in the pH range 1.8–12.2. An overhead stirrer was used, and the temperature was monitored during the course of the measurements.

Analysis of Results. The data in Table 1 were transcribed from the CheqSol program’s analysis of each experiment.

Table 1. Intrinsic Solubility for 100 Compounds

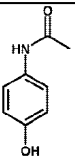
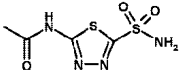
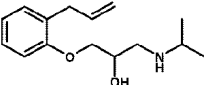

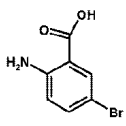
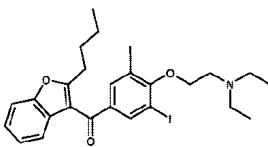
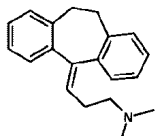
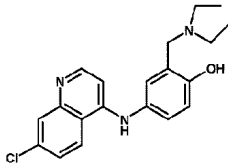
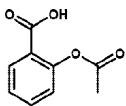
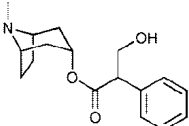
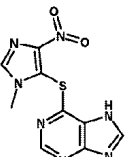
Structure	Name	MW Neutral Form	pKa	Kinetic Solubility μM $\mu\text{g/ml}$	Intrinsic Solubility μM $\mu\text{g/ml}$
	Acetaminophen	151.17	9.52 ± 0.01	$161700 \pm 7000 \mu\text{M}$ $24400 \pm 1060 \mu\text{g/ml}$	$86300 \pm 7000 \mu\text{M}$ $13000 \pm 1060 \mu\text{g/ml}$
	Acetazolamide	222.25	8.75 ± 0.02 7.31 ± 0.04	$6100 \pm 3840 \mu\text{M}$ $1360 \pm 850 \mu\text{g/ml}$	$3670 \pm 80 \mu\text{M}$ $816 \pm 18 \mu\text{g/ml}$
	Alprenolol	249.36	9.47 ± 0.01	$5080 \pm 50 \mu\text{M}$ $1266 \pm 12 \mu\text{g/ml}$	$2320 \pm 40 \mu\text{M}$ $580 \pm 10 \mu\text{g/ml}$
	Amantadine	151.25	10.48 ± 0.01	$17300 \pm 3960 \mu\text{M}$ $2620 \pm 600 \mu\text{g/ml}$	$14000 \pm 1180 \mu\text{M}$ $2120 \pm 180 \mu\text{g/ml}$
	2-Amino-5-bromobenzoic acid	216.04	4.215 ± 0.035 1.743 ± 0.072	$2560 \pm 860 \mu\text{M}$ $553 \pm 186 \mu\text{g/ml}$	$843 \pm 15 \mu\text{M}$ $182 \pm 3 \mu\text{g/ml}$
	Amiodarone	645.32	8.73 ± 0.05	b	$0.0067 \mu\text{M}$ $0.0043 \mu\text{g/ml}$
	Amitriptyline	277.41	9.46 ± 0.02	$29 \pm 3 \mu\text{M}$ $8 \pm 1 \mu\text{g/ml}$	$28.2 \pm 2.9 \mu\text{M}$ $7.8 \pm 0.8 \mu\text{g/ml}$
	Amodiaquine	355.87	7.4 ± 0.2 8.24 ± 0.04 11.5 ± 0.1	$24 \pm 5 \mu\text{M}$ $8.5 \pm 1.8 \mu\text{g/ml}$	$1.62 \pm 0.44 \mu\text{M}$ $0.57 \pm 0.15 \mu\text{g/ml}$
	Aspirin	180.16	3.50 ± 0.01		Decomposes (hydrolysis)
	Atropine	289.38	9.790 ± 0.024	$12400 \pm 4400 \mu\text{M}$ $3900 \pm 1000 \mu\text{g/ml}$	$9900 \pm 460 \mu\text{M}$ $2880 \pm 130 \mu\text{g/ml}$
	Azathioprine	277.27	7.834 ± 0.002 1.064 ± 0.047	$6700 \pm 1600 \mu\text{M}$ $3600 \pm 460 \mu\text{g/ml}$	$620 \pm 16 \mu\text{M}$ $172 \pm 4 \mu\text{g/ml}$

Table 1. Continued

Structure	Name	MW Neutral Form	pKa	Kinetic Solubility μM $\mu\text{g/ml}$	Intrinsic Solubility μM $\mu\text{g/ml}$
	Benzylimidazole	158.20	6.70 ± 0.04	$48500 \pm 8100 \mu\text{M}$ $7700 \pm 1300 \mu\text{g/ml}$	$5500 \pm 1100 \mu\text{M}$ $874 \pm 174 \mu\text{g/ml}$
	5-Bromo-2,4-dihydroxybenzoic acid	233.02	7.138 ± 0.004 2.933 ± 0.006	$8730 \pm 750 \mu\text{M}$ $2030 \pm 170 \mu\text{g/ml}$	$2397 \pm 40 \mu\text{M}$ $559 \pm 9 \mu\text{g/ml}$
	Bromogramine	253.14	9.659 ± 0.004	$390 \pm 90 \mu\text{M}$ $100 \pm 20 \mu\text{g/ml}$	$88 \pm 6 \mu\text{M}$ $22.3 \pm 1.6 \mu\text{g/ml}$
	Bupivacaine	288.44	7.952 ± 0.049	$1970 \pm 70 \mu\text{M}$ $570 \pm 20 \mu\text{g/ml}$	$600 \pm 33 \mu\text{M}$ $172 \pm 9 \mu\text{g/ml}$
	Carprofen	273.72	4.247 ± 0.029	$50 \pm 13 \mu\text{M}$ $14 \pm 4 \mu\text{g/ml}$	$20 \pm 3 \mu\text{M}$ $5.5 \pm 0.8 \mu\text{g/ml}$
	Carvedilol ^c	406.49	7.97 ± 0.01	$63.3 \mu\text{M}$ $25.7 \mu\text{g/ml}$	$55.5 \pm 3.6 \% \mu\text{M}$ $22.6 \pm 3.6 \% \mu\text{g/ml}$
	Cephalothin	396.44	1.963 ± 0.025	$5600 \pm 600 \mu\text{M}$ $2230 \pm 240 \mu\text{g/ml}$	$1154 \pm 3 \mu\text{M}$ $457 \pm 1 \mu\text{g/ml}$
	Chlorpheniramine	274.80	3.819 ± 0.077 9.253 ± 0.063	$2500 \pm 83 \mu\text{M}$ $690 \pm 23 \mu\text{g/ml}$	$2150 \pm 150 \mu\text{M}$ $590 \pm 40 \mu\text{g/ml}$
	Chlorpromazine	318.87	9.24 ± 0.02	$8.9 \pm 0.8 \mu\text{M}$ $2.8 \pm 0.3 \mu\text{g/ml}$	$8.5 \pm 0.1 \mu\text{M}$ $2.70 \pm 0.05 \mu\text{g/ml}$
	Chlorpropamide	276.74	4.669 ± 0.004	$1700 \pm 400 \mu\text{M}$ $470 \pm 120 \mu\text{g/ml}$	$564 \pm 10 \mu\text{M}$ $156 \pm 3 \mu\text{g/ml}$
	Chlorprothixene FORM I	315.87	9.520 ± 0.005	a	$0.178 \pm 0.008 \mu\text{M}$ $0.056 \pm 0.003 \mu\text{g/ml}$
	Chlorprothixene FORM II	315.87	9.520 ± 0.005	$1.6 \pm 0.1 \mu\text{M}$ $0.50 \pm 0.03 \mu\text{g/ml}$	$1.36 \pm 0.12 \mu\text{M}$ $0.43 \pm 0.04 \mu\text{g/ml}$

Table 1. Continued

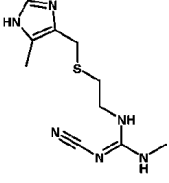
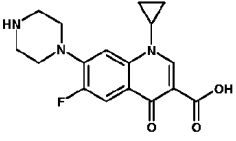
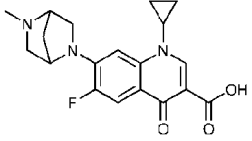
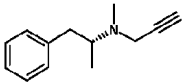
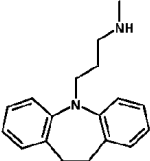
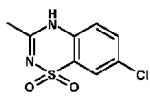
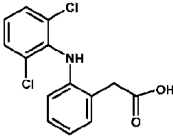
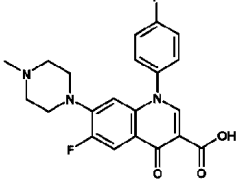
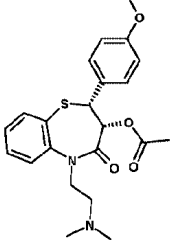
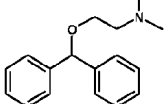
Structure	Name	MW Neutral Form	pKa	Kinetic Solubility μM $\mu\text{g/ml}$	Intrinsic Solubility μM $\mu\text{g/ml}$
	Cimetidine	252.34	6.998 ± 0.039	$86000 \pm 2000 \mu\text{M}$ $21700 \pm 500 \mu\text{g/ml}$	$20340 \pm 480 \mu\text{M}$ $5130 \pm 120 \mu\text{g/ml}$
	Ciprofloxacin	331.35	8.579 ± 0.004 6.234 ± 0.004	$2200 \pm 1200 \mu\text{M}$ $740 \pm 400 \mu\text{g/ml}$	$253 \pm 27 \mu\text{M}$ $84 \pm 9 \mu\text{g/ml}$
	Danofloxacin ^c	357.38	8.727 ± 0.010 6.748 ± 0.084	$1935 \mu\text{M}$ $691 \mu\text{g/ml}$	$1260 \pm 1.68 \% \mu\text{M}$ $450.4 \pm 1.68 \% \mu\text{g/ml}$
	Deprenyl (selegiline)	187.29	7.48 ± 0.05	$3200 \pm 100 \mu\text{M}$ $600 \pm 20 \mu\text{g/ml}$	$3070 \pm 130 \mu\text{M}$ $575 \pm 24 \mu\text{g/ml}$
	Desipramine	266.39	10.26 ± 0.05	$255 \pm 5 \mu\text{M}$ $67 \pm 7 \mu\text{g/ml}$	$236 \pm 8 \mu\text{M}$ $63 \pm 2 \mu\text{g/ml}$
	Diazoxide	230.67	8.629 ± 0.042	$2290 \pm 430 \mu\text{M}$ $530 \pm 100 \mu\text{g/ml}$	$434 \pm 23 \mu\text{M}$ $100 \pm 5 \mu\text{g/ml}$
	Diclofenac	296.16	4.081 ± 0.036	$148 \pm 40 \mu\text{M}$ $44 \pm 12 \mu\text{g/ml}$	$3.5 \pm 0.1 \mu\text{M}$ $1.036 \pm 0.03 \mu\text{g/ml}$
	Difloxacin	399.40	7.287 ± 0.064 5.793 ± 0.025	$820 \pm 10 \mu\text{M}$ $330 \pm 2 \mu\text{g/ml}$	$251 \pm 2 \mu\text{M}$ $100 \pm 1 \mu\text{g/ml}$
	Diltiazem	414.53	8.02 ± 0.03	$880 \pm 50 \mu\text{M}$ $365 \pm 21 \mu\text{g/ml}$	$694 \pm 45 \mu\text{M}$ $287 \pm 19 \mu\text{g/ml}$
	Diphenhydramine	255.36	9.08 ± 0.01	$1240 \pm 130 \mu\text{M}$ $320 \pm 30 \mu\text{g/ml}$	$1130 \pm 28 \mu\text{M}$ $289 \pm 7 \mu\text{g/ml}$

Table 1. Continued

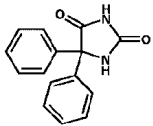
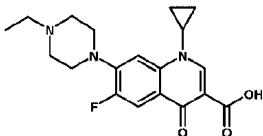
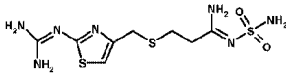
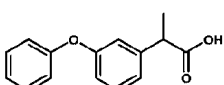
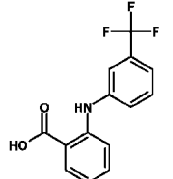
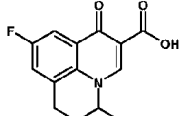
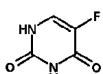
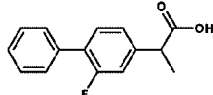
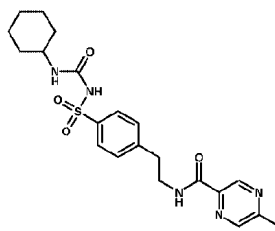
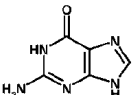
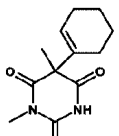
Structure	Name	MW Neutral Form	pKa	Kinetic Solubility μM $\mu\text{g/ml}$	Intrinsic Solubility μM $\mu\text{g/ml}$
	5,5-Diphenylhydantoin (Dilantin)	252.268	8.215 \pm 0.038	480 \pm 230 μM 120 \pm 60 $\mu\text{g/ml}$	139 \pm 16 μM 35 \pm 4 $\mu\text{g/ml}$
	Enrofloxacin	359.40	7.794 \pm 0.004 6.204 \pm 0.004	5000 \pm 1000 μM 1800 \pm 360 $\mu\text{g/ml}$	660 \pm 80 μM 237 \pm 29 $\mu\text{g/ml}$
	Famotidine	337.45	1.80 \pm 0.02 6.77 \pm 0.01 11.01 \pm 0.03	18000 \pm 1250 μM 6080 \pm 420 $\mu\text{g/ml}$	2250 \pm 50 μM 760 \pm 16 $\mu\text{g/ml}$
	Fenopropfen	242.28	3.79 \pm 0.26	270 \pm 28 μM 65 \pm 7 $\mu\text{g/ml}$	200 \pm 30 μM 48 \pm 7 $\mu\text{g/ml}$
	Flufenamic acid	281.24	3.97 \pm 0.05	45.2 \pm 0.6 μM 12.7 \pm 0.2 $\mu\text{g/ml}$	4.42 \pm 0.08 μM 1.24 \pm 0.02 $\mu\text{g/ml}$
	Flumequine	261.26	6.514 \pm 0.017	680 \pm 100 μM 180 \pm 30 $\mu\text{g/ml}$	185 \pm 12 μM 48 \pm 3 $\mu\text{g/ml}$
	5-Fluorouracil	130.08	7.806 \pm 0.003	> 1180000 μM > 153000 $\mu\text{g/ml}$	d
	Flurbiprofen	244.27	3.98 \pm 0.03	200 \pm 130 μM 50 \pm 30 $\mu\text{g/ml}$	70.5 \pm 0.4 μM 17.2 \pm 0.1 $\mu\text{g/ml}$
	Glipizide	445.54	5.13 \pm 0.01	90 \pm 60 μM 40 \pm 28 $\mu\text{g/ml}$	3.25 \pm 0.09 μM 1.45 \pm 0.04 $\mu\text{g/ml}$
	Guanine	151.13	3.181 \pm 0.042 9.118 \pm 0.042 11.9 \pm 0.1	670 \pm 200 μM 100 \pm 30 $\mu\text{g/ml}$	37 \pm 9 μM 5.6 \pm 1.3 $\mu\text{g/ml}$
	Hexobarbital	236.27	8.30 \pm 0.02	6000 \pm 1000 μM 1400 \pm 230 $\mu\text{g/ml}$	2120 \pm 100 μM 500 \pm 26 $\mu\text{g/ml}$

Table 1. Continued

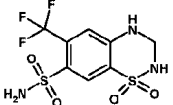
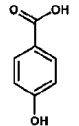
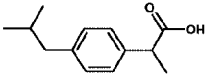
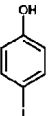
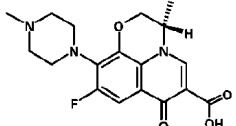
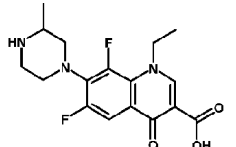
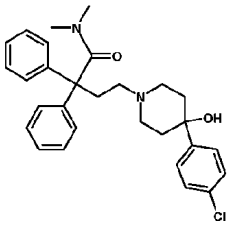
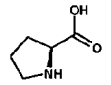
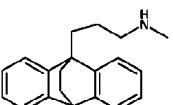
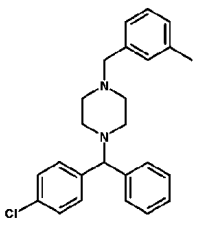
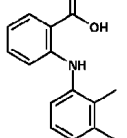
Structure	Name	MW Neutral Form	pKa	Kinetic Solubility μM $\mu\text{g/ml}$	Intrinsic Solubility μM $\mu\text{g/ml}$
	Hydroflumethiazide	331.29	9.750 ± 0.015 8.600 ± 0.031	$8000 \pm 5000 \mu\text{M}$ $2650 \pm 1600 \mu\text{g/ml}$	$1080 \pm 300 \mu\text{M}$ $360 \pm 100 \mu\text{g/ml}$
	4-Hydroxybenzoic acid	138.12	8.952 ± 0.003 4.303 ± 0.003	$53000 \pm 1500 \mu\text{M}$ $7300 \pm 740 \mu\text{g/ml}$	$34320 \pm 670 \mu\text{M}$ $4740 \pm 93 \mu\text{g/ml}$
	Ibuprofen	206.29	4.46 ± 0.16	$940 \pm 60 \mu\text{M}$ $190 \pm 12 \mu\text{g/ml}$	$254 \pm 16 \mu\text{M}$ $52 \pm 3 \mu\text{g/ml}$
	4-Iodophenol	220.01	9.073 ± 0.005	$25000 \pm 8400 \mu\text{M}$ $5500 \pm 1800 \mu\text{g/ml}$	$19300 \pm 600 \mu\text{M}$ $4250 \pm 130 \mu\text{g/ml}$
	Levofloxacin	361.38	8.263 ± 0.045 6.238 ± 0.022	$> 140000 \mu\text{M}$ $> 50000 \mu\text{g/ml}$	d
	Lomefloxacin ^c	351.36	9.100 ± 0.078 5.858 ± 0.049	$55550 \mu\text{M}$ $19500 \mu\text{g/ml}$	$4647 \pm 0.91\% \mu\text{M}$ $1631 \pm 0.91\% \mu\text{g/ml}$
	Loperamide	477.05	8.90 ± 0.07	b	$0.08427 \mu\text{M}$ $0.0402 \mu\text{g/ml}$
	L-Proline	115.13	10.468 ± 0.004 2.016 ± 0.011	$> 535000 \mu\text{M}$ $> 61500 \mu\text{g/ml}$	d
	Maprotiline	277.43	10.330 ± 0.071	$750 \pm 600 \mu\text{M}$ $200 \pm 170 \mu\text{g/ml}$	$20.3 \pm 0.9 \mu\text{M}$ $5.6 \pm 0.2 \mu\text{g/ml}$
	Meclizine	390.96	2.27 ± 0.05 7.28 ± 0.05	b	$0.33 \mu\text{M}$ $0.13 \mu\text{g/ml}$
	Mefenamic acid	241.29	4.22 ± 0.01	$0.60 \pm 0.07 \mu\text{M}$ $0.14 \pm 0.02 \mu\text{g/ml}$	$0.183 \pm 0.001 \mu\text{M}$ $0.0442 \pm 0.0003 \mu\text{g/ml}$

Table 1. Continued

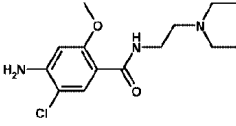
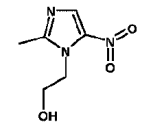
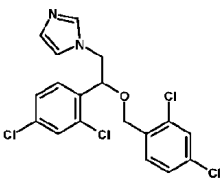
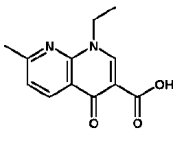
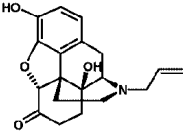

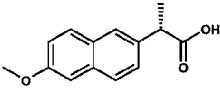
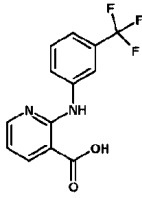
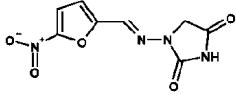
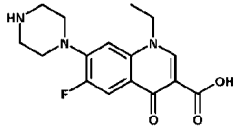
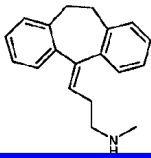
Structure	Name	MW Neutral Form	pKa	Kinetic Solubility μM $\mu\text{g/ml}$	Intrinsic Solubility μM $\mu\text{g/ml}$
	Metoclopramide	299.80	9.240 \pm 0.004	1190 \pm 280 μM 360 \pm 80 $\mu\text{g/ml}$	272 \pm 16 μM 82 \pm 5 $\mu\text{g/ml}$
	Metronidazole (Flagyl)	171.16	2.476 \pm 0.006	92800 \pm 7900 μM 15900 \pm 1360 $\mu\text{g/ml}$	60000 \pm 1500 μM 10280 \pm 260 $\mu\text{g/ml}$
	Miconazole	416.14	6.135 \pm 0.051	27 \pm 12 μM 11 \pm 5 $\mu\text{g/ml}$	8.5 \pm 0.3 μM 3.5 \pm 0.1 $\mu\text{g/ml}$
	Nalidixic acid	232.24	6.036 \pm 0.029 2.1 \pm 0.1	530 \pm 80 μM 123 \pm 18 $\mu\text{g/ml}$	245 \pm 5 μM 57 \pm 1 $\mu\text{g/ml}$
	Naloxone	327.38	9.129 \pm 0.003 8.029 \pm 0.003	6370 \pm 650 μM 2080 \pm 210 $\mu\text{g/ml}$	1265 \pm 89 μM 414 \pm 29 $\mu\text{g/ml}$
	1-Naphthol	144.17	9.17 \pm 0.05	18000 \pm 11800 μM 2600 \pm 1700 $\mu\text{g/ml}$	10400 \pm 400 μM 1500 \pm 60 $\mu\text{g/ml}$
	Naproxen	230.27	4.18 \pm 0.01	86 \pm 14 μM 20 \pm 3 $\mu\text{g/ml}$	31.9 \pm 0.5 μM 7.3 \pm 0.1 $\mu\text{g/ml}$
	Niflumic acid	282.22	2.26 \pm 0.01 4.44 \pm 0.01	160 \pm 70 μM 45 \pm 19 $\mu\text{g/ml}$	26 \pm 5 μM 7 \pm 1 $\mu\text{g/ml}$
	Nitrofurantoin	238.16	7.050 \pm 0.043	1660 \pm 490 μM 390 \pm 100 $\mu\text{g/ml}$	577 \pm 27 μM 137 \pm 6 $\mu\text{g/ml}$
	Norfloxacin ^c	319.34	8.575 \pm 0.005 6.324 \pm 0.005	10640 μM 3396 $\mu\text{g/ml}$	1752 \pm 3.78 μM 559 \pm 3.78 $\mu\text{g/ml}$
	Nortriptyline	263.39	10.21 \pm 0.01	102 \pm 2 μM 27.0 \pm 0.5 $\mu\text{g/ml}$	96 \pm 5 μM 25 \pm 1 $\mu\text{g/ml}$

Table 1. Continued

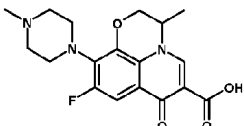
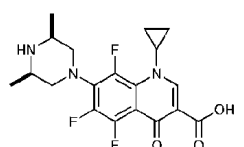
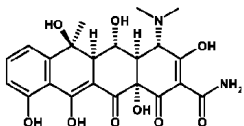
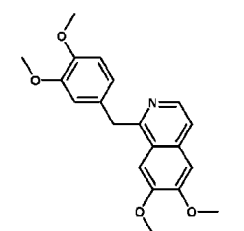
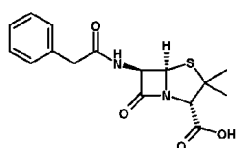
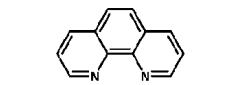
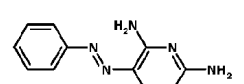
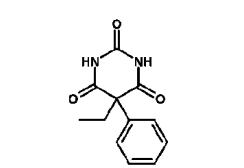
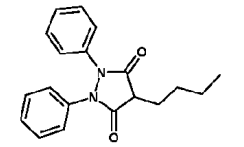
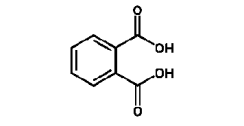
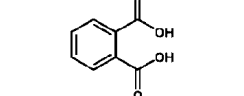
Structure	Name	MW Neutral Form	pKa	Kinetic Solubility μM $\mu\text{g/ml}$	Intrinsic Solubility μM $\mu\text{g/ml}$
	Ofloxacin	361.38	8.193 ± 0.003 6.314 ± 0.003	$80000 \pm 10000 \mu\text{M}$ $30000 \pm 3600 \mu\text{g/ml}$	$54260 \pm 560 \mu\text{M}$ $19600 \pm 200 \mu\text{g/ml}$
	Orbifloxacin	395.38	8.822 ± 0.015 5.903 ± 0.027	$> 26800 \mu\text{M}$ $> 10600 \mu\text{g/ml}$	d
	Oxytetracycline	460.44	12.6 ± 0.7 8.745 ± 0.007 7.272 ± 0.004 3.247 ± 0.007	$19800 \pm 2500 \mu\text{M}$ $9100 \pm 1170 \mu\text{g/ml}$	$820 \pm 180 \mu\text{M}$ $380 \pm 80 \mu\text{g/ml}$
	Papaverine	339.39	5.95 ± 0.01	$2550 \pm 220 \mu\text{M}$ $865 \pm 75 \mu\text{g/ml}$	$136 \pm 12 \mu\text{M}$ $46 \pm 4 \mu\text{g/ml}$
	Pen G (Benzylpenicillin)	334.40	2.519 ± 0.013		DECOMPOSITION (β -lactam hydrolysis)
	Phenantroline ^c	180.21	4.94 ± 0.01	$112000 \mu\text{M}$ $20180 \mu\text{g/ml}$	$24100 \pm 1.75 \% \mu\text{M}$ $4342 \pm 1.75 \% \mu\text{g/ml}$
	Phenazopyridine	213.24	5.154 ± 0.013	$400 \pm 70 \mu\text{M}$ $85 \pm 15 \mu\text{g/ml}$	$64 \pm 2 \mu\text{M}$ $13.7 \pm 0.5 \mu\text{g/ml}$
	Phenobarbital	232.24	7.187 ± 0.002 11.85 ± 0.01	$11200 \pm 3500 \mu\text{M}$ $2600 \pm 800 \mu\text{g/ml}$	$5090 \pm 280 \mu\text{M}$ $1180 \pm 64 \mu\text{g/ml}$
	Phenylbutazone	308.38	4.35 ± 0.01	$162 \pm 7 \mu\text{M}$ $50 \pm 2 \mu\text{g/ml}$	$40.66 \pm 0.04 \mu\text{M}$ $12.54 \pm 0.01 \mu\text{g/ml}$
	Phthalic acid FORM I	166.13	2.70 ± 0.01 4.84 ± 0.01	a	$32160 \pm 2300 \mu\text{M}$ $5350 \pm 380 \mu\text{g/ml}$
	Phthalic acid FORM II	166.13	2.70 ± 0.01 4.84 ± 0.01	$53000 \pm 18000 \mu\text{M}$ $8800 \pm 3000 \mu\text{g/ml}$	$24800 \pm 860 \mu\text{M}$ $4130 \pm 140 \mu\text{g/ml}$

Table 1. Continued

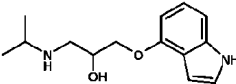
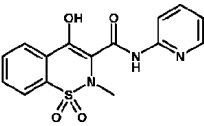
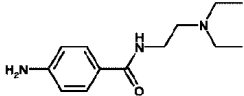
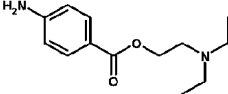
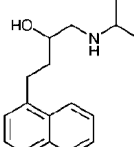
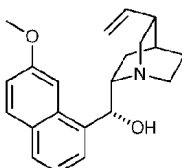
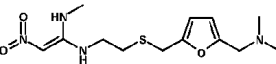
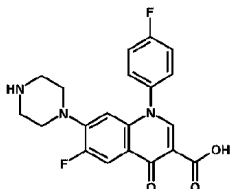
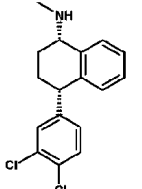
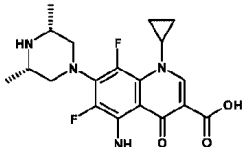
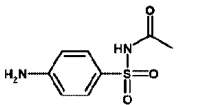
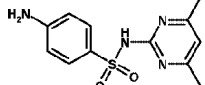
Structure	Name	MW Neutral Form	pKa	Kinetic Solubility μM $\mu\text{g/ml}$	Intrinsic Solubility μM $\mu\text{g/ml}$
	Pindolol	248.33	9.54 ± 0.05	$1930 \pm 1300 \mu\text{M}$ $480 \pm 320 \mu\text{g/ml}$	$163 \pm 4 \mu\text{M}$ $40 \pm 1 \mu\text{g/ml}$
	Piroxicam	331.35	1.87 ± 0.01 5.29 ± 0.01	$590 \pm 109 \mu\text{M}$ $196 \pm 36 \mu\text{g/ml}$	$15.8 \pm 0.5 \mu\text{M}$ $5.2 \pm 0.2 \mu\text{g/ml}$
	Procainamide	235.33	2.83 ± 0.04 9.25 ± 0.01	$> 43800 \mu\text{M}$ $> 10300 \mu\text{g/ml}$	d
	Procaine	236.32	2.29 ± 0.01 9.04 ± 0.01	$19800 \pm 400 \mu\text{M}$ $4700 \pm 100 \mu\text{g/ml}$	$19100 \pm 820 \mu\text{M}$ $4500 \pm 200 \mu\text{g/ml}$
	Propranolol	259.35	9.54 ± 0.01	$1130 \pm 270 \mu\text{M}$ $290 \pm 70 \mu\text{g/ml}$	$320 \pm 16 \mu\text{M}$ $83 \pm 4 \mu\text{g/ml}$
	Quinine	324.43	4.26 ± 0.04 8.55 ± 0.09	$1450 \pm 86 \mu\text{M}$ $469 \pm 28 \mu\text{g/ml}$	$1635 \pm 36 \mu\text{M}$ $530 \pm 12 \mu\text{g/ml}$
	Ranitidine	314.41	8.423 ± 0.005 2.374 ± 0.018	$52200 \mu\text{M}^c$ $16400 \mu\text{g/ml}$	$3160 \pm 500 \mu\text{M}$ $990 \pm 160 \mu\text{g/ml}$
	Sarafloxacin	385.37	8.5 ± 0.1 5.970 ± 0.049	$2050 \pm 200 \mu\text{M}$ $790 \pm 80 \mu\text{g/ml}$	$740 \pm 60 \mu\text{M}$ $280 \pm 20 \mu\text{g/ml}$
	Sertraline	306.24	9.628 ± 0.067	b	$14.8 \mu\text{M}$ $4.5 \mu\text{g/ml}$
	Sparfloxacin	392.41	8.904 ± 0.055 6.403 ± 0.020	$14200 \pm 1700 \mu\text{M}$ $5600 \pm 670 \mu\text{g/ml}$	$425 \pm 35 \mu\text{M}$ $167 \pm 14 \mu\text{g/ml}$
	Sulfacetamide	214.24	5.318 ± 0.005 1.40 ± 0.17	$46600 \pm 6700 \mu\text{M}$ $10000 \pm 1400 \mu\text{g/ml}$	$30200 \pm 240 \mu\text{M}$ $6470 \pm 50 \mu\text{g/ml}$
	Sulfamethazine	278.33	7.64 ± 0.05 2.450 ± 0.002 1.97 ± 0.02	$24000 \pm 10000 \mu\text{M}$ $6700 \pm 2800 \mu\text{g/ml}$	$1852 \pm 23 \mu\text{M}$ $508 \pm 4 \mu\text{g/ml}$

Table 1. Continued

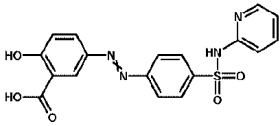
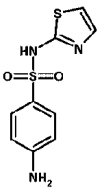
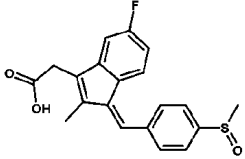
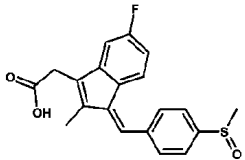
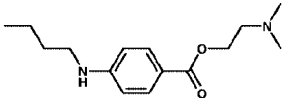
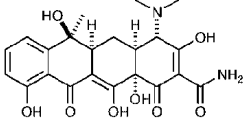
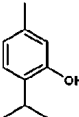
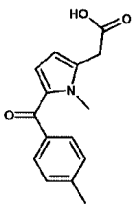
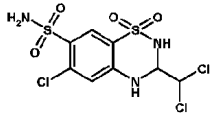
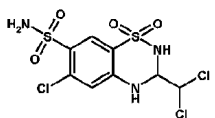
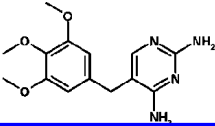
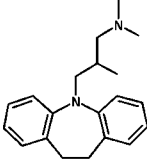
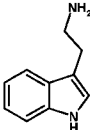
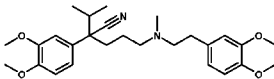
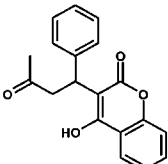
Structure	Name	MW Neutral Form	pKa	Kinetic Solubility μM $\mu\text{g/ml}$	Intrinsic Solubility μM $\mu\text{g/ml}$
	Sulfasalazine	398.40	2.51 ± 0.13 8.11 ± 0.15 11.0 ± 0.3	$34 \pm 2 \mu\text{M}$ $14 \pm 1 \mu\text{g/ml}$	$0.73 \pm 0.04 \mu\text{M}$ $0.29 \pm 0.02 \mu\text{g/ml}$
	Sulfathiazole	255.32	2.11 ± 0.01 7.14 ± 0.01	$2400 \pm 480 \mu\text{M}$ $610 \pm 120 \mu\text{g/ml}$	$2050 \pm 180 \mu\text{M}$ $522 \pm 45 \mu\text{g/ml}$
	Sulindac FORM I	356.42	4.08 ± 0.01	a	$210 \pm 23 \mu\text{M}$ $75 \pm 8 \mu\text{g/ml}$
	Sulindac FORM II	356.42	4.08 ± 0.01	$230 \pm 30 \mu\text{M}$ $82 \pm 11 \mu\text{g/ml}$	$31 \pm 4 \mu\text{M}$ $11 \pm 1 \mu\text{g/ml}$
	Tetracaine	264.37	8.49 ± 0.01 2.39 ± 0.02	$1071 \pm 24 \mu\text{M}$ $283 \pm 6 \mu\text{g/ml}$	$975 \pm 22 \mu\text{M}$ $258 \pm 6 \mu\text{g/ml}$
	Tetracycline	444.45	9.149 ± 0.005 7.506 ± 0.005 3.424 ± 0.005	$25400 \pm 7700 \mu\text{M}$ $11300 \pm 3400 \mu\text{g/ml}$	$1190 \pm 190 \mu\text{M}$ $527 \pm 84 \mu\text{g/ml}$
	Thymol	150.22	10.344 ± 0.01	$9000 \pm 140 \mu\text{M}$ $1350 \pm 21 \mu\text{g/ml}$	$6520 \pm 38 \mu\text{M}$ $979 \pm 6 \mu\text{g/ml}$
	Tolmetin	257.29	3.522 ± 0.017	$395 \pm 95 \mu\text{M}$ $102 \pm 24 \mu\text{g/ml}$	$81 \pm 4 \mu\text{M}$ $21 \pm 1 \mu\text{g/ml}$
	Trichlormethiazide FORM I	380.66	9.63 ± 0.02 6.69 ± 0.01	a	$661 \pm 17 \mu\text{M}$ $251 \pm 7 \mu\text{g/ml}$
	Trichlormethiazide FORM II	380.66	9.63 ± 0.02 6.69 ± 0.01	$4500 \pm 2000 \mu\text{M}$ $1720 \pm 760 \mu\text{g/ml}$	$296 \pm 5 \mu\text{M}$ $113 \pm 2 \mu\text{g/ml}$
	Trimethoprim	290.32	7.21 ± 0.01	$5800 \pm 2200 \mu\text{M}$ $1700 \pm 640 \mu\text{g/ml}$	$1119 \pm 96 \mu\text{M}$ $325 \pm 28 \mu\text{g/ml}$

Table 1. Continued

Structure	Name	MW Neutral Form	pKa	Kinetic Solubility	Intrinsic Solubility
				μM $\mu\text{g/ml}$	μM $\mu\text{g/ml}$
	Trimipramine	294.44	9.34 ± 0.22	$17.3 \pm 2.1 \mu\text{M}$	$16 \pm 3 \mu\text{M}$
				$5.1 \pm 0.6 \mu\text{g/ml}$	$4.8 \pm 0.9 \mu\text{g/ml}$
	Tryptamine	160.22	10.108 ± 0.002	$21000 \pm 6400 \mu\text{M}$	$500 \pm 130 \mu\text{M}$
				$3400 \pm 1000 \mu\text{g/ml}$	$80 \pm 21 \mu\text{g/ml}$
	Verapamil	454.61	8.72 ± 0.07	$105 \pm 3 \mu\text{M}$	$106 \pm 4 \mu\text{M}$
				$48 \pm 2 \mu\text{g/ml}$	$48.1 \pm 1.7 \mu\text{g/ml}$
	Warfarin	308.34	4.94 ± 0.01	$408 \pm 19 \mu\text{M}$	$16.5 \pm 0.3 \mu\text{M}$
				$125 \pm 6 \mu\text{g/ml}$	$5.1 \pm 0.1 \mu\text{g/ml}$

^a The kinetic solubility value is the average between all polymorphic forms. ^b Solubility measured using the cosolvent technique; kinetic solubility is not available. Details of the experiments are in the Supporting Information. ^c Only one experiment performed; no standard deviation available. The error given is the coefficient of variation (CV). ^d Compound too soluble to measure. ^e Only one kinetic value considered.

Data extracted directly from the experiments (solubility, temperature, ionic strength, and number of assays) can be found in the Supporting Information. These data were checked by using a *Java* program to analyze the original data which was recorded as the experiment progressed. The solubilities and recalculated means and standard deviations from this process are included in the Supporting Information. The two methods of gathering the data lead to the same results in all cases.

RESULTS AND DISCUSSION

The thermodynamic solubility of the neutral form of the drugs used in this study was determined by the Chasing Equilibrium method, which has been described in detail elsewhere.^{15,21,29}

Determination of the pK_a . The determination of solubility by CheqSol relies on the knowledge of the pK_a values of the compounds studied. Since an error of one pK_a unit can generate an error of an order of magnitude in the solubility result, very accurate measurement of the pK_a is needed. We routinely measure the pK_a of all the compounds multiple times before measuring the solubility, using exactly the same conditions (temperature and ionic strength) as will be used for the solubility measurement.

Determination of the Kinetic Solubility, S_{kin} . An accurately measured amount of compound is placed in the titration vessel, and a measured volume (usually 10 or 15 mL) of distilled water adjusted to an ionic strength of 0.15 M by use of KCl is added. The solution is then brought to

a basic pH (if the sample is an acid) or to an acidic one (if the sample is a base) by precisely controlled addition of KOH or HCl solutions and kept at that initial pH until total dissolution has been accomplished. The solution is then back-titrated by adding small aliquots of acid or base until precipitation is detected by the use of a spectroscopic dip probe that is positioned inside the titration vessel and records the UV-vis spectra of the solution during the whole experiment. When precipitation occurs there is a large increase in the amount of light absorbed by the solution, indicating that the neutral form of the drug has precipitated. At this point the kinetic solubility is calculated: the values of concentration, the pK_a of the substrate, and the pH of the solution leading up to the precipitation event are used to determine the maximum amount of free-acid or free-base form that existed in solution. Kinetic solubility values are not thermodynamic measurements, and the precipitates formed at this point are kinetically driven. Repeated experiments often lead to a large standard deviation around the mean value for kinetic solubility. When isolation of these first precipitates was possible, solid-state studies show that these precipitates were, in most cases, a different polymorphic form to the precipitates isolated at equilibrium. We observed that these kinetic solubility values are sometimes close to the solubility values given in the literature, suggesting that these literature solubility values are kinetic and not thermodynamic.

Determination of the Intrinsic Solubility, S_0 . After the kinetic precipitation point, the experiment continues with the

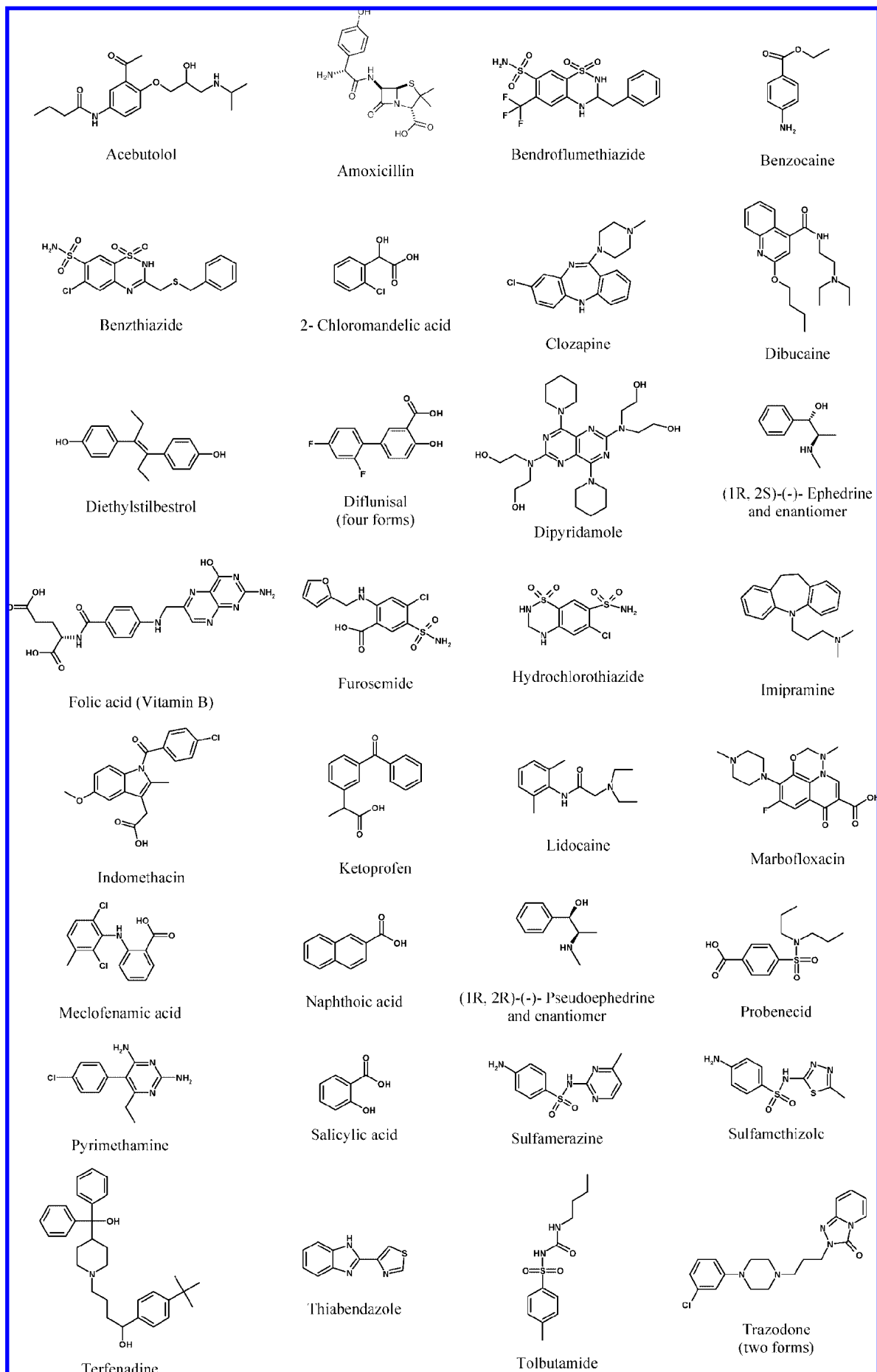


Figure 1. Solubility challenge: predict the intrinsic solubilities of these molecules.

addition of a small aliquot of titrant to produce extra precipitation. If the solution is still supersaturated at this point, then material will be coming out of solution. If it is subsaturated, then some of the precipitate will be redissolving. In either case, the pH of the solution will be changing, and the change in pH with respect to time is monitored. A negative pH-gradient indicates that either the precipitated neutral form of an acid is dissolving or else the un-ionized form of a base is precipitating. A positive pH-gradient means that either the solution is supersaturated and that the neutral form of the acid is precipitating or else the un-ionized form of a base is dissolving into a subsaturated solution. The pK_a of the substrate, which was measured at the start of the procedure, allows us to distinguish between acids and bases. Therefore, the direction of the pH-gradient indicates if the solution is supersaturated (precipitation occurring) or subsaturated (dissolving). Instead of waiting for the system to reach the equilibrium by itself, which can take from minutes to weeks, the CheqSol approach changes the solution from supersaturated to subsaturated and back again by adding tiny amounts of titrant. A change in pH of less than 0.05 is usually sufficient to reverse the direction of the pH-gradient. The zero points of the pH-gradient indicate that the system is at equilibrium, when neither precipitation nor dissolution is observed. The exact pH where the pH-gradient is zero is calculated and used with the mass-balance and charge-balance equations to calculate the intrinsic solubility. Each cycle between a supersaturated solution and a subsaturated one will produce a measure of the intrinsic solubility. Sixteen cycles are normally enough to reach a good determination of the equilibrium point. Usually the intrinsic solubility values are distributed in a tight group around the average value. If these intrinsic solubility values are not randomly distributed and show a drift in either direction, then this means that the equilibrium has not been reached, and the number of cycles is increased until the equilibrium condition is achieved. The spread of the crossing points is used to determine the coefficient of variation (CV)¹⁵ of the mean intrinsic solubility result. Experiments showing a CV higher than 4% were not used in the final calculation of the intrinsic solubility value. The final intrinsic solubility value is calculated from an average of several separate experiments with 16 cycles per experiment.

Polymorphism. Solid-state characterization of the final precipitate is useful because the intrinsic solubility depends on the crystalline form of the solid in equilibrium with the solution. The solid characterization of all the molecules we report here is currently being investigated in our laboratory. A recent survey showed that the difference in solubility of polymorphs is usually less than a factor of 2,²² but our results suggest that polymorphs may have more diverse solubility than this. For example, form I of sulindac is about seven times more soluble than form II.²¹

Ostwald's rule³⁰ suggests that when more than one polymorph is accessible, the less stable ones, and, therefore, the most soluble ones, will come out of solution first, driven by kinetics and not by thermodynamics. If several forms of one compound are reported in Table 1, then the form with the smallest solubility is the one at equilibrium and probably the form of greatest importance.

Very Soluble Compounds. Some compounds are too soluble to measure because of the physical limitations of the

equipment. The most straightforward solution is to increase the amount of the compound and to decrease the volume of initial solution. There is a limit to this process, because the use of more compound means that more titrant is needed. The experiment will be automatically stopped after reaching a total volume of 25 mL, when the titration vessel overflows. In these cases we report a lower limit for the solubility.

Very Insoluble Compounds. If the neutral form of a compound has an intrinsic solubility so low that it is impossible to fully dissolve it even at extreme pH conditions, then the use of cosolvents may lead to a good measurement of solubility. This technique can also be helpful for compounds that form insoluble salts, aggregates, clusters, or micelles. At least six experiments are performed in different ratios of water to cosolvent and then extrapolated to 100% water to obtain the aqueous intrinsic solubility. The pK_a of the compounds must also be measured at different ratios of water to cosolvent for these solubility experiments. The procedure only works if these measurements give sufficient linearity to give confidence that both the pK_a extrapolation and the solubility extrapolation are reasonable processes. These data are recorded in the Supporting Information, and the linearity of the plots is excellent in all cases reported here. The cosolvent technique is time-consuming but showed excellent results when it was required.

Solubility Data. The molecules in Table 1 were selected as a diverse set of available pharmaceutical ingredients. The molecules met the following criteria: (a) an ionizable group with a pK_a between 1 and 13; (b) available and affordable; and (c) safe.

In some cases, such as the quinolone derivatives, we were interested in the properties of a particular subclass of molecules, which we included even though they did not increase the diversity of the data set. For some molecules, it was not possible to measure an experimental solubility, because hydrolysis occurred or the molecule was found to be too soluble resulting in overflow of the vial. The molecular weights range from 115 (proline) to 645 (amiodarone), and their pK_a range from 1 to more than 12. The intrinsic solubilities range over 7 orders of magnitude from 86 mM (acetaminophen) to 6.7 nM (amiodarone), with a reasonably even spread of intermediate values, and a few which are too large for us to determine precisely. With these data available, it should be possible to gain a better understanding of solubility phenomena.

Competition. A 32 molecule test set has been randomly selected from our data of druglike molecules and is shown in Figure 1. Solubilities have been reported in the literature for some of these compounds, but the reported values are not always in close agreement with our measured values. The data are also available in machine-readable form on our Web site: <http://www-jmg.ch.cam.ac.uk/data/solubility/>.

The deadline for submission of predictions of intrinsic solubilities is Sept 15, 2008. The list of compounds with their predicted intrinsic solubilities and an outline of the method used to calculate them should be e-mailed to Anton J. Hopfinger at jcim@unm.edu by Sept 15, 2008. The results will be assessed by a panel, and the groups making the best submissions will be invited to submit papers to this journal describing their models for solubility.

CONCLUSION

We have measured the intrinsic solubility of 100 pharmaceutical drugs at 25 ± 0.2 °C and ionic strength of 0.15 M by a potentiometric technique using the CheqSol approach which ensures that thermodynamic equilibrium is reached. The high accuracy of our data is achieved by rejecting those results with a coefficient of variation (CV) higher than 4% and repeating each experiment several times which makes a statistical treatment possible. We now challenge researchers to predict the intrinsic solubility of 32 other druglike molecules.

ACKNOWLEDGMENT

We thank Pfizer for financial support and helpful suggestions for this project. Sirius Analytical Instruments has provided useful guidance on the use of the CheqSol machine. We specially would like to thank Karl Box for his useful discussions, comments, and suggestions throughout this entire project. We also thank the Editors of the *Journal of Chemical Information and Modeling* and in particular Dr. William Jorgensen, for assistance in setting up this competition.

Supporting Information Available: Details of the results, temperature, ionic strength, and number of assays for each compound and plots for the compounds measured using the cosolvent method. This material is available free of charge via the Internet at <http://pubs.acs.org>.

REFERENCES AND NOTES

- (1) Olcer, M.; Hakyemez, G. Investigations of some physicochemical properties of haloperidol which may affect its activity. *J. Clin. Pharm. Ther.* **1988**, *13*, 341–349.
- (2) Gebbia, V.; Puozzo, C. Oral versus intravenous vinorelbine: clinical safety profile. *Expert Opin. Drug Saf.* **2005**, *4*, 915–28.
- (3) Stark, A. N.; Jackson, G.; Carey, P. J.; Arfeen, S.; Proctor, S. J. Severe Renal Toxicity Due to Intermediate-Dose Methotrexate. *Cancer Chemother. Pharmacol.* **1989**, *24*, 243–245.
- (4) Kennedy, T. Managing the drug discovery/development interface. *Drug Discovery Today* **1997**, *2*, 436–444.
- (5) Palmer, A. M. New horizons in drug metabolism, pharmacokinetics and drug discovery. *Drug News Perspect.* **2003**, *16*, 57–62.
- (6) Paolini, G. V.; Shapland, R. H. B.; van Hoorn, W. P.; Mason, J. S.; Hopkins, A. L. Global mapping of pharmacological space. *Nat. Biotechnol.* **2006**, *24*, 805–815.
- (7) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Delivery Rev.* **2001**, *46*, 3–26.
- (8) Martin, Y. C. A bioavailability score. *J. Med. Chem.* **2005**, *48*, 3164–3170.
- (9) Dollery, C. *Therapeutic Drugs*, 2nd ed.; Churchill Livingstone: 1999.
- (10) Dearden, J. C. In silico prediction of aqueous solubility. *Expert Opin. Drug Discovery* **2006**, *1*, 31–52.
- (11) Yan, A. X.; Gasteiger, J.; Krug, M.; Anzali, S. Linear and nonlinear functions on modeling of aqueous solubility of organic compounds by two structure representation methods. *J. Comput.-Aided Mol. Des.* **2004**, *18*, 75–87.
- (12) Horter, D.; Dressman, J. B. Influence of physicochemical properties on dissolution of drugs in the gastrointestinal tract. *Adv. Drug Delivery Rev.* **2001**, *46*, 75–87.
- (13) Chiarini, A.; Tartarini, A.; Fini, A. pH-Solubility Relationship and Partition-Coefficients for Some Anti-Inflammatory Arylaliphatic Acids. *Arch. Pharm.* **1984**, *317*, 268–273.
- (14) Fini, A.; Laus, M.; Orienti, I.; Zecchi, V. Dissolution and Partition Thermodynamic Functions of Some Nonsteroidal Antiinflammatory Drugs. *J. Pharm. Sci.* **1986**, *75*, 23–25.
- (15) Stuart, M.; Box, K. Chasing equilibrium: Measuring the intrinsic solubility of weak acids and bases. *Anal. Chem.* **2005**, *77*, 983–990.
- (16) Avdeef, A.; Berger, C. M.; Brownell, C. pH-metric solubility. 2: Correlation between the acid-base titration and the saturation shake-flask solubility-pH methods. *Pharm. Res.* **2000**, *17*, 85–89.
- (17) Fini, A.; Fazio, G.; Feroci, G. Solubility and solubilization properties of non-steroidal anti-inflammatory drugs. *Int. J. Pharm.* **1995**, *126*, 95–102.
- (18) Jorgensen, W. L.; Duffy, E. M. Prediction of drug solubility from structure. *Adv. Drug Delivery Rev.* **2002**, *54*, 355–66.
- (19) Loftsson, T.; Hreinsdottir, D. Determination of aqueous solubility by heating and equilibration: A technical note. *AAPS PharmSciTech* **2006**, *7*, 4.
- (20) Huang, L. F.; Tong, W. Q. Impact of solid state properties on developability assessment of drug candidates. *Adv. Drug Delivery Rev.* **2004**, *56*, 321–334.
- (21) Llinàs, A.; Box, K. J.; Burley, J. C.; Glen, R. C.; Goodman, J. M. A new method for the reproducible generation of polymorphs: two forms of sulindac with very different solubilities. *J. Appl. Crystallogr.* **2007**, *40*, 379–381.
- (22) Pudipeddi, M.; Serajuddin, A. T. M. Trends in solubility of polymorphs. *J. Pharm. Sci.* **2005**, *94*, 929–939.
- (23) Sanghvi, T.; Jain, N.; Yang, G.; Yalkowsky, S. H. Estimation of aqueous solubility by the General Solubility Equation (GSE) the easy way. *QSAR Comb. Sci.* **2003**, *22*, 258–262.
- (24) Johnson, S. R.; Chen, X. Q.; Murphy, D.; Gudmundsson, O. A computational model for the prediction of aqueous solubility that includes crystal packing, intrinsic solubility, and ionization effects. *Mol. Pharm.* **2007**, *4*, 513–523.
- (25) Ouvrard, C.; Mitchell, J. B. O. Can we predict lattice energy from molecular structure. *Acta Crystallogr. Sect. B: Struct. Sci.* **2003**, *59*, 676–685.
- (26) Datta, S.; Grant, D. J. W. Crystal structures of drugs: Advances in determination, prediction and engineering. *Nat. Rev. Drug Discovery* **2004**, *3*, 42–57.
- (27) Johnson, J. L. H.; Yalkowsky, S. H. Two new parameters for predicting the entropy of melting: Eccentricity (epsilon) and spirality (mu). *Ind. Eng. Chem. Res.* **2005**, *44*, 7559–7566.
- (28) Palmer, D. S.; Llinàs, A.; Morao, I.; Day, G. M.; Goodman, J. M.; Glen, R. C.; Mitchell, J. B. O. Predicting intrinsic aqueous solubility by a thermodynamic cycle. *Mol. Pharm.* **2008**, *5*, 266–279.
- (29) Llinàs, A.; Burley, J. C.; Box, K. J.; Glen, R. C.; Goodman, J. M. Diclofenac solubility: Independent determination of the intrinsic solubility of three crystal forms. *J. Med. Chem.* **2007**, *50*, 979–983.
- (30) Threlfall, T. Structural and thermodynamic explanations of Ostwald's rule. *Org. Process Res. Dev.* **2003**, *7*, 1017–1027.

CI800058V