Predicting the prevalence of alternative Warfarin tautomers in solution

Alpeshkumar K. Malde,^{1,} Martin Stroet,¹ Bertrand Caron¹ Koen M. Visscher,² and Alan E.
Mark^{1,3,*}*

¹School of Chemistry and Molecular Biosciences, The University of Queensland, Brisbane,
QLD 4072, Australia.

²Division of Molecular Toxicology, VU University Amsterdam, The Netherlands.

³Institute for Molecular Bioscience, The University of Queensland, Brisbane, QLD 4072,
Australia.

KEYWORDS: Tautomerism, Warfarin, Free Energy Calculations, Free Enthalpy of
Formation, Free Enthalpy of Solvation

ABSTRACT

Warfarin, a widely used oral anticoagulant, is prescribed as a racemic mixture. Each enantiomer of neutral Warfarin can exist in 20 possible tautomeric states leading to complex pharmacokinetics and uncertainty as to the relevant species under different conditions. Here the ability of alternative computational approaches to predict the preferred tautomeric form(s) of neutral Warfarin in different solvents is examined. It is shown that varying the method used to estimate the heat of formation in vacuum (direct or via homodesmic reactions), whether vibrational and/or conformational entropic corrections in vacuum were included and the method used to estimate the free enthalpy of solvation (i.e. PCM, COSMO or SMD implicit models or explicit solvent) lead to large differences in the predicted rank and relative populations of the tautomers. In this case only a combination of the free enthalpy of formation using homodesmic reactions and explicit solvent to estimate the free enthalpy of solvation yielded results compatible with the available experimental data. The work also suggests that a small but significant subset of the possible Warfarin tautomers are likely to be physiologically relevant.

Introduction

Detailed knowledge of the structure and thermodynamics of ligand binding at an atomic level is critical for modern drug discovery. This requires knowing not only the thermodynamically preferred conformation and relative orientation of the ligand when bound to the therapeutic target but also its preferred protonation and/or tautomeric state when bound to the target and free in solution or when forming any off target interactions.¹⁻⁴ Consideration of possible alternative tautomeric states (isomers of a molecule that readily interconvert and often associated with changes in the location of hydrogen atoms) is particularly challenging.⁵ The neutral form of the imidazole ring of the amino acid histidine (where the hydrogen may be attached to either of the two nitrogen atoms) is a simple example. Other cases are much more complex.

The anticoagulant Warfarin is a case in point. Warfarin is prescribed as a racemic mixture containing *R* and *S* enantiomers. As illustrated in Figure 1 each enantiomer can exist in one of 2 protonation states. The pKa of Warfarin is ~ 5 ⁶ meaning that in aqueous solution at neutral pH it will essentially all be in the anionic (basic) form. However, under mildly acidic conditions or in non-aqueous (or aprotic) environments such as when interacting with a cell membrane or bound within a hydrophobic pocket of a transport protein, Warfarin will predominantly be in its neutral form. The neutral form of Warfarin is capable of undergoing chemical rearrangement. As discussed in detail by Porter⁷ and references therein, in theory this can lead to 20 possible tautomeric states for each enantiomer (Figure 2). This includes 16 open-chain tautomers (4 x 4-hydroxy-enol tautomers (W01, W05, W06, W13), 4 x 2-hydroxy-enol tautomers (W02, W07, W08, W14), 8 x diketo tautomers (W03, W04, W09, W10, W11, W12, W15, W16) as well as 4 cyclic hemiketal tautomers (W17 to W20). This means that while only one of these species is likely to bind to vitamin K epoxide reductase (VKOR),⁸⁻⁹ the therapeutic

target, there are potentially 42 different chemical moieties that can form off target interactions. For example, Warfarin is known to bind to Human Serum Albumin and be degraded by specific Cytochrome P450s but the form in which Warfarin is transported and/or degraded is not known.⁷ Not unexpectedly, Warfarin shows complex pharmacokinetics interacting with numerous environmental factors, dietary components and drugs. As a consequence, patients on Warfarin therapy require constant monitoring. The question is: Which is the predominant tautomeric form of a ligand (Warfarin in the present study) under a given set of conditions? A number of crystal structures of Warfarin in isolation have been reported. These include structures of the neutral hemiketal tautomers (*R,R* and *S,S* diastereomers; Cambridge Structural Database (CSD) Ids WARFIN¹⁰ and WARFAR10¹¹ respectively) as well as the sodium salt of racemic Warfarin (CSD Ids ZZZKXG¹² and EFIWIZ¹³). Nuclear Magnetic Resonance (NMR) studies of Warfarin in aprotic solvents including CDCl₃ and DMSO-*d*₆ have suggested the presence of a mixture between cyclic hemiketal tautomers (major) and open-chain tautomers (minor).¹⁴⁻¹⁶ NMR studies in aqueous solution indicate a preference for a cyclic hemiketal tautomer under acidic conditions while a mixture of open-chain 4-hydroxyl tautomers and cyclic hemiketal tautomers at neutral pH.¹⁷⁻¹⁹ Solid state ¹³C NMR studies performed on a dry solid sample using cross polarization magic angle spinning at 75.5 MHz indicated the presence of cyclic hemiketal tautomers, however the precise configuration was not determined.²⁰ ¹⁷O NMR studies also suggested the cyclic hemiketal was the major tautomer in the solid state.²¹ In contrast, an NMR study of an inclusion complex between Warfarin and β-cyclodextrin was interpreted as *R*- and *S*-Warfarin binding in the anionic open-chain form (Figure 1).²² Other experimental studies involving a combination of NMR, absorption spectroscopy and fluorescence spectroscopy have suggested the simultaneous presence of multiple different tautomers in alternative environments making any interpretation of the available data challenging.²³⁻²⁴

Some structural data on Warfarin bound to different biological targets is also available. Structures of Warfarin complexed with Cytochrome P450 2C9 (1OG5) and Human Serum Albumin (1HA2, 2BXD and 1HZ9) have been deposited in the Protein Data Bank (PDB). In each of these structures Warfarin was modelled in an open-chain form but given the fact that these structures were solved at resolutions between 2.5 to 3.05 Å it is uncertain precisely which tautomeric form is in reality present.²⁵⁻²⁷ Indeed, drug metabolism studies involving Warfarin and a related anti-coagulant Phenprocoumon (which cannot cyclize to form a hemiketal) suggest that Cytochrome P450 2C9,²⁸ more likely binds the cyclic hemiketal tautomer of Warfarin than the open-chain form modelled into the crystal structure (1OG5).

Given the importance of knowing the most probable tautomeric form of a molecule in a given environment coupled with the difficulties in interpreting the available experimental data in cases where multiple alternative species can co-exist, there is much interest in the use of computational approaches to predict the likelihood of finding a given tautomer under different conditions.²⁹ This is challenging as it involves estimating the free energy of formation of a given tautomer in isolation (a fundamentally quantum mechanical property) as well as the free energy associated with placing that tautomer in a given environment as depicted in Figure 3.

Here we examine the ability of a range of computational methods, including quantum mechanical calculations and free energy calculations, to predict the order/ranking of Warfarin tautomers in a given environment as well as the uncertainties and limitations of these methods in this case.

Methods

To understand the preferred tautomeric state of the neutral form of Warfarin in different environments (vacuum, water, dimethyl sulfoxide and hexane), a series of quantum mechanical calculations, molecular dynamics simulations and free energy calculations of the different possible tautomeric forms of the neutral form of *R*-Warfarin depicted in Figure 2, were performed. The enthalpy of formation (ΔH_f) in vacuum for a given tautomer was obtained either directly from quantum mechanical (QM) thermochemistry calculations or indirectly using so-called isodesmic and/or homodesmic reactions. The use of isodesmic and homodesmic reactions, which allow for experimental data to be incorporated into the calculation of the enthalpy of formation, has been proposed to be more reliable than direct calculations in some cases.³⁰ Potential entropic effects in vacuum were estimated based on vibrational as well as conformational contributions at a given temperature. The solvation free energy (ΔG_{solv}) was calculated using alternative implicit solvation models as well as in explicit solvent using molecular dynamics simulation techniques.

Quantum Mechanical Calculations

All quantum mechanical calculations were performed using Gaussian 09 revision D01³¹ accessed via National Computational Infrastructure (NCI), Australia. Geometry optimizations of the Warfarin tautomers (Figure 2) and the molecules used to construct the homodesmic reactions (listed in Table 1) were performed using density functional theory (the B3LYP density functional).³²⁻³⁴ The basis set was 6-311+G(3df,2p). Harmonic vibrational frequencies were computed in order to a) verify the nature of each stationary point (number of imaginary frequencies = 0); b) determine the zero-point vibrational energy and obtain thermal corrections to the enthalpy and, c) to estimate the difference in the entropy between the different tautomers at 298.15 K in vacuum (see below). Where indicated the effect of solvent was mimicked using

a Self-Consistent Reaction Field (SCRF) in combination with alternate implicit solvent models. The solvent models included the Polarizable Continuum Model (PCM)³⁵, the Self-consistent reaction field Model using charge Density (SMD)³⁶ and the COnductor-like Screening MOdel (COSMO)³⁷ in conjunction with the experimental dielectric constants for water (78.39), dimethyl sulfoxide (46.7) and hexane (1.88). The initial geometries of the open-chain tautomers (4-hydroxy-enol and 2-hydroxy-enol tautomers, Figure 2) were based on the crystal structure of the Na salt of the Warfarin anion (CSD Id EFIWIZ¹³). The diketo tautomers were based on the corresponding enol tautomers. The structures of the cyclic hemiketal tautomers (W17 to W20) were based on the CSD structure WARFIN¹⁰.

The enthalpy of formation (ΔH_f) of the Warfarin tautomers (W01 to W20) were estimated using QM calculations at the B3LYP/6-311+G(3df,2p) level of theory in vacuum. Two approaches were compared. In the first approach, the final enthalpy values were obtained by adding a thermal correction (incorporating zero-point vibrational energy) to the total electronic energy (Figure 3). In the second approach, the enthalpy of formation was estimated using homodesmotic reactions. A few representative homodesmotic reaction schemes are provided as supplementary material (Figure S1). The homodesmotic reactions were constructed depending upon the availability of experimental thermochemical data for the relevant compounds. The enthalpy of formation was calculated by taking into account the computed enthalpies of homodesmotic reaction described by the equations given in Table 2 and the experimental enthalpies of formation in the gas phase of all molecules involved (Table 1) other than the given Warfarin tautomer.

The free enthalpy of formation or Gibbs free energy of formation (ΔG_f) was calculated as $\Delta G_f = \Delta H_f - T\Delta S$. The entropy (S) for each molecule was estimated by combining the

vibrational entropy (S_{B3LYP}) calculated using a single conformation at the B3LYP/6-311+G(3df,2p) level of theory with an estimate of the conformational entropy (S_{conf}):

$$S \approx S_{B3LYP} + S_{conf} \quad (1)$$

The conformational entropy which can be significant for large, flexible molecules, such as Warfarin was estimated as:

$$S_{conf} = R \ln(\Omega) \quad (2)$$

where Ω is the number of unique conformations (states) and R is the gas constant.

Molecular Dynamics Simulations (Solvation Free Enthalpy Calculations)

All molecular dynamics simulations were performed using the GROMOS 11³⁸⁻⁴⁰ simulation package. Force field parameters for Warfarin tautomers were obtained from the ATB (Automated Topology Builder, <https://atb.uq.edu.au>).⁴¹⁻⁴³ Van der Waals and bonded parameters (bonds, angles, improper dihedrals and torsions) were based on the GROMOS 54A7⁴⁴⁻⁴⁵ force field. The partial atomic charges were based on fitting to the electrostatic potential (ESP) using the Kollmann-Singh⁴⁶ scheme at the B3LYP/6-31G* level of theory. A major challenge in this study is that multiple different tautomeric forms of the same molecule are considered. This means that even slight differences in the parameters assigned to chemically equivalent atoms could lead to artifacts. It is well known that the assignment of partial atomic charges using ESP fitting can be sensitive to changes in local geometry. While the ATB algorithm minimizes such effects by enforcing intramolecular symmetry, small variations between equivalent groups in different molecules can still arise. To ensure the charges assigned to equivalent atoms are consistent throughout the entire set of molecules, a multi-molecule ESP fitting procedure was used.⁴⁷ This involved fitting the partial charges of

all similar/identical groups across all 20 tautomers of R-Warfarin (W01 to W20) simultaneously using the second-neighbour chemically equivalent symmetry groups shown in Figure 4.

Before commencing the simulations, each Warfarin tautomer was placed in a rectangular periodic box and solvated with either ~1600 simple point charge (SPC)⁴⁸ water molecules or ~1000 dimethyl sulfoxide (DMSO)⁴⁹ molecules or ~550 hexane⁴⁴ molecules. The structure was energy minimized and the system equilibrated for 200 ps with the heavy atoms of the solute positionally restrained before a series of unrestrained molecular dynamics simulations was commenced. All simulations were performed at constant temperature (298 K) and pressure (1 atm) using a Berendsen thermostat (coupling time of 0.1 ps) and barostat (coupling time of 1.0 ps and isothermal compressibility of 4.575×10^{-4} (kJ/mol/nm³)⁻¹).⁵⁰ A triple-range cutoff was used. Interactions within a shorter-range cutoff of 0.8 nm were updated every step (0.002 ps). Interactions within the longer-range cutoff of 1.4 nm were updated every 0.010 ps (5 steps) together with the pairlist. To correct for the truncation of electrostatic interactions beyond the 1.4 nm long-range cutoff a reaction-field correction was applied using a dielectric permittivity of 61 for water,⁵¹ 38 for DMSO⁴⁹ and 2 for hexane. Note, the experimental dielectric permittivity of water (78) and DMSO (47) were used for the quantum mechanical calculations. However, the dielectric permittivity values used during the molecular dynamics simulations were those of the SPC water and DMSO models. The equations of motion were integrated using the leapfrog scheme and a time step of 2 fs. Initial velocities at a given temperature were taken from a Maxwell-Boltzmann distribution. All bonds were constrained using the SHAKE⁵² algorithm with a geometric tolerance of 0.0001.

Solvation free enthalpies were calculated using the thermodynamic integration (TI) in which the difference in free enthalpy between two states of a system *A* and *B* is expressed as:

$$\Delta G_{AB} = \int_{\lambda_A}^{\lambda_B} \left\langle \frac{\partial V(\mathbf{r})}{\partial \lambda} \right\rangle_{\lambda} d\lambda \quad (4)$$

where $V(\mathbf{r})$ is the potential energy of the system as a function of the coordinate vector $\mathbf{r}$ and λ is a parameter that couples the two states A and B .⁵³ In this case the coupling parameter λ was used to scale the inter- and intramolecular non-bonded interactions involving the solute from 0 to 1 (where 0 represents the full interaction and 1 no interaction). To avoid sampling singularities in the potential energy function and in the derivative with respect to λ (as well as numerical instabilities during the simulations) the non-bonded interactions were scaled using the λ -dependent soft-core interaction function of Beutler *et al.* with $\alpha_{LJ} = 0.5$ and $\alpha_{\text{electrostatic}} = 0.5 \text{ nm}^2$.⁵⁴ Note, using the λ -dependent soft-core interaction function of Beutler *et al.* as implemented in GROMOS, there is no requirement or advantage in performing the removal of the charge and Lennard-Jones interactions in separate stages. Eq. (4) was evaluated by calculating the ensemble average of the derivative $\langle \partial V(\mathbf{r}) / \partial \lambda \rangle_{\lambda}$ at a series of discrete λ -values. The values of $\langle \partial V(\mathbf{r}) / \partial \lambda \rangle_{\lambda}$ were then integrated using the Trapezoidal approximation. The solvation free enthalpy (in water, DMSO and hexane) was calculated as:

$$\Delta G_{\text{solv}} = \Delta G_{0 \rightarrow I}(\text{vacuum}) - \Delta G_{0 \rightarrow I}(\text{solvent}) \quad (5)$$

$$= \int_{\lambda_0}^{\lambda_1} \left\langle \frac{\partial V(\mathbf{r})}{\partial \lambda} \right\rangle_{\lambda}^{\text{vac}} d\lambda - \int_{\lambda_0}^{\lambda_1} \left\langle \frac{\partial V(\mathbf{r})}{\partial \lambda} \right\rangle_{\lambda}^{\text{solvent}} d\lambda \quad (6)$$

$$= \int_{\lambda_0}^{\lambda_1} \left[\left\langle \frac{\partial V(\mathbf{r})}{\partial \lambda} \right\rangle_{\lambda}^{\text{vac}} - \left\langle \frac{\partial V(\mathbf{r})}{\partial \lambda} \right\rangle_{\lambda}^{\text{solvent}} \right] d\lambda \quad (7)$$

Note that for the combined integral in Eq. (6) to hold as written, the same λ -values must be sampled in solvent and in vacuum. Initially the value of $\langle \partial V(\mathbf{r}) / \partial \lambda \rangle_{\lambda}$ at 11 equally spaced points between $\lambda = 0$ and $\lambda = 1$ was determined to obtain a first estimate of shape of the underling curve. An automated protocol based on the analysis of errors was used to add more

λ -points and extend the simulations at each λ -point.⁴³ The target error of 1.0 kJ/mol was used to calculate the solvation free enthalpies.

The vacuum systems were generated from a given configuration in the corresponding solvent by simply deleting all solvent molecules within the simulation box. Note that during TI calculations, stochastic coupling with a reference temperature of 298 K and an atomic friction coefficient of 1 ps⁻¹ was applied to prevent thermal decoupling of solvent and solute components of the system.

Results and Discussion

Warfarin continues to be widely-used as an anti-coagulant. This is despite the fact that when taken orally as a racemic mixture Warfarin can in principle give rise to 42 unique chemical entities not including potential metabolites. Below we assess the ability of different computational approaches to predict the preferred tautomeric form for Warfarin in different environments. Note, all calculations have been performed using *R*-Warfarin. The results for *S*-Warfarin would be identical in vacuum as well as in non-chiral solvents including water, DMSO and hexane.

Heat of formation:

The most basic means to predict the preferred tautomeric state of a molecule is to compare the standard enthalpy of formation (ΔH_f) calculated quantum mechanically for the different species in vacuum (gas-phase). Warfarin contains 22 non-hydrogen atoms. As the size of the molecule can affect the accuracy of ΔH_f calculations, two alternate approaches were used to estimate the relative enthalpy of formation ($\Delta\Delta H_f$) of 20 tautomers of *R*-Warfarin. Either ΔH_f was calculated directly (see Methods) or calculated indirectly using the experimental heats of formation of more simple compounds and homodesmic reactions to transform these compounds into a given tautomer of Warfarin. All calculations were performed at the B3LYP/6-311+G(3df,2p) level of theory. An isodesmic reaction is a hypothetical chemical reaction in which the nature of chemical bonds broken in the reactant are the same as those formed in the product. A homodesmic reaction is a special case of an isodesmic reaction, where the orbital hybridization is also taken into account and there is no change in the number of carbon to hydrogen bonds. The advantage of using isodesmic or homodesmic reactions is the possibility for the cancellation of systematic errors. This makes isodesmic or homodesmic

reactions a computationally efficient way to calculate the enthalpy of formation for relatively large molecules such as Warfarin with high accuracy.

The relative enthalpy of formation ($\Delta\Delta H_f$) of 20 tautomers obtained using the B3LYP/6-311+G(3df,2p) level of theory directly and by combining the results from a series of homodesmotic reactions are given in Table 3 (columns 2 and 3). Using the direct approach at this level of theory the open-chain tautomer W01 is predicted to be the preferred tautomer followed by the cyclic hemiketal tautomer W17 (+8.5 kJ/mol). Using homodesmotic reactions with the same underlying theory and basis set the cyclic hemiketal tautomer W17 is predicted to be the most stable followed by its diastereomer W18 (+5.4 kJ/mol). Using homodesmotic reactions the enthalpy of the open-chain tautomer W01 is 6.3 kJ/mol higher than W17. To place these numbers in perspective, at 298.15K a difference of 5.7 kJ/mol in the free energy of formation corresponds to approximately an order of magnitude difference in the relative probability of finding a given tautomer. Considering just the heat of formation the direct approach would predict ~96% would be W01, ~3% would be W17 and each of the remaining species much less than 1%. In contrast using homodesmotic reactions and this level of theory one would predict ~84% W17, ~9% W18 and ~6% W01.

Free Energy of Formation:

Formally, one should not simply consider the heat of formation but rather the free enthalpy (or Gibbs's energy) of formation (ΔG_f). This is especially important in cases where the structures of the tautomers vary significantly such as in Warfarin. Potential entropic contributions in vacuum were estimated as a combination of a vibrational contribution ($ZPVE$) derived from the Hessian determined at the B3LYP/6-311+G(3df,2p) level of theory (Table 3 column 4) and a conformational contribution (S_{conf}) calculated as described in the methods (Table 3 column 5). Note, the exact determination of the absolute conformational entropy would require

knowledge of all conformational states accessible to each system and their relative energies at 298.15 K. Here the relative contribution of the conformational entropy has been estimated assuming the free phenyl group can adopt two energetically equivalent conformations and the pyran ring can also adopt two equivalent conformations leading to 4 possible states for cyclic hemiketal (W17, W18, W19, W20) tautomers. In the case of the diketo tautomers it is assumed that the phenyl ring as well as the acetone(enol) group each can adopt three energetically equivalent conformations leading to 9 possible states. It is further assumed that the 4-hydroxy-enol and 2-hydroxy-enol tautomers can adopt an additional two energetically equivalent states by rotation around the bond between the chiral carbon and benzopyran ring leading to 18 possible states. The overall entropic contribution at room temperature (298.15K) varied by up to 19.7 kJ/mol between species with W18 having the lowest entropy and W06 the highest. This is sufficient to shift the relative probability of finding a given species by up to three-orders of magnitude. The conformational contributions to the entropy are small compared to the vibrational contributions.

The relative standard free enthalpy of formation of the 20 tautomers derived from the direct and homodesmotic calculations are given in columns 6 and 7 of Table 3 respectively. In the case of the direct method, the inclusion of the correction for entropy does not alter the prediction that W01 is the preferred tautomer but lowers the relative probability of finding any other tautomer dramatically. Essentially, the concentrations of all other species would be negligible. In case of the homodesmotic calculations the order of the species changes. W01 is now predicted to be the dominant species (~87%) with W17 making up around 12% of the population. All other species would be negligible.

Effect of solvation:

To distinguish between the alternative tautomers in different environments the effect of solvation must also be considered (Figure 3). Here, three solvents were compared: water, DMSO and hexane. The free enthalpy of solvation was calculated using three alternative implicit solvation models (PCM, COSMO and SMD) as well as using explicit solvent based on the GROMOS force field. The absolute solvation free enthalpy values (in kJ/mol) for the 20 tautomers of *R*-Warfarin considered are given in Table 4. Using both the PCM and COSMO models the differences between the alternative solvents is primarily determined by their respective dielectric constants. As the dielectric constants of water and DMSO differ by less than a factor of 2 (78 as compared to 47) the differences in the predicted solvation free enthalpies in water and DMSO using these models is small (< 1.0 kJ/mol). The range of values in water and DMSO for both the PCM and COSMO models was ~ 15 kJ/mol. W17 was favoured over W01 by ~ 5 kJ/mol. Using the PCM and COSMO models all tautomers are predicted to be much less soluble in hexane than in either water or DMSO (> 20 kJ/mol). The solvation free enthalpy in hexane was systematically ~ 3 -5 kJ/mol lower using COSMO compared to PCM. The range of values was similar in both cases (~ 5 kJ/mol). W17 was favoured over W01 by ~ 2 kJ/mol in both cases. Much greater variation was observed using the SMD method. The tautomers were predicted to be significantly more soluble in all three solvents than when using the PCM or COSMO models. The solvation free enthalpy in DMSO was ~ 6 -24 kJ/mol lower than that in water. The values in hexane were in general similar to, or lower, than in water, in stark contrast to that predicted using either the PCM or COSMO models. The range of solvation free enthalpies was ~ 24 kJ/mol in water, ~ 14 kJ/mol in DMSO and ~ 6 kJ/mol in hexane using the SMD model.

In addition to the use of the three implicit solvation models described above the solvation free enthalpies in water, DMSO and hexane were also calculated using thermodynamic integration and explicit solvent as described in the methods. The values obtained using explicit

solvent (Table 4) are systematically lower than those obtained using either the PCM, COSMO or SMD implicit approaches. The greatest differences were observed in the case of DMSO. Of the three implicit solvation models considered, the results obtained using the SMD model were closest to those obtained in explicit solvent for all three solvents. Not only were the values of the solvation free energy significantly lower in explicit solvent but the range of values in water and in particular DMSO was greater. The range of solvation free energies for the 20 tautomers considered was ~25 kJ/mol for water, ~36 kJ/mol for DMSO and ~6 kJ/mol for hexane. Note, while the general magnitude and the range of values obtained in explicit water and using the SMD model for water are similar, the values of individual tautomers vary. This means that the ranking of the tautomers differed significantly. For example, using the SMD approach solvation favours W17 over W01 by ~9 kJ/mol in water but only ~1 kJ/mol in DMSO. In contrast, using SPC water solvation favours W17 over W01 by 11 kJ/mol but by 18 kJ/mol using the GROMOS DMSO model. Even in hexane, which shows the least variation between the different models, the tautomer predicted to have the most favourable solvation was W06 using PCM and COSMO, W03 using SMD and W19 in explicit hexane.

Comparison to experiment:

To compare to the values derived from the calculations to experiment we must consider the overall process of solvation described in Figure 3. That is to combine the free enthalpy of formation of the isolated molecule in vacuum (ΔG_f) with the solvation free enthalpy (ΔG_{solv}). Using the direct estimates of the free enthalpy of formation together with the PCM, COSMO and SMD solvation models, one would predict that the open-chain tautomer W01 is preferred in water, DMSO and hexane (the concentrations of all other species would be negligible). Estimating the free enthalpy of formation based on homodesmotic reactions combined with either the PCM or COSMO solvation models one would predict that a combination open-chain and

cyclic tautomers would be observed. Specifically, one would predict W17 (~50%), W01 (~40%) and W18 (~8%) in both water or DMSO and W01 (~73%), W17 (~23%) and W18 (~2%) in hexane. Using the homodesmic method and the SMD solvation model one would predict W17 (~73%), W01 (~13%), and W18 (~13%) in water, W01 (~77%), W17 (~20%) and W18 (~2%) in DMSO and W01 (~87%) and W17 (~12%) in hexane. The values of the relative free enthalpies of solution for the PCM, COSMO, SMD models are given in the supplementary material (Tables S1, S2, S3). Table 5 shows the results obtained by combining the direct estimates of the free enthalpy of formation together with the free enthalpy of solvation calculated using explicit solvent one would predict a combination of W01 (~95%) and W17 (~4%) in water, W01 (~65%) and W17 (~34%) in DMSO and W01 (> 99%) in hexane. In contrast using the homodemic method together with the free enthalpy of solvation calculated using explicit solvents one would predict W17 (~93%) and W01 (~6%) in water, W17 (~99%) in DMSO and W01 (~87%) and W17 (~12%) in hexane.

As noted in the introduction, in aqueous solution at neutral pH or higher the open-chain anionic form (Figure 1) is dominant. The neutral forms considered here are found under mildly acidic conditions (pH < 5), in aprotic solvents or in non-aqueous environments such as bound within the hydrophobic pocket of a protein. UV spectroscopy and solution NMR studies suggest that in aqueous solution below pH 5 a mixture of tautomers is found.^{14, 55} The cyclic hemiketal form W17 is the major component while the open-chain W01 is present in minor amounts. Based on solution NMR the cyclic hemiketals W17 (~70%) and W18 (~30%) are the dominant tautomers observed in DMSO-*d*6.¹⁴⁻¹⁵ To the best of our knowledge the preferred tautomeric state of Warfarin in hexane is not known.

Using the direct method to calculate the enthalpy of formation in vacuum the open chain W01 tautomer is predicted to be preferred over the W17 tautomer by 19.6 kJ/mol. This

difference is greater than the difference in the free energy of solvation between the alternative tautomers. As a consequence, W01 is predicted to be the predominant tautomer irrespective of the solvent or solvation model used. This is not in agreement with experiment and suggests that the direct approach, at the chosen level of theory, is not sufficiently accurate to distinguish between the alternative tautomers in this case. Results using the direct approach will not be discussed further.

Using the homodesmotic approach, which has been proposed to be more appropriate for larger molecules such as Warfarin at the chosen level of theory, the open chain W01 tautomer is predicted to be preferred over the cyclic hemiketals W17 and W18 tautomers in vacuum by 4.8 kJ/mol and 11.2 kJ/mol respectively. In this case the PCM and COSMO models predict that the W17 tautomer is the major species in water closely followed by the W01 tautomer. However, the PCM and COSMO models are unable to reproduce the observed differences between water and DMSO. Using the SMD solvation model W17 is predicted to be the dominant species in water followed by W01 in line with experiment. However, using the SMD model W01 is predicted to be the major species in DMSO in clear contrast to experiment. The SMD model also suggests that W18 is present in significant concentrations in water a finding that is also not supported by experiment. The use of explicit water and explicit DMSO yields the best agreement with experiment in regard to the relative populations of the different tautomers. In SPC water, W17 is predicted to be dominant with W01 present as a minor species. Using the GROMOS DMSO model, W17 is predicted to be the predominant species with W01 and W18 making up < 1% of the population. This suggests stability of W18 relative to W17 in this DMSO model is underestimated by ~10 kJ/mol.

As noted above, we do not know the preferred tautomeric state of Warfarin in hexane. However, as hexane is a low dielectric medium it is expected to be closest to vacuum. Indeed,

all of the solvation models predict that W01 is the major species and that W17 is a minor species in hexane.

In addition to the relative populations of specific tautomers in a given environment it is also possible to use partition data to compare the alternative solvation models. The solubility of Warfarin in water at pH 4 is 5 mg/L. The solubility of Warfarin in DMSO is 3 orders of magnitude higher at 5 g/L. This would translate to a difference in solvation free enthalpy of ~ 18 kJ/mol. The hexane/water partition coefficient ($\log P$) has been reported in the range 0.05 - 0.2 at pH 2-4^{56,57} and the solubility of Warfarin in heptane has been reported to be 6.4 mg/L. This suggests that the free enthalpy of solvation of Warfarin in hexane should be 0-3 kJ/mol more favorable than that in water.

In addition to failing to correctly predict the preferred tautomer, the PCM and COSMO models also fail to predict the difference in solubility of Warfarin in water and DMSO. In addition, both models predict Warfarin is much more (by 3 order of magnitude) soluble in water than in hexane. Clearly, these models are incapable of reproducing the experimental trends in this system. Although the SMD model does not predict the preferred tautomer correctly it does predict the increased solubility in DMSO and the fact that the neutral forms of Warfarin are slightly more soluble in hexane than in water. The explicit solvent models overestimate the relative solubility of Warfarin in DMSO as compared to water but correctly predict the relative solubility of Warfarin in water as compared to hexane. We would note in this regard that the parameters generated by the ATB^{41,43} have been extensively validated for use with the SPC water model and the GROMOS united atom hexane model used in this study. The average error in the free enthalpy of solvation in both water and hexane for compounds containing just carbon, hydrogen and oxygen is less than 4 kJ/mol. While the GROMOS

DMSO model reproduces a range of properties of liquid DMSO it has not been extensively validated against solvation data.

Warfarin in other environments:

Clearly, the precise tautomeric form of Warfarin will depend on the local environment. For example, while the open-chain W01 tautomer can be crystallized from a solution containing water and isopropanol, the cyclic hemiketal W17 tautomer is obtained when crystalized from acetone or methanol.^{10-11, 13} Our results would suggest that in a membrane environment the open-chain W01 tautomer would be preferred if fully buried while the cyclic hemiketal W17 tautomer (or the anionic form) may be preferred in the more polar head group region. Predicting which form of Warfarin might bind to a particular biological target is more challenging. Not only are the binding sites within a protein by definition chiral in nature but specific tautomeric forms of Warfarin can form specific pairwise interactions. The affinity of Warfarin for its target protein vitamin K epoxide reductase (VKOR) is 2 μM with the free enthalpy associated with binding being in the order of ~ 32 kJ/mol. The affinity of Warfarin for Human Serum Albumin and cytochrome P450 (CYP) 2C9 which together govern the pharmacokinetics of Warfarin are 10 μM , (~ 29 kJ/mol) and 20 μM ⁵⁸ (~ 27 kJ/mol). Eight neutral Warfarin tautomers (*R* and *S* enantiomers of W01, W02, W17 and W18) are separated by less than 16 kJ/mol (Table 8). This suggests that there is sufficient free enthalpy associated with these interactions for different species to be bound preferentially in different cases.

When considering which mixture of species will be present in a given environments the pathway and kinetics of interconversion between tautomers may also play a role. For example, W01 can be converted to W17 in a single step whereas the conversion of W02 to W17 requires W01 as an intermediate. The energy barriers associated with these interconversions are yet to be investigated. The population of the neutral form will also play a role. The pKa of Warfarin

free in solution is ~ 5 . Thus, at neutral pH only about 1% of Warfarin would be in the neutral form. The therapeutic plasma concentration of Warfarin is in the μmolar range meaning that the alternative tautomeric forms described above will be present at pharmacologically important nanomolar concentrations. However, small changes in the apparent pK_a of Warfarin would mean that in different environments the population of the neutral form with its corresponding tautomers could be much larger.

Uncertainty:

Although in this case the combination of the use of homodesmotic reactions to estimate the free enthalpy of formation in vacuum together with the use of explicit solvent to estimate the free enthalpy of solvation was able to reproduce most of the available experimental data, the available experimental data is limited to only a small number of the tautomers potentially present under different conditions. While the populations of the 2 or 3 most common species may have been correctly predicted, the order of the remaining species varied dramatically. Variations of only a few kJ/mol in any of the terms considered was sufficient to change the rank order. This energy is comparable to the intrinsic uncertainty in each stage of the calculation. For example, the direct and homodesmotic approaches showed a relative difference of up to $\sim 22 \text{ kJ/mol}$ for a given tautomer. Even if a higher-level theory was used to estimate the heats of formation and convergence between the direct and homodesmotic methods was obtained, the potential residual uncertainty would still mean there is no guarantee that the ranking of all 20 tautomers would be correct. It should be noted that in the current work only one conformer of each tautomer was considered when estimating the heat of formation and the vibrational contribution to the entropy. Equally, the approach used to estimate the conformational entropy was very crude. What we can conclude in this case is that the implicit solvation models considered were not of sufficient accuracy to capture even basic experimental

trends such as the difference in solubility between water, DMSO and hexane. Indeed, the practice of combining high level quantum mechanical calculations with very simplistic solvation models must be questioned.

Conclusions

The role of possible alternative tautomeric states is increasingly recognised as a key challenge in structure-based drug design. To unambiguously determine which protomeric and/or tautomeric state of a ligand is bound to an acceptor often requires a combination of both high-resolution X-ray data together with neutron diffraction data and/or NMR data. For this reason, there is much interest in the use of computational approaches to determine the preferred tautomeric states in a given environment.^{2, 59} Here, different combinations of quantum mechanical calculations, molecular dynamics simulation and free energy calculations have been used in an attempt to identify the thermodynamically preferred tautomeric state of the anti-coagulant Warfarin. It was found that only a combination of homodesmic reactions to estimate the free enthalpy of formation in vacuum and an explicit solvent model to estimate the free enthalpy of solvation was able to reproduce the available experimental data. At the same time, we have shown that the intrinsic uncertainty in the calculation of each of the terms that must be considered when estimating the relative stability of alternative tautomeric states in a given solvent, could alter the predicted relative populations of each of the species significantly, highlighting the importance of ensuring that each term is calculated with high fidelity. Finally, we have shown that eight neutral Warfarin tautomers (*R* and *S* enantiomers of W01, W02, W17 and W18) are predicted to be separated by less than 16 kJ/mol in water, underlining the importance of considering a range of potential species when performing structural studies of Warfarin complexes or attempting to understand how Warfarin might interact with biological membranes.

ASSOCIATED CONTENT

Supporting Information

Figure S1: Representative homodesmic reactions. Table S1, S2 and S3: The total free enthalpy in solution ($\Delta G_f + \Delta G_{solv}$) for 20 tautomers of *R*-Warfarin in water, DMSO and hexane solvation free enthalpies obtained using implicit solvent models PCM, COSMO and SMD respectively.

AUTHOR INFORMATION

Corresponding Author

School of Chemistry and Molecular Biosciences

University of Queensland

St Lucia, QLD 4072

Australia

Phone: +61 7 336 54616

Fax: +61 7 336 53872

E-mail: a.malde@uq.edu.au

School of Chemistry and Molecular Biosciences

University of Queensland

St Lucia, QLD 4072

Australia

Phone: +61 7 336 54180

Fax: +61 7 336 53872

E-mail: a.e.mark@uq.edu.au

Funding Sources

This work was funded from the Australian Grants Commission (Discovery grants DP130102153 and DP160103414) with the assistance of high-performance computing resources provided by through the National Computational Merit Allocation Scheme supported by the Australian Government (projects n63 and m72).

ACKNOWLEDGMENT

ABBREVIATIONS

ATB, Automated Topology Builder; COSMO, COnductor-like Screening MOdel; DMSO, DiMethyl SulfOxide; MD, Molecular Dynamics; PCM, Polarizable Continuum Model; QM, Quantum Mechanical; SMD, Self-consistent reaction field Model using charge Density; SPC, Simple Point Charge

FIGURES

Figure 1. Chemical structures of a representative neutral form and anionic form of *R*-Warfarin.

Figure 2. Chemical structures of 20 tautomeric forms of *R*-Warfarin.

Figure 3. Steps involved in the formation of a molecule in vacuum and its subsequent solvation.

Figure 4. Groups identified as similar for the multi-molecule ESP charge fitting scheme.

TABLES

Table 1. Experimental values for enthalpy of formation of gas phase at standard conditions (1 atm pressure, 298.15 K).

Molecule		ΔH_f (expt.) (kJ/mol)
Coumarin	A	-163.4
Chromone	B	-126.1
Dihydrobenzopyran	C	-82.4
Hydrogen	D	0.0
Carbon dioxide	E	-393.5
Water	F	-241.8
Methane	G	-74.4
Ethane	H	-83.8
Ethene	I	52.3
Methanol	J	-205.0
Ethanol	K	-234.0
Isopropanol	L	-272.8
Z-propen-1-ol	M	-174.0
Propen-2-ol	N	-176.0
Acetone	O	-218.5
Toluene	P	50.0

Table 2. Homodesmotic reactions used to calculate the relative standard enthalpy of formation (ΔH_f°) and relative standard free enthalpy of formation (ΔG_f°) of 20 tautomers of *R*-Warfarin in the gas phase.

Tautomer	Homodesmotic Reaction
W01	$W01 + I + 3G \rightarrow A + P + H + M + O$
W02	$W02 + I + 3G \rightarrow B + P + H + M + O$
W03	$W03 + 5G \rightarrow C + P + 2H + O + E$
W04	$W04 + 5G \rightarrow C + P + 2H + O + E$
W05	$W05 + I + 3G \rightarrow A + P + H + M + N$
W06	$W06 + I + 3G \rightarrow A + P + H + M + N$
W07	$W07 + I + 3G \rightarrow B + P + H + M + N$
W08	$W08 + I + 3G \rightarrow B + P + H + M + N$
W09	$W09 + 5G \rightarrow C + P + 2H + N + E$
W10	$W10 + 5G \rightarrow C + P + 2H + N + E$
W11	$W11 + 5G \rightarrow C + P + 2H + N + E$
W12	$W12 + 5G \rightarrow C + P + 2H + N + E$
W13	$W13 + I + 3G \rightarrow A + P + H + M + N$
W14	$W14 + I + 3G \rightarrow B + P + H + M + N$
W15	$W15 + 5G \rightarrow C + P + 2H + N + E$
W16	$W16 + 5G \rightarrow C + P + 2H + N + E$
W17	$W17 + I + 4G + F \rightarrow A + P + H + M + L + J$
W18	$W18 + I + 4G + F \rightarrow A + P + H + M + L + J$
W19	$W19 + I + 4G + F \rightarrow B + P + H + M + L + J$
W20	$W20 + I + 4G + F \rightarrow B + P + H + M + L + J$

Table 3. The enthalpy of formation (ΔH_f) and the free enthalpy of formation (ΔG_f) in vacuum for 20 tautomers of *R*-Warfarin derived directly from geometry optimization of 20 tautomers versus the ones calculated using homodesmotic reactions at the B3LYP/6-311+G(3df,2p) level of theory. The entropy (S) for each molecule was estimated by combining the vibrational entropy ($ZPVE$) calculated using a single conformation at the B3LYP/6-311+G(3df,2p) level of theory at a temperature, $T = 298.15\text{K}$, with an estimate of the relative conformational entropy calculated based on the relative number of equivalent conformers as described in the text. The relative values are shown in kJ/mol.

	$\Delta\Delta H_f$		$-T\Delta\Delta S$ (relative)		$\Delta\Delta G_f$	
	direct	homodesmotic	ZPVE	S_{conf}	direct	homodesmotic
W01	0.0	6.3	-8.3	-3.8	0.0	0.0
W02	29.8	33.3	-8.0	-3.8	30.2	27.3
W03	35.4	51.4	-15.8	-2.0	29.7	39.4
W04	32.7	48.7	-13.6	-2.0	29.2	38.9
W05	64.0	55.1	-8.0	-3.8	64.3	49.1
W06	81.0	72.1	-15.9	-3.8	73.5	58.2
W07	81.7	69.9	-7.8	-3.8	82.2	64.1
W08	101.4	89.6	-15.5	-3.8	94.3	76.2
W09	78.5	79.3	-9.7	-2.0	78.9	73.4
W10	83.4	84.2	-14.4	-2.0	79.1	73.6
W11	70.8	71.5	-8.6	-2.0	72.3	66.8
W12	86.0	86.8	-14.5	-2.0	81.5	76.0
W13	87.7	78.7	-9.0	-3.8	87.0	71.7
W14	104.6	92.9	-7.8	-3.8	105.2	87.1
W15	95.4	96.2	-8.9	-2.0	96.6	91.1
W16	93.6	94.4	-8.8	-2.0	94.8	89.3
W17	8.5	0.0	-1.0	-0.0	19.6	4.8
W18	13.9	5.4	-0.0	-0.0	26.0	11.2
W19	41.7	30.4	-0.5	-0.0	53.3	35.6
W20	47.0	35.6	-0.2	-0.0	58.8	41.2

Table 4. The free enthalpy of solvation (ΔG_{solv}) for 20 tautomers of *R*-Warfarin using three different implicit solvation models (PCM, COSMO and SMD) and in explicit solvent, for water, DMSO and hexane as described in the methods. The values are in kJ/mol. The statistical error in the explicit solvent calculations is ± 1.0 kJ/mol. The range is the difference between the highest and lowest values.

	PCM			COSMO			SMD			Explicit solvent		
	Water	DMSO	hexane	Water	DMSO	hexane	Water	DMSO	hexane	Water	DMSO	hexane
W01	-38.3	-37.6	-12.4	-38.7	-38.2	-16.3	-45.9	-66.0	-54.9	-60.9	-96.3	-70.3
W02	-37.2	-36.5	-11.8	-37.5	-37.1	-15.7	-45.9	-64.8	-54.7	-79.5	-110.3	-66.3
W03	-43.7	-43.0	-14.8	-44.0	-43.6	-19.0	-57.1	-76.6	-58.8	-66.3	-99.7	-65.6
W04	-36.9	-36.3	-12.4	-37.3	-36.9	-16.1	-43.8	-67.4	-56.2	-66.4	-99.4	-70.0
W05	-42.7	-42.0	-13.9	-43.1	-42.6	-18.1	-53.8	-67.1	-54.7	-73.7	-118.9	-69.2
W06	-49.5	-48.7	-15.9	-50.0	-49.4	-20.9	-67.1	-73.6	-56.9	-73.0	-120.0	-68.2
W07	-39.5	-38.8	-12.9	-39.8	-39.4	-16.7	-47.9	-60.6	-53.4	-77.8	-114.7	-69.7
W08	-44.0	-43.3	-14.4	-44.4	-44.0	-18.8	-59.5	-64.7	-54.7	-79.5	-117.5	-67.7
W09	-35.0	-34.5	-11.7	-35.4	-35.0	-15.2	-44.0	-65.7	-54.9	-65.2	-106.9	-64.0
W10	-38.9	-38.3	-13.4	-39.2	-38.8	-17.1	-54.1	-69.2	-55.8	-65.9	-102.1	-65.3
W11	-35.4	-34.8	-11.6	-35.7	-35.3	-15.3	-43.5	-65.2	-54.5	-64.6	-101.3	-69.5
W12	-39.4	-38.7	-12.7	-39.8	-39.4	-16.9	-53.7	-69.2	-55.2	-68.0	-107.4	-64.7
W13	-46.1	-45.4	-15.4	-46.5	-46.0	-19.5	-58.7	-70.1	-54.9	-86.0	-132.5	-67.9
W14	-42.5	-41.9	-14.3	-42.9	-42.4	-18.1	-52.6	-63.0	-53.0	-84.4	-123.5	-66.1
W15	-39.8	-39.1	-13.1	-40.2	-39.8	-17.3	-49.3	-70.3	-54.6	-65.6	-104.4	-66.1
W16	-34.3	-33.7	-11.5	-34.6	-34.2	-14.9	-50.8	-63.4	-53.0	-70.9	-108.3	-70.1
W17	-43.6	-42.9	-13.9	-44.1	-43.6	-18.3	-54.9	-67.5	-54.8	-72.4	-114.3	-70.2
W18	-45.2	-44.4	-14.3	-45.7	-45.2	-19.1	-57.1	-68.9	-55.0	-63.2	-105.7	-68.3
W19	-42.7	-41.9	-13.6	-43.2	-42.7	-18.1	-55.6	-66.6	-54.8	-74.4	-113.5	-71.0
W20	-44.1	-43.3	-13.9	-44.6	-44.1	-18.7	-58.1	-67.9	-54.7	-69.2	-105.7	-70.0
Range	15.2	15.0	4.4	15.4	15.2	6.0	23.6	16.0	5.8	25.1	36.2	7.0

Table 5. The total free enthalpy in solution ($\Delta G_f + \Delta G_{solv}$) for 20 tautomers of *R*-Warfarin in water, DMSO and hexane solvation free enthalpies obtained using explicit solvent. The relative values are shown in kJ/mol.

	Water		DMSO		Hexane	
	direct	homodesmic	direct	homodesmic	direct	homodesmic
W01	0.0	6.7	0.0	13.2	0.0	0.0
W02	11.6	15.4	16.2	26.5	34.2	31.3
W03	24.3	40.7	26.3	49.2	34.4	44.1
W04	23.7	40.1	26.1	49.0	29.5	39.2
W05	51.5	43.0	41.7	39.7	65.4	50.2
W06	61.4	52.8	49.8	47.7	75.6	60.3
W07	65.3	53.9	63.8	58.9	82.8	64.7
W08	75.7	64.3	73.1	68.2	96.9	78.8
W09	74.6	75.8	68.3	76.0	85.2	79.7
W10	74.1	75.3	73.3	81.0	84.1	78.6
W11	68.6	69.8	67.3	75.0	73.1	67.6
W12	74.4	75.6	70.4	78.1	87.1	81.6
W13	61.9	53.3	50.8	48.7	89.4	74.1
W14	81.7	70.3	78.0	73.1	109.4	91.3
W15	91.9	93.1	88.5	96.2	100.8	95.3
W16	84.8	86.0	82.8	90.5	95.0	89.5
W17	8.1	0.0	1.6	0.0	19.7	4.9
W18	23.7	15.6	16.6	15.0	28.0	13.2
W19	39.8	28.8	36.1	31.6	52.6	34.9
W20	50.5	39.6	49.4	45.0	59.1	41.5

REFERENCES

1. Malde, A. K.; Mark, A. E., Binding and enantiomeric selectivity of threonyl-tRNA synthetase. *Journal of the American Chemical Society* **2009**, *131* (11), 3848-9.
2. Malde, A. K.; Mark, A. E., Challenges in the determination of the binding modes of non-standard ligands in X-ray crystal complexes. *Journal of Computer-Aided Molecular Design* **2011**, *25* (1), 1-12.
3. Nair, P. C.; Malde, A. K.; Mark, A. E., Using Theory to Reconcile Experiment: The Structural and Thermodynamic Basis of Ligand Recognition by Phenylethanolamine N-Methyltransferase (PNMT). *Journal of Chemical Theory and Computation* **2011**, *7* (5), 1458-1468.
4. Nair, P. C.; Malde, A. K.; Drinkwater, N.; Mark, A. E., Missing fragments: Detecting cooperative binding in fragment-based drug design. *ACS Medicinal Chemistry Letters* **2012**, *3*, 322-326.
5. Martin, Y. C., Let's not forget tautomers. *Journal of Computer-Aided Molecular Design* **2009**, *23* (10), 693-704.
6. Hiskey, C. F.; Bullock, E.; Whitman, G., Spectrophotometric study of aqueous solutions of warfarin sodium. *Journal of Pharmaceutical Sciences* **2006**, *51* (1), 43-46.
7. Porter, W. R., Warfarin: history, tautomerism and activity. *Journal of Computer-Aided Molecular Design* **2010**, *24* (6-7), 553-573.
8. Tie, J. K.; Stafford, D. W., Structure and Function of Vitamin K Epoxide Reductase. In *Vitamins and Hormones*, Academic Press: 2008; Vol. 78, pp 103-130.
9. Wallin, R.; Wajih, N.; Hutson, S. M., VKORC1: A Warfarin-Sensitive Enzyme in Vitamin K Metabolism and Biosynthesis of Vitamin K-Dependent Blood Coagulation Factors. In *Vitamins and Hormones*, Academic Press: 2008; Vol. 78, pp 227-246.
10. Bravic, G.; Gaultier, J.; Hauw, C., Crystal-Structure Of antivitamin-K Warfarin. *Comptes Rendus Hebdomadaires Des Seances De L Academie Des Sciences Serie C* **1973**, *277* (22), 1215-1218.
11. Valente, E. J.; Trager, W. F.; Jensen, L. H., Crystal and molecular-structure and absolute-configuration of (-)-(S)-Warfarin. *Acta Crystallographica Section B-Structural Science* **1975**, *31* (APR15), 954-960.
12. Hiskey, C. F.; Melnitchenko, V., Clathrates of sodium Warfarin. *Journal of Pharmaceutical Sciences* **1965**, *54* (9), 1298-302.
13. Sheth, A. R.; Young, V. G., Jr.; Grant, D. J. W., Warfarin sodium 2-propanol solvate. *Acta Crystallographica Section E* **2002**, *58* (5), m197-m199.
14. Valente, E. J.; Lingafelter, E. C.; Porter, W. R.; Trager, W. F., Structure of Warfarin in solution. *Journal of Medicinal Chemistry* **1977**, *20* (11), 1489-1493.
15. Guasch, L.; Peach, M. L.; Nicklaus, M. C., Tautomerism of Warfarin: Combined Chemoinformatics, Quantum Chemical, and NMR Investigation. *Journal of Organic Chemistry* **2015**, *18* (20), 9900-9909.
16. Giannini, D. D.; Chan, K. K.; Roberts, J. D., Carbon-13 Nuclear Magnetic Resonance Spectroscopy. Structure of the Anticoagulant Warfarin and Related Compounds in Solution. *Proceedings of the National Academy of Sciences* **1974**, *71* (10), 4221-4223.
17. Stella, V. J.; Mooney, K. G.; Pipkin, J. D., Dissolution and Ionization of Warfarin. *Journal of Pharmaceutical Sciences* **1984**, *73* (7), 946-948.
18. Moser, A. C.; Kingsbury, C.; Hage, D. S., Stability of warfarin solutions for drug-protein binding measurements: Spectroscopic and chromatographic studies. *Journal of Pharmaceutical and Biomedical Analysis* **2006**, *41* (4), 1101-1109.

19. Tode, C.; Takeuchi, A.; Iwakawa, S.; Tatsumi, A.; Sugiura, M., Hydrogen-Deuterium (H-D) Exchange Reaction of Warfarin in D₂O Solution. *Chemical and Pharmaceutical Bulletin* **2009**, *57* (7), 653-656.
20. Pisklak, M.; Maciejewska, D.; Herold, F.; Wawer, I., Solid state structure of coumarin anticoagulants: Warfarin and Sintrom. C-13 CPDAS NMR and GIAO DFT calculations. *Journal of Molecular Structure* **2003**, *649* (1-2), 169-176.
21. Kong, X.; Dai, Y.; Wu, G., Solid-state ¹⁷O NMR study of 2-acylbenzoic acids and warfarin. *Solid State Nuclear Magnetic Resonance* **2017**, *84* (Supplement C), 59-64.
22. Tárkányi, G.; Németh, K.; Mizsei, R.; Tőke, O.; Visy, J.; Simonyi, M.; Jicsinszky, L.; Szemán, J.; Szente, L., Structure and stability of Warfarin-sodium inclusion complexes formed with permethylated monoamino- β -cyclodextrin. *Journal of Pharmaceutical and Biomedical Analysis* **2013**, *72*, 292-298.
23. Karlsson, B. C. G.; Rosengren, A. M.; Andersson, P. O.; Nicholls, I. A., The spectrophysics of warfarin: Implications for protein binding. *Journal of Physical Chemistry B* **2007**, *111* (35), 10520-10528.
24. Karlsson, B. C. G.; Rosengren, A. M.; Andersson, P. O.; Nicholls, I. A., Molecular Insights on the Two Fluorescence Lifetimes Displayed by Warfarin from Fluorescence Anisotropy and Molecular Dynamics Studies. *Journal of Physical Chemistry B* **2009**, *113* (22), 7945-7949.
25. Petitpas, I.; Bhattacharya, A. A.; Twine, S.; East, M.; Curry, S., Crystal structure analysis of warfarin binding to human serum albumin - Anatomy of drug site I. *Journal of Biological Chemistry* **2001**, *276* (25), 22804-22809.
26. Williams, P. A.; Cosme, J.; Ward, A.; Angova, H. C.; Vinkovic, D. M.; Jhoti, H., Crystal structure of human cytochrome P450C₉ with bound warfarin. *Nature* **2003**, *424* (6947), 464-468.
27. Ghuman, J.; Zunszain, P. A.; Petitpas, I.; Bhattacharya, A. A.; Otagiri, M.; Curry, S., Structural basis of the drug-binding specificity of human serum albumin. *Journal of Molecular Biology* **2005**, *353* (1), 38-52.
28. He, M.; Korzekwa, K. R.; Jones, J. P.; Rettie, A. E.; Trager, W. F., Structural Forms of Phenprocoumon and Warfarin That Are Metabolized at the Active Site of CYP2C₉. *Archives of Biochemistry and Biophysics* **1999**, *372* (1), 16-28.
29. Nagy, P. I., Theoretical Consideration of In-Solution Tautomeric Equilibria in Relation to Drug Design. In *Tautomerism*, Antonov, L., Ed. 2016.
30. Raghavachari, K.; Curtiss, L. A., *Quantum-Chemical Methods for Accurate Theoretical Thermo- chemistry*. Kluwer: New York, 2002.
31. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J., Gaussian 09, Revision D.01. Wallingford CT, 2016.
32. Lee, C. T.; Yang, W. T.; Parr, R. G., Development of the colle-salvetti correlation-energy formula into a functional of the electron-density. *Physical Review B* **1988**, *37* (2), 785-789.

33. Perdew, J. P.; Wang, Y., Accurate and simple analytic representation of the electron-gas correlation-energy. *Physical Review B* **1992**, *45* (23), 13244-13249.
34. Becke, A. D., Density-functional thermochemistry .3. The role of exact exchange. *Journal of Chemical Physics* **1993**, *98* (7), 5648-5652.
35. Miertuš, S.; Scrocco, E.; Tomasi, J., Electrostatic interaction of a solute with a continuum. A direct utilization of ab initio molecular potentials for the prevision of solvent effects. *Chemical Physics* **1981**, *55* (1), 117-129.
36. Marenich, A. V.; Cramer, C. J.; Truhlar, D. G., Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *The Journal of Physical Chemistry B* **2009**, *113* (18), 6378-6396.
37. Klamt, A.; Schuurmann, G., COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *Journal of the Chemical Society, Perkin Transactions 2* **1993**, (5), 799-805.
38. Christen, M.; Hunenberger, P. H.; Bakowies, D.; Baron, R.; Burgi, R.; Geerke, D. P.; Heinz, T. N.; Kastenholz, M. A.; Krautler, V.; Oostenbrink, C.; Peter, C.; Trzesniak, D.; van Gunsteren, W. F., The GROMOS software for biomolecular simulation: GROMOS05. *Journal of Computational Chemistry* **2005**, *26* (16), 1719-1751.
39. Kunz, A. P. E.; Allison, J. R.; Geerke, D. P.; Horta, B. A. C.; Hunenberger, P. H.; Riniker, S.; Schmid, N.; van Gunsteren, W. F., New functionalities in the GROMOS biomolecular simulation software. *Journal of Computational Chemistry* **2012**, *33* (3), 340-353.
40. Schmid, N.; Christ, C. D.; Christen, M.; Eichenberger, A. P.; van Gunsteren, W. F., Architecture, implementation and parallelisation of the GROMOS software for biomolecular simulation. *Computer Physics Communications* **2012**, *183* (4), 890-903.
41. Malde, A. K.; Zuo, L.; Breeze, M.; Stroet, M.; Poger, D.; Nair, P. C.; Oostenbrink, C.; Mark, A. E., An Automated force field Topology Builder (ATB) and repository: Version 1.0. *Journal of Chemical Theory and Computation* **2011**, *7* (12), 4026-4037.
42. Canzar, S.; El-Kebir, M.; Pool, R.; Elbassioni, K.; Malde, A. K.; Mark, A. E.; Geerke, D. P.; Stougie, L.; Klau, G. W., Charge Group Partitioning in Biomolecular Simulation. In *Research in Computational Molecular Biology*, Chor, B., Ed. Springer Berlin / Heidelberg: 2012; Vol. 7262, pp 29-43.
43. Koziara, K. B.; Stroet, M.; Malde, A. K.; Mark, A. E., Testing and validation of the Automated Topology Builder (ATB) version 2.0: Prediction of hydration free energies. *Journal of Computer-Aided Molecular Design* **2014**, *28* (3), 221-233.
44. Oostenbrink, C.; Villa, A.; Mark, A. E.; van Gunsteren, W. F., A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *Journal of Computational Chemistry* **2004**, *25* (13), 1656-1676.
45. Schmid, N.; Eichenberger, A.; Choutko, A.; Riniker, S.; Winger, M.; Mark, A. E.; van Gunsteren, W. F., Definition and testing of the GROMOS force-field versions 54A7 and 54B7. *European Biophysics Journal* **2011**, *40* (7), 843-856.
46. Singh, U. C.; Kollman, P. A., An approach to computing electrostatic charges for molecules. *Journal of Computational Chemistry* **1984**, *5* (2), 129-145.
47. Dupradeau, F.-Y.; Pigache, A.; Zaffran, T.; Savineau, C.; Lelong, R.; Grivel, N.; Lelong, D.; Rosanski, W.; Cieplak, P., The R.E.D. tools: advances in RESP and ESP charge derivation and force field library building. *Physical Chemistry Chemical Physics* **2010**, *12* (28), 7821-7839.

48. Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J., Interaction models for water in relation to protein hydration. In "*Intermolecular Forces*", Pullman, B., Ed. Reidel: Dordrecht, The Netherlands, 1981; pp 331-342.
49. Geerke, D. P.; Oostenbrink, C.; van der Vegt, N. F. A.; van Gunsteren, W. F., An Effective Force Field for Molecular Dynamics Simulations of Dimethyl Sulfoxide and Dimethyl Sulfoxide–Water Mixtures. *The Journal of Physical Chemistry B* **2004**, *108* (4), 1436-1445.
50. Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Dinola, A.; Haak, J. R., Molecular-dynamics with coupling to an external bath. *Journal of Chemical Physics* **1984**, *81* (8), 3684-3690.
51. Heinz, T. N.; van Gunsteren, W. F.; Hünenberger, P. H., Comparison of four methods to compute the dielectric permittivity of liquids from molecular dynamics simulations. *The Journal of Chemical Physics* **2001**, *115* (3), 1125-1136.
52. Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C., Numerical-Integration of Cartesian Equations of Motion of a System with Constraints - Molecular-Dynamics of N-Alkanes. *Journal of Computational Physics* **1977**, *23* (3), 327-341.
53. van Gunsteren, W. F.; Beutler, T. C.; Fraternali, F.; King, P. M.; Mark, A. E.; Smith, P. E., *Computer Simulation of Biomolecular Systems, Theoretical and Experimental Applications*, . ESCOM Science Publishers: Leiden, 1989; p 315-348.
54. Beutler, T. C.; Mark, A. E.; van Schaik, R. C.; Gerber, P. R.; van Gunsteren, W. F., Avoiding singularities and numerical instabilities in free-energy calculations based on molecular simulations. *Chemical Physics Letters* **1994**, *222* (6), 529-539.
55. Rosengren, A. M.; Karlsson, B. C. G., Spectroscopic evidence for the presence of the cyclic hemiketal form of warfarin in aqueous solution: Consequences for bioavailability. *Biochemical and Biophysical Research Communications* **2011**, *407* (2), 318-320.
56. Illum, L.; Bundgaard, H., Sorption of drugs by plastic infusion bags. *International Journal of Pharmaceutics* **1982**, *10*, 339-351.
57. Caron, G.; Ermondi, G., Calculating Virtual log P in the Alkane/Water System (log PNalk) and Its Derived Parameters $\Delta\log P_{\text{Noct-alk}}$ and $\log D_{\text{pHalk}}$. *Journal of Medicinal Chemistry* **2005**, *48* (9), 3269-3279.
58. Afzelius, L.; Zamora, I.; Masimirembwa, C. M.; Karlén, A.; Andersson, T. B.; Mecucci, S.; Baroni, M.; Cruciani, G., Conformer- and Alignment-Independent Model for Predicting Structurally Diverse Competitive CYP2C9 Inhibitors. *Journal of Medicinal Chemistry* **2004**, *47* (4), 907-914.
59. Furet, P.; Meyer, T.; Strauss, A.; Raccuglia, S.; Rondeau, J. M., Structure-based design and protein X-ray analysis of a protein kinase inhibitor. *Bioorganic & Medicinal Chemistry Letters* **2002**, *12* (2), 221-224.

TOC Graphics