Analytische Entwicklung polarisierbarer Kraftfelder für Wasser

Philipp Tröster

München 2014

Analytische Entwicklung polarisierbarer Kraftfelder für Wasser

Philipp Tröster

Dissertation
an der Fakultät für Physik
der Ludwig-Maximilians-Universität
München

vorgelegt von
Philipp Tröster
aus München

München, den 28. April 2014

Erstgutachter: Prof. Dr. Paul Tavan
Zweitgutachter: Prof. Dr. Christian Ochsenfeld
Tag der mündlichen Prüfung: 14. 10. 2014

Zusammenfassung

Wasser ist ohne Zweifel die wichtigste Flüssigkeit unseres Planeten. Aufgrund seiner höchst ungewöhnlichen Eigenschaften konnte das Leben aus dem Wasser entstehen. So bleibt es am Grunde von Seen und Meeren flüssig, weil seine Dichte (unter Normaldruck $p_0 = 1$ bar) bei der Temperatur $T^{\mathrm{md}} = 277,134$ K maximal wird, weil es erst bei der Schmelztemperatur des Eises $T^{\mathrm{m}} = 273,15$ K in die sehr viel weniger dichte feste Phase übergehen kann und weil schließlich deshalb das tieferliegende Wasser vom darüber schwimmenden Eis von der Umgebungskälte isoliert wird.

Daher ist auch die theoretische Erklärung, wie sich die makroskopische Physik des Wassers aus den mikroskopischen Eigenschaften seiner H_2O Moleküle ergibt, von großem wissenschaftlichem Interesse. Die vorliegende Dissertation leistet dazu einen Beitrag, indem sie durch quantenmechanische Beschreibungen einzelner Moleküle im Rahmen der Dichtefunktionaltheorie (DFT), welche in polarisierbare molekülmechanische (PMM) Modelle ihrer flüssigen Umgebung eingebettet sind, neue und verbesserte PMM Potentialfunktionen zunehmender Komplexität ableitet.

Dazu stellt sie drei kürzlich erschienene Publikationen /4-6/ vor. Das erste Resultat /4/ ist dabei die Entwicklung einer selbstkonsistenten Methode zur Parametrisierung von PMM Wassermodellen, welche sich auf eine neue Hybridtechnik zur DFT/PMM Molekulardynamik-(MD-)Simulation /3/, auf bekannte Eigenschaften des H_2O Moleküls in der Gasphase (z.B. Dipolmoment, Polarisierbarkeit) und auf DFT/MM Vorarbeiten zu seiner Polarisierbarkeit in der Flüssigkeit [z.B. Schropp und Tavan (2010). *J. Phys. Chem B*, 114, 2051-2057] stützen konnte. Dieses DFT/PMM gestützte Vorgehen liefert die elektrostatischen Eigenschaften der H_2O Modelle. Daher müssen lediglich drei Parameter von van der Waals Modellpotentialen an drei Messwerte zur flüssigen Phase bei p_0 und $T_0 \equiv 300$ K angepasst werden.

Die Anwendung dieser Methode ergab drei durch den Parameter $\nu = 4, 5, 6$ abgezählte und als TLνP bezeichnete PMM Modelle zunehmender Komplexität /4,5/, wobei $\nu - 1$ die Anzahl der Punktladungen angibt, die zur Modellierung des statischen Anteils der elektrostatischen Signatur eines Wassermoleküls in flüssiger Phase verwendet werden. Nachdem die Elektrostatik der TLνP Modelle anhand von DFT/PMM Rechnungen und ihre van der Waals Potentiale durch PMM-MD Simulationen optimiert waren, konnten die Vorhersagen der damit erzeugten PMM Modelle für viele Eigenschaften von Wasser durch eine Vielzahl weiterer Simulationen, die auch den Temperaturbereich (250-320 K) der Dichteanomalie und des Schmelzpunktes umfassten, getestet werden.

Es zeigte sich, dass sich die TLνP Vorhersagen mit zunehmender Komplexität ν zwar immer weiter vielen Beobachtungsdaten annäherten, aber bis zu $\nu = 5$ in einigen Aspekten noch deutlich vom quantenmechanischen Vorbild (Quadrupolmomente) und von der experimentellen Evidenz [Dichteverlauf $n(T, p_0)$] abwichen /4/. Überraschenderweise reproduzierte das Sechspunktmodell TL6P jedoch plötzlich hervorragend sowohl die DFT/MM Elektrostatik von H_2O in Wasser als auch den Dichteverlauf $n(T, p_0)$. So sagte es für T^{md} mit 277,005 K einen im Rahmen der Statistik ununterscheidbaren Wert vorher /6/. Ebenfalls viel besser als Vorhersagen früherer PMM Modelle ist die Vorhersage für T^{m}, die T^{m} um weniger als 10 K unterschätzt /5/. Die physikalische Bedeutung dieser Befunde wird erklärt /4-6/.

Verzeichnis der im Rahmen dieser Arbeit entstandenen Publikationen

/1/ J VandeVondele, P Tröster, P Tavan, and G Mathias (2012). Vibrational Spectra of Phosphate Ions in Aqueous Solution Probed by First Principles Molecular Dynamics *J. Phys. Chem. A* 116, 2466-2474.

/2/ K Lorenzen, M Schwörer, P Tröster, S Mates, and P Tavan (2012). Optimizing the Accuracy and Efficiency of Fast Hierarchical Multipole Expansions for MD Simulations. *J. Chem. Theory Comput.* 8, 3628-3636.

/3/ M Schwörer, B Breitenfeld, P Tröster, S Bauer, K Lorenzen, P Tavan, and G Mathias (2013). Coupling DFT to Polarizable Force Fields for Efficient and Accurate Hamiltonian Molecular Dynamics Simulations. *J. Chem. Phys.* 138, 244103 (1-13).

/4/ P Tröster, K Lorenzen, M Schwörer, P Tavan (2013). Polarizable Water Models from Mixed Computational and Empirical Optimization. *J. Phys. Chem. B* 117, 9486-9500.

/5/ P Tröster, K Lorenzen, and P Tavan (2014). Polarizable Six-Point Water Models from Computational and Empirical Optimization. *J. Phys. Chem. B* 118, 1589-1602.

/6/ P Tröster and P Tavan (2014). The Microscopic Physical Cause for the Density Maximum of Liquid Water. *J. Phys. Chem. Lett.* 5, 138-142.

Die mit blauer Farbe hervorgehobenen Arbeiten sind in den Text der Dissertation eingearbeitet und dort nachgedruckt.

Inhaltsverzeichnis

1 Einleitung

Da Leben im Wasser entstanden ist, kann man Wasser unzweifelhaft als das wichtigste Lösungsmittel der Erde bezeichnen [1, 2]. Der Einfluss dieser höchst ungewöhnlichen Flüssigkeit spiegelt sich in den Formen und Funktionen der biologischen Makromoleküle wider, die für das Leben verantwortlich sind. Noch immer bergen die sonderbaren Eigenschaften des Wassers Geheimnisse, die Chemiker und Physiker aufzuklären trachten. Die Bedeutung der Wasserforschung lässt sich nicht nur an den zahllosen Veröffentlichungen auf diesem Gebiet ablesen, sondern auch an der Einführung des "Stockholm Water Prize", des hochdotierten sogenannten Nobelpreises für Wasser, der jedes Jahr verliehen wird.

1.1 Wasser, eine ungewöhnliche Flüssigkeit

Auf den ersten Blick könnte man Wasser als einen langweiligen Stoff bezeichnen: Es ist geschmacklos, geruchlos und vor allem ist es allgegenwärtig. Allerdings liegt schon in der Allgegenwart eine erste Besonderheit. Wasser ist die einzige chemische Verbindung, welche auf der Erdoberfläche in allen drei Aggregatszuständen vorkommt. Es ist ferner die einfachste Verbindung der beiden sehr häufigen und reaktiven Elemente Wasserstoff und Sauerstoff.

Dass das Wasser, trotz der vermeintlichen Einfachheit des H_2O-Moleküls, diejenige Flüssigkeit ist, die mit Abstand die meisten Besonderheiten aufweist, ist eines der Wunder der Natur. Um den Geheimnissen des Wassers auf die Spur zu kommen, müssen die dynamischen Abläufe, die aus der Wechselwirkung der Moleküle untereinander resultieren, im Detail verstanden werden.

Die vorliegende Arbeit, welche im Gebiet der theoretischen chemischen Physik angesiedelt ist, wird zeigen, wie man mittels computergestützter Molekulardynamik (MD)-Simulationen, die polarisierbare molekülmechanische (PMM) Kraftfelder mit quantenmechanischen (QM) Berechnungen kombinieren, Einblicke in die ungewöhnlichen Eigenschaften des Wassers gewinnen und Ursachen derselben finden kann [die grundlegenden Eigenschaften molekülmechanischer (MM) Kraftfelder und die verwendete Darstellung der Polarisierbarkeit werden in Abschnitt 1.2 erläutert].

Aus den mehr als 70 Anomalien des Wassers [3] sind einige der interessantesten die ungewöhnlich hohe Dielektrizitätskonstante, der ungewöhnlich hohe Schmelz- und Siedepunkt, sowie der hohe kritische Punkt, das Phänomen unterkühlten Wassers und die Tatsache, dass dieses durch Erwärmen gefriert, die hohe Viskosität, die hohe Wärmekapazität, die ungewöhnlich kleine Kompressibilität und die wohl bekannteste, die Dichteanomalie [3, 4].

Die mikroskopische physikalische Ursache, die der Dichteanomalie zugrunde liegt, war bis vor kurzem nicht geklärt. Eines der wichtigsten und spannendsten Ergebnisse der vorliegen-

den Arbeit ist die eindeutige Klärung dieser Ursache der Anomalie. Es ist uns nämlich gelungen, ein PMM Modellpotential für Wasser zu entwickeln [5], welches das experimentell gemessene Temperatur-Dichte Profil $n(T, p_0)$ [6] in einem großen Bereich von Temperaturen T beim Normaldruck $p_0 \equiv 1$ bar mit bisher unerreichter Genauigkeit vorhersagt [7], obwohl bei der Entwicklung dieses PMM Models namens TL6P[1] lediglich der experimentelle Wert $n(T_0, p_0)$ der Dichte bei T_0 K vorausgesetzt wurde. Da zusätzlich zwei weitere, weniger komplexe Modellpotentiale (TL4P und TL5P) für Wasser mit identischen Prozeduren der Parametrisierung entwickelt wurden [8], welche das beobachtete Temperatur-Dichte-Profil weit verfehlen, konnte der mikroskopische Mechanismus, der zur Dichteanomalie führt, eindeutig identifiziert werden [7].

Ermöglicht wurden diese Erkenntnisse durch die Entwicklung eines neuen Verfahrens zur Parametrisierung polarisierbarer Modellpotentiale von Lösungsmittelmolekülen [5, 8], welches ein weiteres Hauptergebnis dieser Dissertation darstellt. Die entscheidende technische Voraussetzung dafür bildete eine neu entworfene Hamiltonsche Kopplung [9] des PMM-MD Programms IPHIGENIE [9, 10, 11], mit dem gitterbasierten Dichtefunktionaltheorie (DFT) Programm CPMD [12]. Dabei stellt IPHIGENIE eine gründliche Überarbeitung des parallelisierten MM-MD Simulationsprogramms EGO [13, 14] dar.

DFT/PMM Hybridsysteme, welche mit der Programmkombination CPMD/IPHIGENIE berechenbar sind und in Abschnitt 1.2.3 näher erläutert werden, erlauben vermittels der DFT detaillierte Einblicke in die mikroskopischen Eigenschaften eines kleinen Moleküls oder eines Teils eines biologischen Makromoleküles in kondensierter Phase, die bei genau definierten thermodynamischen Bedingungen durch ein PMM-MD Simulationssystem repräsentiert wird.

Eine erste Kombination des MM-MD Simulationsprogramms EGO und des DFT-Programms CPMD, welches durch die Verwendung ebener Wellen als Basis der Kohn-Sham Wellenfunktionen effizient parallelisiert ist, wurde von Eichinger et al. 1999 entwickelt [15]. Durch diese DFT/MM Hybridmethode konnten Schropp und Tavan zwei wichtige Erkenntnisse über die Polarisierbarkeit gelöster Wassermoleküle gewinnen und daraus entsprechende Vorschläge zur Entwicklung von PMM Wassermodellen ableiten [16, 17].

Aufgrund der Bedeutung des Wassermoleküls und der erwähnten Arbeiten von Schropp und Tavan lag es nahe, das neue Parametrisierungsverfahren zunächst auf die Entwicklung von PMM Wassermodellen anzuwenden. Dabei gebietet der immer nötige Kompromiss zwischen Aufwand und Genauigkeit, das zu konstruierende Modellpotential so einfach wie möglich und lediglich so komplex wie nötig zu gestalten.

Darüber hinaus war der PMM Modellentwurf von der Überzeugung geleitet, dass jeder vorkommende Parameter physikalisch klar motiviert sein sollte. Dies ist ein beim Entwurf von (P)MM Wassermodellen kaum je beachteter Grundsatz [Abschnitt 1.2.3 gibt einen Überblick über existierende Wassermodelle]. So verfügt etwa das Wassermodell iAMOEBA [18], welches das experimentelle Temperatur-Dichte Profil relativ gut reproduzieren kann, über 19 freie Parameter, die anhand dieses Profils sowie einer Vielzahl weiterer (zumeist empirisch

[1] Der Name TL6P rührt von den Autoren Tröster und Lorenzen und von der Verwendung von sechs Ansatzpunkten für intermolekulare Kräfte her.

bestimmter) Größen in einer globalen Optimierung festgelegt wurden. Dabei geht aber jeder Zusammenhang zwischen den mikroskopischen und makroskopischen Eigenschaften des Modells verloren. Insbesondere kann die mikroskopische physikalische Ursache, welche der Dichteanomalie zugrunde liegt, auf diesem Wege nicht bestimmt werden.

1.1.1 Das Temperatur-Dichte Profil und die Temperatur maximaler Dichte

Abbildung 1.1 zeigt das beobachtete [6] Temperatur-Dichte Profil $n(T, p_0)$ flüssigen Wassers bei dem Normaldruck $p_0 = 1$ bar im Temperaturbereich $T \in [250, 320]$ K und das Dichtemaximum bei der Temperatur $T^{\mathrm{md}} \approx 4^\circ$ C (genau: 3.984° C [6]), die durch eine senkrechte Linie gekennzeichnet ist. Da die Dichte von Eis, dessen Schmelzpunkt bei der Temperatur $T^{\mathrm{m}} = 0^\circ$ C liegt, sehr viel kleiner ist, schwimmt es im Winter oben. Damit isoliert es die Flüssigkeit von der Umgebungskälte, so dass Seen nicht komplett gefrieren und Lebewesen in der Tiefe bei 4° C überleben können.

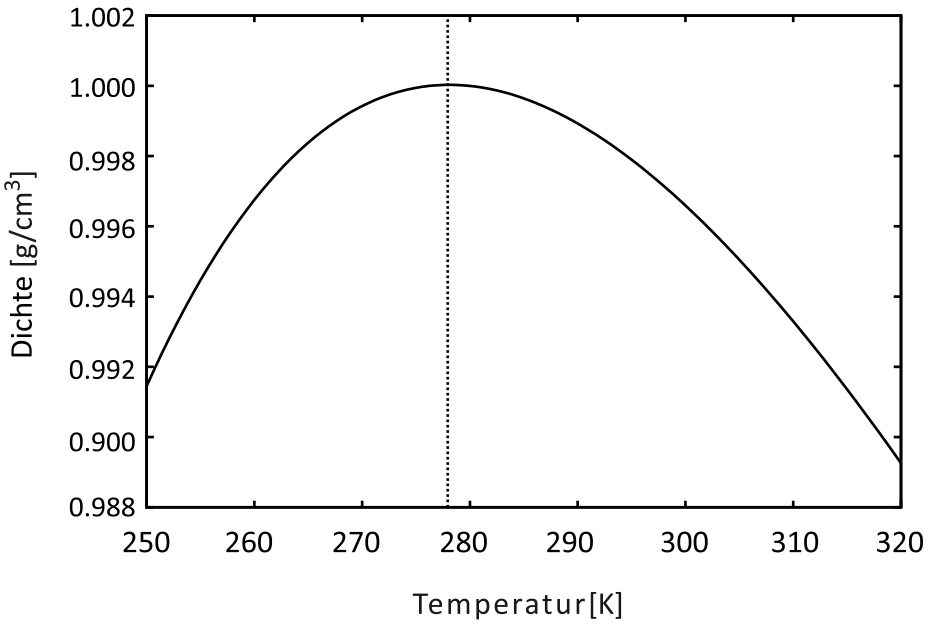


Abbildung 1.1: Die Graphik zeigt die von Kell [6] gemessene Dichte $n(T, p_0)$ im Temperaturbereich $T \in [250, 320]$ K. Die Temperatur T^{md} maximaler Dichte bei 277,134 K (3,984° C) ist ebenfalls eingezeichnet.

Für eine Flüssigkeit ist das in Abb. 1.1 gezeigte Verhalten ungewöhnlich und kommt nur bei wenigen Stoffen vor. In der Regel erwartet man bei abnehmender Temperatur eine zunehmende Dichte, da Moleküle dann durch ihre geringere kinetische Energie kleinere Volumina ausfüllen, also weniger Platz brauchen. Dieser Effekt ist aus der kinetischen Gastheorie bekannt [19]. Bei flüssigem Wasser ist dies oberhalb von 4° C ebenso. Aber unterhalb setzt ein Effekt ein, der zu einer Abnahme der Dichte führt. Diese Dichteabnahme setzt sich bis weit unter den Schmelzpunkt von Eis bei T^{m} fort, wobei man beachten muss, dass es sich hier

um unterkühltes Wasser, also H_2O in flüssiger Phase, und nicht um Eis handelt. Da im Alltag Wasser lediglich in Verbindung mit gelösten Stoffen, wie beispielsweise Mineralien, die einen Kristallisationskeim bilden, vorkommt, tritt in der Natur bei T^{m} immer der Phasenübergang in die feste Phase auf. Im Labor hingegen kann gezeigt werden, dass sich reines flüssiges Wasser bis zu einer Temperatur von etwa 232 K (-41° C) herunterkühlen lässt, bevor es zur Kristallisation kommt [20].

Am Beispiel eines Eiskristalls lässt sich der Effekt abnehmender Dichte bei abnehmender Temperatur qualitativ verstehen. Von den vielen Eiskristallen, die sich im Labor unter verschiedenen thermodynamischen Bedingungen bilden können[2], kommt in der Natur, also bei p_0, ausschließlich sogenanntes Eis Ih vor, dessen Dichte bei T^{m} um 8 % niedriger ist als die Dichte unterkühlten Wassers. Der Grund hierfür ist dessen aufgelockerte hexagonale Struktur[21], welche sich beim Festkörper über weite Distanzen erstreckt. Im Gegensatz dazu existiert in flüssiger Phase lediglich eine kurzreichweitige Ordnung [22].

Obwohl unterkühltes Wasser aufgrund fehlender Kristallisationskeime nicht gefriert, so wirken doch ähnliche strukturbildende Kräfte wie beim Eiskristall. Diese Kräfte stehen der thermischen Kontraktion bei abnehmender Temperatur entgegen. Die beiden Effekte gleichen sich bei 4° C gerade aus, wodurch das Dichtemaximum entsteht. Makroskopisch ist dies verstanden [23], allerdings fehlt bis heute eine Erklärung der mikroskopischen physikalischen Natur der strukturbildenden Kräfte [3].

Um diese Ursache aufklären zu können, benötigt man einen detaillierten Einblick in die Elektronendichte eines gelösten Wassermoleküls, was lediglich durch QM Beschreibungen möglich ist. Da man aber dazu auch Ensembles flüssigen Wassers bei p_0 und Temperaturen T aus dem in Abb. 1.1 gezeigten Bereich benötigt, scheiden QM Beschreibungen aufgrund ihres numerischen Aufwandes aus.

Einen Ausweg aus diesem Dilemma bietet die erwähnte DFT/PMM Hybridmethode [9] (siehe Abschnitt 1.2.3), bei der man ein einzelnes Wassermolekül als QM-, beziehungsweise DFT-Fragment auswählen und die wässrige Umgebung durch ein PMM Kraftfeld beschreiben kann. Durch die vereinfachte PMM Beschreibung der Umgebung können Systeme behandelt werden, die viele Tausend H_2O Moleküle bei klar definierten thermodynamischen Bedingungen umfassen.

1.1.2 Parametrisierung von Wassermodellen

Schropp und Tavan [16, 17] konnten anhand von DFT/MM Rechnungen zwei wichtige Erkenntnisse über die Polarisierbarkeit gelöster Wassermoleküle gewinnen. Dabei wählten sie ein H_2O Molekül aus einem flüssigen Ensemble von Wassermolekülen als DFT Fragment und näherten den großen Rest des umgebenden Wassers durch die bekannten MM Modellpotentiale TIP3P [24], SPC/E [25] und TIP4P [24]. Diese DFT/MM Ergebnisse stellen wahrscheinlich das bislang genaueste verfügbare Modell eines in der Flüssigkeit gelösten Wassermoleküls dar und bilden somit eine gute Vorlage für ein zu parametrisierendes PMM Modellpotential.

[2] http://www1.lsbu.ac.uk/water/ice.html

Der Entwurf eines neuen DFT/PMM Hybridverfahrens [9] und seine Realisierung in Form der Kombination CPMD/IPHIGENIE ermöglichte nun die Erweiterung dieses DFT/MM Ansatzes zu einer selbst-konsistenten DFT/PMM Parametrisierungsstrategie für PMM Modellpotentiale von Lösungsmittelmolekülen. Die zu entwickelnden Modellpotentiale sollten sich dabei so genau wie möglich am Vorbild des DFT-Fragments orientieren.

Der elektronischen Polarisierbarkeit α der PMM Modellpotentiale kommt dabei eine entscheidende Bedeutung zu. Da nicht-polarisierbare MM Modelle Polarisationseffekte lediglich im statistischen Mittel durch einen erhöhten statischen Dipol erfassen, sind sie an homogene Systeme bei bestimmten thermodynamischen Bedingungen gebunden und nicht in andere Umgebungen bzw. auf andere Bedingungen transferierbar. Als Beispiel sei die Arbeit von Klaehn et al. genannt, in der durch DFT/MM Rechnungen die Infrarotspektren von gelösten Phosphatanionen bestimmt wurden [26], wobei als Lösungsmittelmodell das sehr einfache MM Modell TIP3P [24] verwendet wurde. Die Defizite des auf diese Weise berechneten Spektrums wurden unter anderem auf die fehlende Polarisierbarkeit der Lösungsmittelmoleküle zurückgeführt [26, 27].

Man sollte nun annehmen, dass PMM Modellpotentiale für α den experimentell in der Gasphase bestimmten Wert [28] α^{g} einsetzen sollten. Stattdessen passen beispielsweise die PMM Wassermodelle SWM4-DP [29] und SWM4-NDP [30] die dort durch einen Drude-Oszillator dargestellte Polarisierbarkeit α empirisch auf deutlich kleinere Werte an, damit die experimentell bekannte Dielektrizitätskonstante reproduziert werden kann.

Wie Schropp und Tavan durch ihre DFT/MM Hybridrechnungen festgestellt haben [16], wird diese Reduktion von α bei Modellen mit induzierbaren Punktdipolen durch die starke Inhomogenität des elektrischen Feldes, welches benachbarte Wassermoleküle im Volumen eines gelösten H_2O Moleküls erzeugen, erzwungen, weil dieses Feld am Ort des Sauerstoffatoms um etwa 40 % größer ist als im Volumenmittel. Quantenmechanisch zählt für die Polarisierung jedoch das Volumenmittel, das durch die den Kern umgebende ausgedehnte Elektronendichte vorgenommen wird. Dagegen wird bei punktpolarisierbaren PMM Modellen nur die zu große Feldstärke im Zentrum des Moleküls berücksichtigt, was bei Verwendung von α^{g} zu einer Überschätzung der Polarisation in der simulierten flüssigen Phase und damit der Dielektrizitätskonstante führt. Diese Einsicht [16] begründet im Nachhinein die empirische Reduktion von α bei den SWM4 Modellen.

Durch weitere DFT/MM Hybridrechnungen konnten Schropp und Tavan anschließend zeigen [17], dass sich die Polarisierbarkeit von Wassermolekülen beim Transfer aus der Gasphase in die Flüssigkeit nicht ändert, obwohl sich ihre Geometrie dabei, wie wir gleich genauer sehen werden, relativ stark ändert.

Aus diesen Ergebnissen haben Schropp und Tavan für die Konstruktion von PMM Wassermodellen die theoretisch begründeten Vorschläge abgeleitet, man solle für das molekulare Dipolmoment den experimentellen Gasphase-Wert μ^{g} und, bei Modellierung der Polarisation durch einen Punktdipol, eine im Vergleich zu α^{g} um 40 % verminderte Polarisierbarkeit verwenden.

Alternativ dazu erlaubt es jedoch der Einsatz polarisierbarer Gaußscher Dipolverteilungen, welche eine ausgedehnte polarisierbare Elektronendichte modellieren, auch bei PMM Be-

schreibungen explizit über das polarisierende Feld mitteln und für α den Wert α^{g} verwenden. Mit der Breite σ der Gaußschen induzierten Dipolverteilung erhält man dann einen zusätzlichen Parameter, der geeignet gewählt werden muss [16]. Wir sind im Rahmen unserer eigenen Konstruktion von PMM Modellen diesen Vorschlägen gefolgt [5, 8].

PMM Modelle bieten darüber hinaus den Vorteil, dass sie im feldfreien Fall lediglich das isolierte Wassermolekül darstellen müssen, dessen Eigenschaften, wie etwa sein Dipolmoment und seine Polarisierbarkeit, zum Teil experimentell gut bekannt sind.

Eigenschaften des Wassermoleküls

Die Dreiecksform der molekularen Geometrie $\mathbf{G}_{\mathrm{m}}^{\mathrm{g}}$ eines H_2O Moleküls in der Gasphase konnte 1932 erstmals experimentell nachgewiesen werden, indem das Rotationsschwingungsspektrum von Wasserdampf vermessen wurde [31]. Die genauen Werte für die Bindungslänge l_{OH} und den Bindungswinkel φ_{HOH} wurden 1956 von Benedict, Gailar und Plyler mit 104.52° und 0.9572 Å angegeben [32].

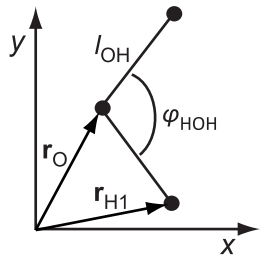


Abbildung 1.2: Die molekulare Geometrie $\mathbf{G}_{\mathrm{m}}$ des Wassermoleküls ist durch den Abstand zwischen Sauerstoff und Wasserstoff l_{OH} und den Winkel φ_{HOH} charakterisiert. Ein internes Koordinatensystem kann über die Winkelhalbierende des HOH-Dreiecks (x-Achse) und die Verbindungslinie zwischen den Wasserstoffen (y-Achse) definiert werden.

In Abbildung 1.2 ist die molekulare Geometrie $\mathbf{G}_{\mathrm{m}}$ des H_2O Moleküls dargestellt und ein internes Koordinatensystem definiert, welches so gewählt ist, dass die Winkelhalbierende des HOH-Dreiecks in x-Richtung und die Verbindung zwischen den beiden Wasserstoffen in y-Richtung zeigt. Die z-Achse zeigt entsprechend aus der Molekülebene heraus. Diese Wahl der Geometrie wurde bei allen im Rahmen dieser Dissertation vorgestellten Veröffentlichungen verwendet.

Geht man von isolierten Molekülen in der Gasphase zu Clustern mehrerer Moleküle oder zu Wassermolekülen in Lösung über, so ändert sich $\mathbf{G}_{\mathrm{m}}$ durch den Einfluss benachbarter Teilchen. Die molekulare Geometrie $\mathbf{G}_{\mathrm{m}}^{\mathrm{l}}$ von H_2O in Lösung wurde 1982 von Thiessen und Narten vermessen und mit $l_{\mathrm{OH}} = 0.968$ Å und $\varphi_{\mathrm{HOH}} = 105.3°$ angegeben [33]. Der Einfluss des elektrischen Feldes in flüssigem Wasser auf das Wassermolekül resultiert also in einer leichten Vergrößerung der Bindungslänge und Aufweitung des Bindungswinkels.

In der Gasphase hat ein H_2O Molekül das elektrische Dipolmoment $\mu^{\mathrm{g}} = 1.855$ D, wie 1973

von Clough et. al. [34] mittels elektronischer Resonanz-Spektroskopie gemessen und im gleichen Jahr durch Dyke und Muenter [35] bestätigt wurde.

Aufgrund seiner elektronischen Polarisierbarkeit verändert sich das Dipolmoment eines von anderen Molekülen umgebenen Wassermoleküls durch die von diesen erzeugten elektrischen Felder. Die Polarisierbarkeit ist bei Molekülen üblicherweise richtungsabhängig und wird durch den Polarisierbarkeitstensor $\boldsymbol{\alpha}$ angegeben. Der Polarisierbarkeitstensor des H_2O Moleküls wurde 1977 von Murphy in der Gasphase durch Raman-Streuexperimente vermessen [28]. Die gemessenen Werte waren $\alpha_{xx} = 1.47$ Å^3, $\alpha_{yy} = 1.53$ Å^3, und $\alpha_{zz} = 1.42$ Å^3. Die Polarisierbarkeit von Wasser ist also annähernd isotrop ist und kann durch den Skalar $\alpha^{\mathrm{g}} \equiv 1.47$ Å^3 approximiert werden.

Für das durch Polarisation im Wasserdimer geänderte Dipolmoment wurde experimentell der Wert[3] von 2.1 D bestimmt [36]. Dagegen gibt es für den mittleren Dipol eines gelösten Wassermoleküls keine gesicherten experimentellen Messdaten. Ab initio MD Rechnungen sagen hohe Werte von bis zu 2.9 D vorher [37, 38], klassische (P)MM-MD Rechnungen gehen eher von Werten zwischen 2.4 D und 2.6 D aus [29, 39]. Sicher ist, dass das mittlere elektrische Feld in Lösung in positive x-Richtung zeigt.

Der Quadrupolmoment-Tensor des Wassermoleküls bezüglich des Schwerpunkts eines H_2O Moleküls ist ebenfalls nur in der Gasphase bekannt [40]. Es wurde die für die Komponenten des Quadrupoltensors $\mathbf{Q}$ die Werte $Q_{xx} = -0.13$ DÅ, $Q_{yy} = 2.63$ DÅ und$Q_{zz} = -2.50$ DÅ- durch Zeeman Spektroskopie ermittelt.

Die Solvatstruktur von Wasser

Eine wichtige Eigenschaft des Wassers ist seine Solvatstruktur, die durch abstandsabhängige radiale Verteilungsfunktionen $g_{ij}(r)$ angegeben wird. Dabei charakterisieren i und j die Atomsorten (H, O), deren Abstände r betrachtet werden. Diese $g_{ij}(r)$ messen die Häufigkeit, mit der man ausgehend von einem Teilchen im Abstand r ein weiteres findet, bezogen auf die als konstant angenommene Dichte. Durch diese Normierung sind die $g_{\mathrm{ij}}(r)$ dimensionslos und haben für große Abstände den Grenzwert eins, da die Teilchen dann unkorreliert sind.

Radiale Verteilungsfunktionen können nicht direkt experimentell gemessen werden.[4] Gemessen werden bei Röntgen- oder Neutronenstreuungs-Experimenten die Intensitäten der Streuamplitude, abhängig vom Impulsübertrag. Dabei geht jedoch die Phaseninformation verloren. Die Rückschlüsse auf die radiale Verteilungsfunktionen von Wasser sind deshalb mit einiger Unsicherheit behaftet, wie große Unterschiede der veröffentlichten Daten deutlich machen [41, 42, 43, 44]. Durch neue und bessere Messmethoden, auf die hier nicht näher eingegangen werden soll, konnte die Unsicherheit der Messdaten erheblich eingeschränkt werden [45, 46]. Aus diesem Grund, und da speziell die Sauerstoff-Sauerstoff Verteilungsfunktion für die Parametrisierung und Evaluation von Wassermodellen eine große Rolle spielt, werden in Abbildung 1.3 zwei, in jüngerer Vergangenheit veröffentlichte, Messkurven gezeigt [45, 46].

[3] Dies ist ein Mittelwert zwischen Donor und Akzeptor.

[4] Diese Verteilungsfunktionen werden im nachfolgenden Text dennoch als experimentell bezeichnet.

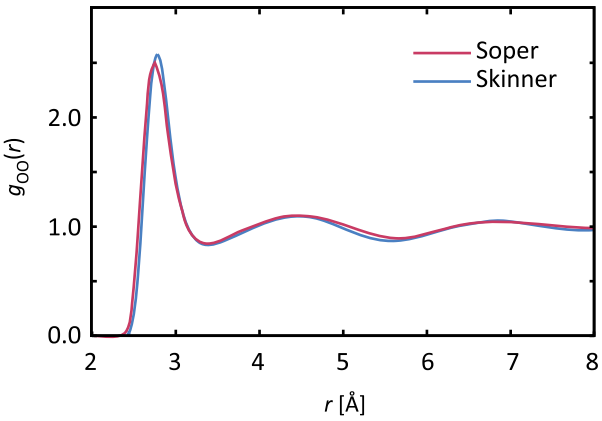


Abbildung 1.3: Eine von Skinner et al. [46] durch Röntgenstreuung gemessene und eine von Soper et al. [45] durch Neutronenstreuung bestimmte Sauerstoff-Sauerstoff Verteilungsfunktion im Bereich [2, 8] Å. Die Nahordnung in Form von drei Maxima ist ebenso wie der Grenzwert von 1 für große Abstände gut zu erkennen.

Verglichen werden eine durch Röntgenstreuung (blau) und eine durch Neutronenstreuung (rot) erfasste Verteilungsfunktion bei Raumtemperatur (298 K) [45, 46]. Beide Messkurven, welche auch im Rahmen unserer DFT/PMM Parametrisierung verwendet werden [5, 8], verlaufen sehr ähnlich. Leichte Unterschiede sind lediglich im Bereich des ersten Maximums des zweiten Minimums auszumachen. Deutlich wird, dass sich Wassermoleküle nicht näher als 2 Å kommen. Dies liegt an der Pauli-Repulsion besetzter Elektronenschalen (siehe Abschnitt 1.2.1). Die Nahordnung erstreckt sich ungefähr bis 10 Å und umfasst etwa drei Maxima, die auch oft Solvatisierungs-Schalen genannt werden.

1.2 Methoden

Unsere DFT/PMM Parametrisierungsstrategie, beruht, wie oben erwähnt wurde, auf der Verwendung des parallelisierten PMM-MD Simulationsprogramm-Pakets IPHIGENIE [10, 13, 14] und seiner neuartigen Kopplung [9] mit dem gitterbasierten DFT Programmpaket CPMD [12]. Es sollen nun die grundlegende Konzepte dieser Verfahren sowie ein Überblick über einige verbreitete polarisierbare Wassermodelle skizziert werden.

1.2.1 MM-MD Simulationen

Molekulardynamik-(MD) Simulationen behandeln die N Atome eines Systems als klassische Punktteilchen der Massen m_i an den Orten $\mathbf{r}_i$ ($i \in 1, ..., N$), wobei die Orte zum Konfigurationsvektor $\mathbf{R} \equiv \{\mathbf{r}_1, ...\mathbf{r}_N\}$ zusammengefasst werden. Diese Punktteilchen werden unter periodischen Randbedingungen simuliert, um Randeffekte zu vermeiden und um den Druck kontrollieren zu können. Wenn man die bei Gittersummenmethoden zur Behandlung der langreichweitigen Elektrostatik möglichen Periodizitätsartefakte vermeiden will, so kann man stattdessen auch, wie in IPHIGENIE implementiert ist, toroidale Randbedingungen[47] unter

Beachtung der *minimum image convention* mit einem Reaktionsfeldverfahren kombinieren [10, 14].

Ein MM-Kraftfeld ist eine analytische Energiefunktion $E_{\mathrm{MM}}(\mathbf{R})$, deren negativer Gradient bezüglich der Atomkoordinaten die Kraft $\mathbf{F}_i = -\nabla_i E_{\mathrm{MM}}(\mathbf{R})$ auf das jeweilige Atom i ergibt und daher die numerische Integration der Newtonschen Bewegungsgleichungen ermöglicht. Die übliche zeitliche Schrittweite Δt einer solchen Verlet-Integration [48] muss klein genug gewählt werden, um die Freiheitsgrade glatt abzutasten. Üblicherweise wird für Δt eine Femtosekunde gewählt.

Die MM-Kraftfelder für Wasser modellieren die Moleküle zumeist als starre Körper, weil schon die energetisch niedrigste Schwingungsmode des H_2O Moleküls bei 300 K im quantenmechanischen Schwingungsgrundzustand eingefroren ist. Die starre Geometrie kann mithilfe von Algorithmen wie SHAKE [49], SETTLE [50], LINCS [51] oder M-SHAKE [52] gewährleistet werden.

MM-Kraftfelder für Wasser kennen daher nur intermolekulare Wechselwirkungen, die sogenannten Wechselwirkungen $E_{\mathrm{nb}}(\mathbf{R})$ nicht gebundener Atome. Sie werden auf Paarwechselwirkungen beschränkt, die daher Funktionen des Abstands

$$r_{ij} = |\mathbf{r}_i - \mathbf{r}_j| \tag{1.1}$$

zwischen zwei Atomen i und j sind. E_{nb} besteht aus der Van der Waals Wechselwirkung E_{vdW} und der elektrostatischen Wechselwirkung E_{elstat}.

Van der Waals Wechselwirkung

E_{vdW} beschreibt die sehr kurzreichweitige Pauli-Repulsion zweier Atome, die durch Abstoßung besetzter Elektronenschalen entsteht, und eine attraktive Wechselwirkung, die sogenannte Dispersion [53]. Hervorgerufen wird die Dispersion durch Fluktuationen der Elektronenhüllen der Atome in Bezug zu den Atomkernen. Während die Abstandsabhängigkeit der

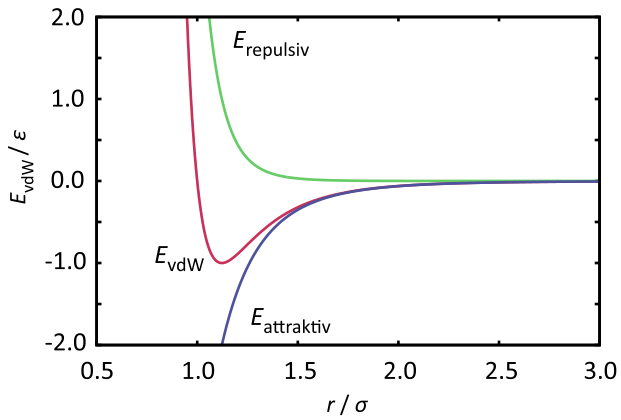


Abbildung 1.4: Das Van der Waals Potential (rot) besteht aus einem attraktiven Anteil (blau) und einem repulsiven Teil (grün). Für zwei Atome gibt der Parameter ϵ gibt die Tiefe des Minimums der Funktion (1.2) und σ deren Nulldurchgang an.

Dispersion sich durch eine einfache Herleitung [54] mit der London-Formel als proportional zu r_{ij}^{-6} angeben lässt, muss die Form des repulsiven Terms empirisch gewählt werden.

Für MM-Kraftfelder wird hierfür üblicherweise eine Abstandsabhängigkeit von r_{ij}^{-12} angenommen, wodurch sich das 12-6 Lennard-Jones Potential

$$E_{\text{vdW}}(\mathbf{R}) = E_{\text{LJ}}(\mathbf{R}) = \frac{1}{2}\sum_{i=1}^{N}\sum_{j>i}^{N} 4\epsilon\left[\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^{6}\right] \tag{1.2}$$

ergibt [55]. Wie in Abbildung 1.4 gezeigt wird, gibt der Parameter ϵ die Tiefe des Minimums der Funktion (1.2) und σ deren Nulldurchgang an. Die kurze Reichweite des Van der Waals Potentials hat zur Folge, dass seine Wirkung auf die Dynamik der Atome schon bei kurzen Atomabständen oberhalb von 10 Å relativ klein ist und daher ohne allzu große algorithmische Artefakte vernachlässigt wird. Üblicherweise wird der energetische Effekt der vernachlässigten Wechselwirkungen durch eine Molekularfeldnäherung beschrieben [47].

Im Jahre 1947 hat Buckingham eine alternative Form für das Van der Waals Potential, das nach ihm benannte Potential

$$E_{\text{vdW}}(\mathbf{R}) = E_{\text{B}}(\mathbf{R}) = \frac{1}{2}\sum_{i=1}^{N}\sum_{j>i}^{N}\left[a\mathrm{e}^{-cr_{ij}} - \frac{b}{r_{ij}^{6}}\right] \tag{1.3}$$

vorgestellt [56], die den abstoßenden Anteil durch eine Exponentialfunktion beschreibt, während der attraktive Anteil weiterhin die gleiche Ortsabhängigkeit hat. Dadurch konnte er deutlich bessere Ergebnisse bei der analytischen Berechnung des zweiten Virialkoeffizienten von Neon und Argon erzielen.

In Abbildung 1.5 kann man die Unterschiede zwischen den beiden Varianten des Van der Waals Potentials erkennen. Während das 12-6 Lennard-Jones Potential eine relative steile repulsive Flanke hat, verläuft der durch eine Exponentialfunktion beschriebene repulsive Anteil des Buckingham Potentials etwas flacher.

Zur Veranschaulichung der Abstände ist in Abbildung 1.5 die experimentelle radiale Sauerstoff-Sauerstoff Verteilungsfunktion von Soper et al. [45], die schon aus Abschnitt 1.1.2 bekannt ist, ebenfalls eingezeichnet. Die Form des repulsiven Teils des Van der Waals Potentials beeinflusst aufgrund seiner kurzen Reichweite vor allem die Position und Form des ersten Maximums der Paarverteilungsfunktion. Diese wird durch den etwas flacheren repulsiven Teil des Buckingham Potentials deutlich besser beschrieben, als durch die steile Flanke des 12-6 Lennard-Jones Potentials [42, 57, 58]. Aus diesem Grund entschieden wir uns, im Rahmen der DFT/PMM Parametrisierungsstrategie, für die Beschreibung des Van der Waals Potentials durch ein Buckingham Potential am Ort des Sauerstoffs.

Elektrostatische Wechselwirkung

Die Elektrostatik wird in nicht-polarisierbaren MM-Kraftfeldern durch statische Partialladungen an Orten von Atomen sowie an masselosen Ladungspunkten im Molekül modelliert, deren

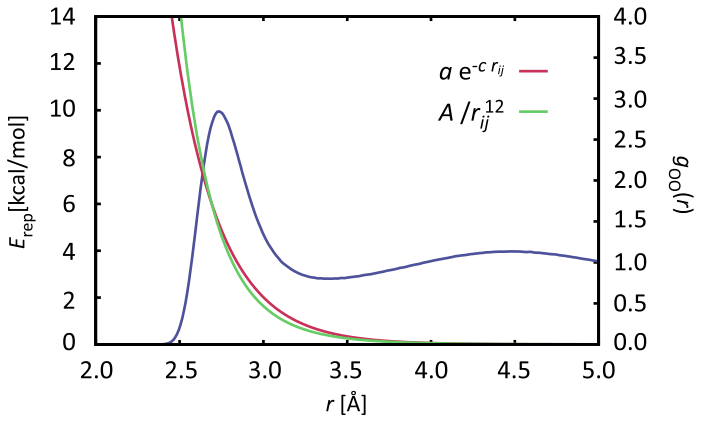


Abbildung 1.5: Vergleich der repulsiven Anteile eines 12-6 Van der Waals Potentials (grün) und eines Buckingham Potentials (rot). Um eine bessere Vorstellung der Abstände r_{ij} zu bekommen ist die experimentell bestimmte Verteilungsfunktion auf der rechten y-Achse eingezeichnet. Die Parameter für beide Potentiale wurden Referenz [5] entnommen, die im Rahmen dieser Arbeit vorgestellt wird.

Wechselwirkung durch die Coulomb-Energie

$$E_{\text{elstat}}(\mathbf{R}) = E_{\text{C}}(\mathbf{R}) = \frac{1}{2}\sum_{i=1}^{N}\sum_{j>i}^{N}\frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \tag{1.4}$$

beschrieben wird. Im Gegensatz zur Van der Waals Wechselwirkung ist die Coulomb Wechselwirkung langreichweitig, und kann nicht ab einem bestimmten Abstand vernachlässigt werden [59, 60]. Da eine explizite Auswertung der Coulomb-Summe eine Skalierung des Rechenaufwandes von N^2 nach sich zieht, müssen geeignete Näherungsverfahren angewendet werden, um die langreichweitige Coulomb-Wechselwirkung angemessen zu beschreiben.

Die von Ewald [61] für die Festkörperphysik entwickelte Gittersummenmethode und weitere darauf aufbauende Verfahren [62] nutzen die bei MD Simulationen verwendeten periodischen Randbedingungen aus. Sie skalieren mit mit $N\log(N)$ [62]. Ein Nachteil sind mögliche Periodizitätsartefakte [63, 64, 65, 66, 67], welche durch die künstlich eingeführte Periodizität des elektrostatischen Potentials erzeugt werden könne.

Um Periodizitätsartefakte zu vermeiden lässt sich die langreichweitige Elektrostatik auch durch die Kombination [14] der schnellen Multipolmethode SAMM [10, 68, 69] mit einem Reaktionsfeldverfahren (RF) berechnen, welche im MD-Simulationsprogramm IPHIGENIE implementiert ist. Der komplexe SAMM/RF Algorithmus, der während der Laufzeit meiner Dissertation (i) auf die Dispersionswechselwirkung erweitert und (ii) dessen Genauigkeit und Effizienz bedeutend gesteigert wurde (Ref. 10 und die laufende Dissertation von K. Lorenzen), so dass die umfangreichen Simulationen [5, 7] mit dem TL6P Wassermodell durchführbar wurden, kann hier nicht ansatzweise skizziert werden. Es müssen daher die angeführten Verweise genügen.

1.2.2 Polarisierbarkeit in PMM-MD Simulationen

Der große Vorteil von MM-Kraftfeldern liegt darin, dass durch die vereinfachten Potentialfunktionen Systeme, die einige hunderttausend Atome umfassen, auf Zeitskalen von Nanosekunden simuliert werden können. Der Nachteil ist die eingesetzte Molekularfeldnäherung für Effekte der elektronische Polarisierbarkeit α, die nur auf homogene Systeme und bestimmte thermodynamische Bedingungen mit guter Genauigkeit anwendbar ist. Die Transferierbarkeit eines Wassermodells in andere Umgebungen und Bedingungen kann auf diese Weise nicht gewährleistet werden [70, 71, 72]. Daher muss die Polarisierbarkeit so effizient, aber auch so genau wie möglich in MM-Kraftfelder eingebunden werden, die dadurch zu PMM-Kraftfeldern werden.

Im Rahmen der linearen Antwortnäherung ist der induzierte Dipol proportional zum induzierenden elektrischen Feld. Es gilt also $\boldsymbol{\mu} = \alpha \mathbf{E}$. Für Wassermoleküle in wässriger Umgebung diese Näherung korrekt, wie DFT/MM Hybridrechnungen von Schropp und Tavan gezeigt haben [16]. Das polarisierende Feld in PMM-Kraftfeldern setzt sich also zusammen aus dem Feld der statischen Partialladungen, und dem Feld der polarisierbaren Dipole, welche selbstkonsistent bestimmt werden müssen. Die zugrundeliegenden Prinzipien [73, 74, 75], sollen hier skizziert werden, bevor in Abschnitt 1.2.3 eine Übersicht über polarisierbare Wassermodelle gegeben wird.

Induzierbare Dipole

Eine Methode die elektronische Polarisierbarkeit in MM-Kraftfelder einzubinden besteht darin, induzierbare Dipole (ID) an den Orten der Atome [76, 77, 78] oder entlang der kovalenten Bindungen [79] einzuführen. Dabei ist darauf zu achten, dass die sogenannte Polarisationskatastrophe verhindert wird, die dadurch zustande kommen kann, dass sich zwei induzierbare Dipole zu nahe kommen, sich immer weiter induzieren und somit divergieren [80, 81]. So kann durch geeignete Wahl der Van der Waals Parameter verhindert werden, dass sich zwei Atome zu nahe kommen [82]. Als Alternative kann man eine abstandsabhängige Dämpfung der Polarisierbarkeit α einführen [83, 84, 85]. Eleganter ist die Verwendung Gaußscher Dipol- oder Ladungsverteilungen [86], da deren Potentiale bei kleinen Atomabständen nicht divergieren.

Drude Oszillator

Um die Einbindung polarisierbarer Dipole in den Programmcode eines PMM-MD Programmes zu umgehen, lässt sich die sogenannte Drude-Oszillator (DO) Methode nutzen [87, 88, 89]. Bei dieser Methode wird eine masselose, zumeist negativ gewählte Ladung durch ein harmonisches Federpotential an das zu polarisierende Atom gebunden. Die Positionen dieser Hilfsteilchen in einer bestimmten Konfiguration $\mathbf{R}$ der Atome, müssen selbstkonsistent so bestimmt werden, dass ein Minimum der potentiellen Energie gefunden wird. Dadurch muss lediglich die Coulomb-Summe (1.4) um diese Hilfsteilchen erweitert werden, und es müssen keine Dipol-Dipol oder Dipol-Ladung Wechselwirkungen berechnet werden.

Im Vergleich zum Einsatz eines induzierten Dipols bewirkt die eine zusätzliche Ladung des DO Modells, dass der Rechenaufwand steigt. Aus den genannten Bequemlichkeitsgründen der leichteren Implementierung ist die DO Methode jedoch deutlich stärker verbreitet als die ID Methode. Die Anwendungen reichen von ionischen Kristallen [90, 91], über die Hydratisierung kleiner Ionen [92, 93], bis zu einfachen Flüssigkeiten [94, 95] und ganzen Proteinkraftfeldern [96]. Auch für QM/MM Systeme wurden DO Methoden verwendet [97].

Veränderliche (fluktuierende) Ladungen

Eine weitere Möglichkeit Polarisierbarkeit in MM-MD Simulationen zu integrieren besteht darin, eine Fluktuation der Partialladungen, abhängig vom externen elektrischen Feld oder Potential, zu gestatten (FQ-Methode). Die Werte der Partialladungen werden durch selbstkonsistente Minimierung der elektrostatischen Energie ermittelt. Die Coulomb-Summe (1.4) wird hierbei durch eine komplexere Entwicklung [98] ersetzt, die im Grenzfall großer Entfernungen in das Coulomb-Potential übergeht. Der Rechenaufwand der FQ Methode ist etwas größer als bei der ID Methode.

1.2.3 Modellpotentiale für Wasser

Seit den ersten Computersimulationen von flüssigem Wasser [99, 100] zu Beginn der 1970iger Jahre wurden fast unzählbar viele Modellpotentiale für H_2O veröffentlicht. Dieser Modelle lassen sich in etwa durch die Anzahl der Ortspunkte im Molekül, an denen Kräfte berechnet und ausgewertet werden müssen, klassifizieren.

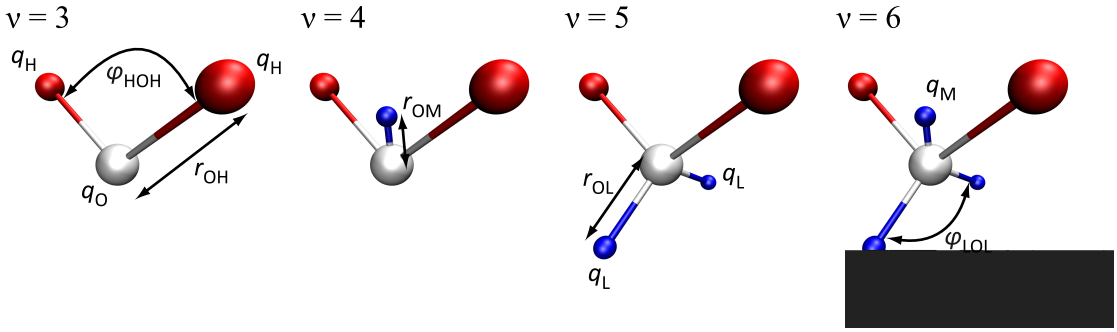


Abbildung 1.6: Einteilung der Wassermodelle in Drei-, Vier-, Fünf und Sechspunktmodelle, abhängig von der elektrostatischen Signatur und der damit verbundenen Anzahl $\nu \in [3, ..., 6]$ an Ortspunkten eines Modells, an denen Kräfte berechnet werden. Die positive Ladungen q_H befinden sich an den Orten der Wasserstoffe, während die Positionen der negativen Ladungen q_O, q_M und q_L je nach Modellklasse variieren.

Abbildung 1.6 verbildlicht die Einteilung der Wassermodelle in vier Gruppen. Die Gruppen sind nach der Anzahl $\nu \in [3, ..., 6]$ an Orten charakterisiert, an denen Kräfte berechnet werden. Die einfachsten Wassermodelle, die Dreipunktmodelle beschreiben die elektrostatische Signatur eines Wassermoleküls durch positive Partialladungen q_H an den Orten der Wasserstoffe und eine negative Partialladung q_O am Ort des Sauerstoffes. Auch das Van der Waals

Potential wird üblicherweise am Ort des Sauerstoffs ausgewertet. Das TIP3P Modell von Jorgensen [24] oder die beiden Modelle SPC und SPC/E von Berendsen [101, 102] sind solche MM-MD Dreipunktmodelle, die aufgrund ihrer Einfachheit sehr effizient simuliert werden können, und deshalb immer noch weite Verbreitung finden. Die Geometrie ist durch Angabe von r_{OH} und φ_{HOH} vollständig festgelegt. Wie in Abschnitt 1.1.2 erwähnt wurde, sind für diese Parameter experimentelle Werte in der Gasphase und der flüssigen Phase bekannt. q_{H} und q_{H} werden so gewählt, dass das Dipolmoment einen gewünschten Wert annimmt.

Vierpunktmodelle wie TIP4P[24], TIP4P/2005[103] oder TIP4Q [104], stellen die einfachsten Weiterentwicklungen von Dreipunktmodellen dar. Bei Vierpunktmodellen wird die negative Ladung vom Ort des Sauerstoffs um eine Strecke r_{OM} in Richtung der Wasserstoffe, auf einen masselosen Aufpunkt $\mathbf{r}_{\mathrm{M}}$, verschoben. Die Ladung q_{O} wird dadurch zu q_{M}. Durch den zusätzlichen Freiheitsgrad r_{OM} lassen sich höhere Multipolmomente anpassen, ohne das Dipolmoment zu verändern

Im Falle der Fünfpunktmodelle wie TIP5P [105] oder TIP5P/E [106] werden zwei masselose Ladungspunkte spiegelsymmetrisch in der $xz-$Ebene, senkrecht zur $xy-$Ebene, in der sich die Wasserstoffe und der Sauerstoff befinden, angeordnet (vgl. Abb. 1.2). Die beiden masselosen Ladungen q_{L} befinden sich an den Orten $\mathbf{r}_{\mathrm{L1}}$ und $\mathbf{r}_{\mathrm{L2}}$, und haben von $\mathbf{r}_{\mathrm{O}}$ den Abstand r_{OL}. Ein zusätzlicher Parameter ist der Winkel φ_{LOL}, der das Dreieck zwischen Sauerstoff und den beiden q_{L} definiert. Sechspunktmodelle [107] stellen eine einfache Kombination von Vier- und Fünfpunktmodellen dar. Der Parameterraum ist entsprechend erweitert.

Die bisherige Einteilung der Wassermodelle in Drei-, Vier-, Fünf und Sechspunktmodelle ist, angesichts der Vielfalt der gewählten Ansätze für Modellpotentiale des H_2O Moleküls, nicht vollständig. Generell wurden beim Wassermolekül beinahe alle erdenklichen Kombinationsmöglichkeiten und Parametersätze, seien es polarisierbare Van der Waals Potentiale [108], feldabhängige Polarisierbarkeit [109], induzierbare Punktdipole auf den Wasserstoffen [110], Ladungstransfer zwischen Wassermolekülen [111] oder beliebig komplexe Formen des Van der Waals Potentials [112], vorgestellt, obwohl die meisten dieser Modellannahmen fundierter physikalischer Grundlagen entbehren.

DFT/PMM Hybridmethoden und die Elektrostatische Signatur von H_2O

Die DFT basiert auf zwei Theoremen, die von Hohenberg und Kohn [113] und von Kohn und Sham [114] aufgestellt wurden. Sie ist die Grundlage eines weit verbreiteten numerischen QM Verfahrens zur Berechnung der Grundzustandseigenschaften eines wechselwirkenden Vielelektronensystems. Um einen umfassenden Überblick über die DFT zu erhalten, sei der Leser auf das Buch von Dreizler und Gross [115] verwiesen. Hier sei lediglich erwähnt, dass das erste Theorem die Verbindung zwischen der hoch-dimensionalen Wellenfunktion des Grundzustandes und der von nur drei Ortskoordinaten abhängenden Elektronendichte herstellt . Aufgrund der Reduktion der Freiheitsgrade stellt dieser Schritt für numerische Programme eine wichtige Vereinfachung dar. Die Berechnung der Grundzustandsenergie, der Elektronendichte und daraus folgender Größen, wie beispielsweise der Bindungslängen zwischen Atomen, sind dann in diversen Programmen implementiert. Wichtige Programmpackete im Bereich der Computerchemie sind beispielsweise Gaussian, CPMD oder CP2K [12, 116, 117].

All diese Programme sind aufgrund der Komplexität wechselwirkender Elektronensysteme für konkrete Berechnungen auf Näherungen angewiesen, von denen die wichtigste die *local density approximation* ist [118]. Bei Gittermethoden besteht eine weitere Näherung darin, nicht alle Elektronen, sondern lediglich die Valenzelektronen explizit zu berechnen. Die Kerne und die kernnahen Elektronen werden in sogenannten Pseudopotentialen zusammengefasst, die den abschirmenden Effekt dieser Elektronen auf die Kerne beschreiben sollen.

Bei meinen DFT/PMM Rechnungen habe ich das gitterbasierte DFT Programm CPMD [12] und die von Troullier und Martins entwickelten [119] Pseudopotentiale verwendet. Ferner habe ich das Funktional BP86 [120, 121] gewählt, obwohl bekannt ist, dass es einige Nachteile [27] gegenüber anderen Funktionalen, wie beispielsweise B3LYP [122], hat. Der Grund für diese Wahl war die Sicherstellung der Vergleichbarkeit mit den DFT/MM Rechnungen von Schropp und Tavan, die seinerzeit das selbe Funktional und Pseudopotential verwendet hatten [16, 17].

QM/(P)MM Hybridmethoden bilden, seit der Veröffentlichung der Methode von Warshel und Karplus [123], ein weites Feld. Hier soll lediglich die im Rahmen dieser Arbeit verwendete DFT/PMM Kopplung thematisiert werden . Bei Interesse sei der Leser auf den umfangreichen Übersichtsartikel [124] zu QM/(P)MM Methoden verwiesen.

Im DFT/PMM Programm CPMD/IPHIGENIE wird die Van der Waals Wechselwirkung überall vermittels Gleichung (1.2) oder (1.3) berechnet. Dagegen muss die elektrostatische Wechselwirkung vom PMM- in das DFT-Fragment importiert werden und umgekehrt. Dazu müssen die potentiellen Energien, welche die Atome des PMM-Fragments an den Gitterpunkten des DFT-Fragments erzeugen, berechnet werden. Anschließend muss die SCF Iteration des DFT-Fragments bis zur Konvergenz durchgeführt werden, um aus der Elektronendichte des DFT-Fragments seine Rückwirkung auf das PMM-Fragment zu bestimmen.

Die hohe Anzahl an Gitterpunkten, der damit verbundene Aufwand zur Berechnung der elektrostatischen Wechselwirkungen zwischen den Fragmenten und die gleichzeitige Verwendung zweier iterativer, selbst-konsistenter Vorgänge, der DFT-SCF Iteration und der Iteration der polarisierbaren Dipole, machen einen effizienten DFT/PMM Algorithmus zu einer schwierigen Aufgabe. Wie jedoch von Eichinger et al. gezeigt wurde [15], kann man die elektrostatische Wechselwirkung im Rahmen des SAMM/RF Algorithmus [10, 13, 14, 68, 69] sehr effizient beschreiben. Eine weitere Effizienzsteigerung dieses Imports und Exports der elektrostatischen Potentiale in das DFT Fragment und die effiziente Einbindung der Polarisierbarkeit wurde von Schwörer et al. [9] entworfen.

Abbildung 1.7 skizziert einen Ausschnitt eines DFT/PMM Systems, bei dem ein Wassermolekül als DFT Fragment gewählt wurde, während die restlichen PMM Moleküle das Lösungsmittel bilden. Bei unserer DFT/PMM Optimierung von PMM Wassermodellen [5, 8] wurden solche Systeme verwendet.

Wie im vorigen Abschnitt erklärt wurde, wird im feldfreien Fall die elektrostatische Signatur eines PMM Wassermodells durch die Partialladungen und deren Lage im moleküleigenen Koordinatensystem bestimmt (vergleiche Abbildung 1.6). Diese Ladungen werden so gewählt, dass der statische Dipol einen bestimmten Wert annimmt. Die Quadrupol- und alle höheren elektrostatischen Momente sind dann durch die dadurch noch nicht festgelegten Größen

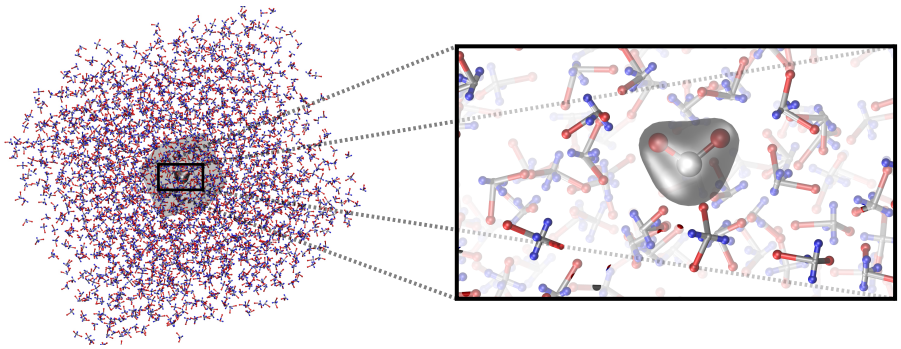


Abbildung 1.7: Quantenmechanisch beschriebenes H_2O Molekül eingebettet in eine PMM Wasserumgebung. Die Elektronendichte $\rho(\mathbf{r})$ ist durch eine gräulich hinterlegte Oberfläche angedeutet.

und Lagen der Ladungen bestimmt. Experimentelle Befunde für die höheren Momente des Wassermoleküls in Lösung liegen nicht vor. Durch DFT/PMM Rechnungen können sie aber bestimmt werden [5, 8].

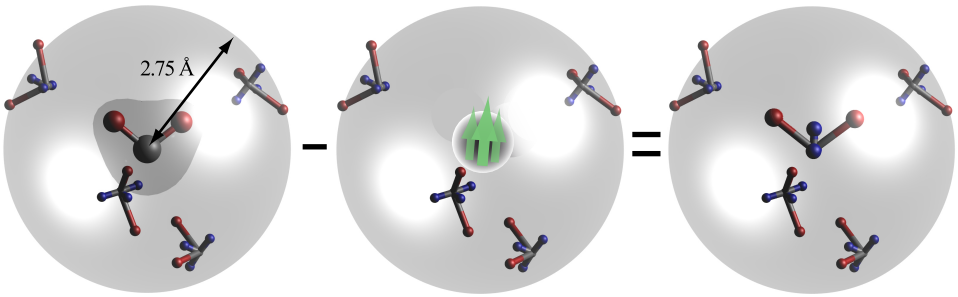


Abbildung 1.8: Grundprinzip der Optimierung der elektrostatischen Signatur des PMM Modells TL6P [5] und seiner Vorgänger. Das elektrostatische Potential eines DFT Wassermoleküls wird auf einer Kugel mit Radius 2.75Å ausgewertet (linke Seite). Nachdem von diesem Potential der Teil, der vom induzierten Dipolmoment (grüne Pfeile) herrührt abgezogen wurde, erhält man das zum statischen Anteil der Ladungsverteilung eines Wassermoleküls in Lösung gehörige Potential. Die Parameter der elektrostatischen Signatur (vgl. Abb. 1.6) können dann so gewählt werden, dass sie diesen statischen Anteil so gut wie möglich reproduziert.

Alle Multipolmomente sind im elektrostatischen Potential, welches ein gelöstes Wassermolekül in seiner Umgebung erzeugt, kodiert. Das Ziel sollte es deshalb sein, den statischen Anteil dieses Potentials so gut wie möglich durch die Partialladungen eines PMM Modells abzubilden [5, 8]. Wie in Abbildung **??** skizziert wird, muss dazu vom Potential, das durch die Ladungsverteilung des DFT-Fragments erzeugt wird, das Potential des des induzierten Dipols abgezogen werden, um so den statischen Anteil des Potentials zu erhalten. Die elektrostatische Signatur eines PMM Modells kann dann so gewählt werden, dass dieser optimal reproduziert wird.

1.3 Ziele und Gliederung

Mein Forschungsprojekt war durch den Sonderforschungsbereich 749 zur Erforschung der Dynamik und Intermediate molekularer Transformationen finanziert. Die elektronische Polarisierbarkeit der Lösungsmittel spielt bei solchen Prozessen eine wichtige Rolle. Daher sollten im Rahmen dieser Dissertation PMM Kraftfelder für das biologisch wichtigste Lösungsmittel Wasser entwickelt werden. Die technischen Voraussetzungen dafür waren vorhanden: Nach einer gründlichen Überarbeitung [9, 10, 11] des Programmpakets EGO [13, 14], das anschließend in in IPHIGENIE umbenannt wurde, ließen sich PMM Kraftfelder unter Verwendung von Gaußschen Dipolen und masselosen Ladungspunkten effizient behandeln.

Anknüpfend an die Arbeiten von Schropp und Tavan [16, 17] und aufbauend auf das weiterentwickelte DFT/PMM Programmpaket CPMD/IPHIGENIE [9] wird in der Publikation [8], die in Abschnitt 2.1 nachgedruckt ist, eine neue Methode zur Parametrisierung von PMM Wassermodellen vorgestellt, welche die Grundlage der gesamten Arbeit bildet. Dabei handelt es sich um ein auf DFT/PMM Rechnungen beruhendes, iteratives und selbstkonsistentes Verfahren zur Bestimmung derjenigen elektrostatischen Eigenschaften von PMM Modellen, welche durch die Vorgaben des Dipolmoments μ^{g} und der Polarisierbarkeit α^{g} in der Gasphase sowie der Flüssigphasengeometrie $\mathbf{G}_{\mathrm{m}}^{\mathrm{l}}$ noch nicht spezifiziert sind. Die wenigen Parameter des am Sauerstoff zentrierten Van der Waals Modellpotentials wurden jedoch durch PMM-MD Simulationen bei den Standardbedingungen $n(p_0, T_0)$ und T_0 an entsprechend wenige experimentell bekannte Größen wie die Solvatisierungsenthalpie oder den Druck p_0 angepasst. Die resultierenden3-, 4- und 5-Punkt PMM Wassermodelle wurden gründlich evaluiert, indem eine Vielzahl von Observablen bei T_0 und p_0 durch geeignete Simulationen berechnet wurden.

Die Evaluation der Modellpotentiale hat gezeigt, dass die aus der neuen Optimierungsmethode abgeleiteten PMM Kraftfelder den bislang besten, empirisch entwickelten Kraftfeldern zumindest ebenbürtig, in mancher Hinsicht aber auch überlegen waren. Insbesondere das 4- und das 5-Punktmodell lieferten gute bis sehr gute Ergebnisse für alle Observablen mit Ausnahme des thermischen Expansionskoeffizienten $\alpha_p(T_0)$, der die logarithmische Ableitung des Temperatur-Dichte Profils nach der Temperatur darstellt.

Eine zutreffende Vorhersage von $\alpha_p(T_0)$ war auch anderen Entwicklern von PMM Modellpotentialen für Wasser bis dahin nicht gelungen. Um zu herauszufinden, ob diese Schwachstelle durch Erhöhung der Modellkomplexität beseitigt werden kann entschieden wir uns, ein Sechspunktmodell zu berechnen. Seine Entwicklung, Evaluation und der Vergleich mit den Vorgänger-Modellen, sind Inhalt der Veröffentlichung [5], welche Kapitel 2.2 dieser Dissertation bildet. Hier zeigte sich, dass das 6-Punktmodell nicht nur alle bislang untersuchten Observablen bei T_0 und p_0 besser als seine Vorgänger beschrieb sondern vor allem auch $\alpha_p(T_0)$ sehr genau traf.

Aufgrund von Forderungen der Fachgutachter mussten wir die Evaluation des Sechspunktmodells und seiner Vorgänger um einige Observablen erweitern. Speziell die Vorhersagekraft des Modells außerhalb der flüssigen Phase sollte von uns dokumentiert werden. Daher beinhaltet die Supporting Information zu Veröffentlichung [5] auch Gasphasen- sowie Festkörpereigenschaften der Modelle.

Die korrekte Vorhersage auch von $\alpha_p(T_0)$ durch unser PMM 6-Punktmodell war der Anlass, das von den 4-, 5-, und 6-Punktmodellen vorhergesagte Temperatur-Dichte Profil durch 20 ns Replika-Austauschsimulationen mit außergewöhnlich großer statistischer Genauigkeit zu bestimmen. Hier lieferte unser 6-Punktmodell im Gegensatz zu seinen Vorgängern eine hervorragende Beschreibung, die, zusammen mit der Diskussion der daraus abgeleiteten mikroskopischen physikalischen Ursachen der Dichteanomalie, in einer weiteren Veröffentlichung zusammengefasst sind [7]. Diese ist in Kapitel 2.3 abgedruckt.

2 Entwicklung einer PMM-gestützten Optimierungsmethode für Wassermoleküle

Zunächst wird eine Optimierungsmethode eingeführt, welche in der Lage ist aus DFT/PMM Rechnung und PMM Simulationen optimale PMM Modellpotentiale für Wassermoleküle zu generieren. Entsprechende PMM 3-, 4-, und 5-Punktmodelle werden einer gründlichen Evaluation bei den Standardbedingungen $T_0 = 300$ K und $p_0 = 1$ bar unterzogen.

2.1 DFT/PMM Optimierung von Wassermodellen

Die nachfolgende Publikation[1]

> „Polarizable Water Models from Mixed Computational and Empirical Optimization“, Philipp Tröster, Konstantin Lorenzen, Magnus Schwörer, and Paul Tavan, *J. Phys. Chem. B*, **117**, 9486-9500, (2013),

die von mir zusammen mit Konstantin Lorenzen, Magnus Schwörer, und Paul Tavan verfasst wurde, beinhaltet die Entwicklung einer selbstkonsistenten Optimierungsmethode für PMM Wassermodelle. Eine umfangreiche Evaluation der resultierenden 3-, 4- und 5-Punktmodelle zeigt, dass durch die DFT/PMM Optimierungsmethode und den damit verbundenen Einblick in die elektronische Ladungsverteilung eines gelösten Moleküls, qualitativ hochwertige Modellpotentiale entwickelt werden können.

[1] Reproduced with permission from the Journal of Physical Chemistry, **117**, 9486-9500, 2013. Copyright 2013 American Chemical Society.

THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

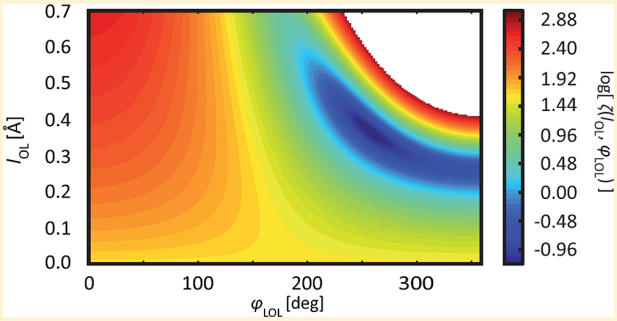
Polarizable Water Models from Mixed Computational and Empirical Optimization

Philipp Tröster, Konstantin Lorenzen, Magnus Schwörer, and Paul Tavan*

Lehrstuhl für Biomolekulare Optik, Fakultät für Physik, Ludwig-Maximilians-Universität München, Oettingenstrasse 67, D-80538 München, Germany

Supporting Information

ABSTRACT: Here we suggest a mixed computational and empirical approach serving to optimize the parameters of complex and polarizable molecular mechanics (PMM) models for complicated liquids. The computational part of the parameter optimization relies on hybrid calculations combining density functional theory (DFT) for a solute molecule with a PMM treatment of its solvent environment at well-defined thermodynamic conditions. As an application we have developed PMM models for water featuring ν = 3, 4, and 5 points of force action, a Gaussian inducible dipole and a Buckingham potential at the oxygen, the experimental liquid phase geometry, the experimental gas phase polarizability α^{g}_{exp} = 1.47 Å ^{3}, and, for ν = 4 and 5, the gas phase value μ^{g}_{exp} = 1.855 D for the static dipole moment. The widths of the Gaussian dipoles and, for ν = 4 and 5, also the electrostatic geometries of these so-called TLνP models are derived from self-consistent DFT/PMM calculations, and the parameters of the Buckingham potentials (and the static TL3P dipole moment) are estimated from molecular dynamics (MD) simulations. The high quality of the resulting models is demonstrated for the observables targeted during optimization (potential energy per molecule, pressure, radial distribution functions) and a series of predicted properties (quadrupole moments, density at constant pressure, dielectric constant, diffusivity, viscosity, compressibility, heat capacity) at certain standard conditions. Remaining deficiencies and possible ways for their removal are discussed.



1. INTRODUCTION

Water is undoubtedly the most important liquid on earth, because life originates from aqueous solution. The properties and functions of biological macromolecules are shaped by this polar and polarizable solvent, which features many unusual properties.[1] Therefore, atomistic simulations of biomolecular systems,[2,3] which use so-called molecular mechanics (MM) force fields such as CHARMM,[4] Amber,[5] or Gromos[6] require model potentials for water.

Unfortunately, MD simulations of biomolecular systems usually employ extremely simplified model potentials for the water molecules, such as the "three-point transferable intermolecular potential" (TIP3P) of Jorgensen[7] or the various "simple point charge" (SPC, SPC/E) models of Berendsen,[8,9] although these models can hardly reproduce all important properties of the bulk liquid at once, which include, e.g., the local structure as represented by various radial distribution functions[10−14] (RDFs), the dielectric constant, the density maximum at 4 °C and ambient pressure, and so forth.[1,15]

Figure 1 illustrates how such a so-called "three-point model" simplifies the complex electrostatic signature of a water molecule, which is generated by the electron density surrounding the three nuclei at positions $\mathbf{r}_O$, $\mathbf{r}_{H1}$, and $\mathbf{r}_{H2}$, by assigning a negative partial charge q_O to $\mathbf{r}_O$ and positive partial charges $q_H = -q_O/2$ to $\mathbf{r}_{H1/2}$. The molecular geometry $\mathbf{G}_m$, as defined by the chosen bond lengths l_{OH} and bond angle φ_{HOH}, is usually assumed to be fixed. Then the absolute value μ selected for the dipole moment $\boldsymbol{\mu}$ fixes the partial charge q_O as well as the higher multipole moments of the water model. Lennard-Jones potentials[16,17] centered at $\mathbf{r}_O$ are additionally employed to model the Pauli repulsion and dispersion attraction acting between different molecules.

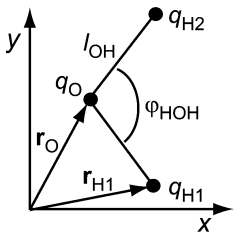


Figure 1. The molecular geometry $\mathbf{G}_m$ of H_2O is defined by the parameters l_{OH} and φ_{HOH}. In three-point models, the electrostatic signature of H_2O is generated by partial charges $q_O = -2q_H$, $q_H \equiv q_{H1} = q_{H2} > 0$, localized at the positions of the nuclei. The Cartesian axes are attached to the molecular plane as indicated.

Received: May 7, 2013
Revised: July 10, 2013
Published: July 11, 2013

   dx.doi.org/10.1021/jp404548k | J. Phys. Chem. B 2013, 117, 9486−9500

The key drawbacks of such three-point models are (i) the severe restrictions imposed to the higher multipole moments by the requirement that the partial charges q_i, $i \in \{O,H1,H2\}$, are located at the positions $\mathbf{r}_i$ of the nuclei and (ii) the neglect of the large polarizability α of the water molecule, which follows from the choices of fixed partial charges and of a rigid geometry $\mathbf{G}_m$. While the restrictions (i) are mainly responsible for the generally weak performance of three-point models concerning the local solvation structures, e.g., of small ions[18] or water molecules,[1,19] the neglect (ii) of the polarizability[20−27] leads to a poor transferability of such three-point models from the bulk at certain standard conditions (e.g., temperature T_0 = 300 K, pressure p_0 = 1 atm) toward other conditions[1] or into different environments, such as, for example, the interior of a protein.[28]

1.1. Four- and Five-Point Models. In view of these drawbacks there have been many suggestions for improved water models (for reviews see refs 1 and 29), which have been almost exclusively parametrized for and applied to MD simulation studies of the pure bulk liquid at standard conditions. Here, with the aim of remedying the restrictions, (i) additional massless points carrying fixed partial charges were introduced, leading to so-called four- and five-point models.[7,30−33]

Figure 2 characterizes the geometries of such more complex and, thus, computationally more demanding models. In fact,

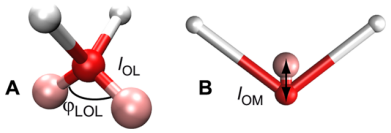


Figure 2. (A) The electrostatic geometry $\mathbf{G}_e$ of a five-point model is defined by the distance l_{OL} between the red oxygen atom O and each of the two pink massless points with charges $q_L = -q_H$, which are symmetrically located above and below the molecular plane, and by the angle φ_{LOL} between the connections of the pink points with O. Depicted is the tetrahedral TIP5P model[30] (l_{OL} = 0.7 Å, φ_{LOL} = 109.47°). For φ_{LOL} = 360° the two charges q_L become degenerate and merge into the single charge $q_M = 2q_L$, thus rendering the four-point model (B), whose $\mathbf{G}_e$ is defined by the massless M site point on the HOH bisectrix at the distance $l_{OM} \equiv l_{OL}$. For l_{OM} = 0 one recovers the three-point model of Figure 1, for which $\mathbf{G}_e = \mathbf{G}_m$.

such models were capable of reproducing the local solvation structures in water as measured by the various RDFs much better than their three-point predecessors.[15,19,34] While they were initially chosen as nonpolarizable, thus attempting to approximate the enhanced dipole moments of the water molecules in the bulk[21−26] by a mean field approach, subsequently also polarizable four-point,[35−58] five-point,[59,60] or even six-point potentials[61−64] were suggested with the aim of tackling also the drawback (ii) of the poor transferability. Furthermore, also polarizable three-point models such as AMOEBA[65] were suggested, which place, in addition to electrostatic monopoles, also dipole and quadrupole moments to the three atoms of H_2O, attempting in this way to properly model the higher static multipole moments of the molecule. Whether this alternative model class, which offers many adjustable parameters, can provide a better compromise between accuracy and computational efficiency than polarizable four-, five-, or six-point models is unclear.

1.2. Polarizable Models. As nicely reviewed in ref 29, which provides an almost complete set of references, four main routes were taken to include the polarizability into water models, i.e. the use of induced molecular point dipoles[35−42] (ID), which may be equivalently replaced by so-called Drude oscillators[43−50,60,66] (DO), of fluctuating charges[52−54,63] (FQ), and of induced atomic dipoles[55−58,65] (3-ID), where all the given references pertain to four-point models.

Quite generally one may state that the search for suitable parameters characterizing the electrostatic properties of molecules is in principle easier for polarizable molecular mechanics (PMM) force fields than for nonpolarizable ones,[29] because the latter effectively try to include the average dipole moment of a water molecule, which is induced in the liquid phase by the surrounding molecules, into the choice of the static partial charges. Therefore, all MM parameters are usually derived by comparing bulk properties obtained in MD simulations at certain thermodynamic conditions with corresponding experimental data.

In contrast, PMM models describe the electronic polarization explicitly during a simulation and therefore can use the dipole moment μ^g_{exp} = 1.855 D and isotropic polarizability α^g_{exp} = 1.470 Å^3 of an isolated water molecule, which are experimentally well-known,[20,67,68] as corner pillars of a parametrization. Here, the polarizability α can be safely assumed to be isotropic, because the deviations from isotropy are small.[68] Then solely the higher multipole moments of a water molecule in the liquid phase and the parameters entering a suitable van der Waals potential of the Lennard-Jones[16,17] or Buckingham[69] type remain to be specified.

The higher multipole moments of (P)MM water models are determined by their electrostatic geometries $\mathbf{G}_e$, that is by the numbers and locations of the partial charges generating the static electrostatic signatures of the various models (cf. Figures 1 and 2). In the case of the three-point models, the choice of a molecular geometry $\mathbf{G}_m$ automatically fixes also $\mathbf{G}_e$. Here, choosing μ^g_{exp} for the zero-field dipole moment determines the partial charges and, thus, all higher multipole moments.

Attempts of constructing such polarizable three-point PMM models[42,70−73] either yielded highly suboptimal liquid−vapor coexistence curves, when applied to the study of critical phenomena in a Monte Carlo simulation setting,[42,71] or showed strong underestimates of the dimer binding energies.[73−75] The latter underestimate led several authors[72−75] to employ a larger static dipole moment of 1.9−2.1 D, which in some cases[72−74] enforced a reduction of the polarizability from α^g_{exp} to values of 0.9−1.1 Å^3 to avoid an overpolarization of the liquid.

From the results of MD simulations of the liquid at standard conditions one can furthermore conclude that polarizable three-point models[72−76] generally yield RDFs $g_{OO}(r)$ values for the oxygen−oxygen distances r, which exhibit much less structure in the region beyond the first solvation shell than the corresponding experimental data[10−14] and, hence, resemble the rather structureless RDFs of the TIP3P[7] and SPC[8] models (for SPC see Figure 4 in ref 73). Such failures may be partially avoided, if one additionally includes empirically parametrized interaction potentials,[70] which however lack any physical motivation, render the model computationally more expensive, and most likely represent an overspecialization to the liquid at standard conditions.

Moreover, the careful analysis of Yu et al.[73] has demonstrated that polarizable three-point models grossly overestimate the dielectric constant at standard conditions by 55−117%, where the smaller overestimate could only be achieved by reducing the polarizability by 37% as compared to the gas phase value

dx.doi.org/10.1021/jp404548k | J. Phys. Chem. B 2013, 117, 9486−9500

α^g_{exp}. Hence it seems that there is no combination of μ^g_{exp} and α^g_{exp} with a reasonable three-point geometry $\mathbf{G}_m$, which leads to acceptable bulk phase properties. Consequently many authors tried to choose empirically adjusted values for μ, α, or both, for a better match of these properties.[72−75] However, the limited success of these attempts seemed to indicate that a larger model complexity, as represented, e.g., by an additional charge point, is necessary for substantially improved bulk properties. As a consequence, many authors suggested polarizable four-point models[35−50,52−58,65] as documented by the large list of corresponding references.

1.3. Choice of α. However, even with the added flexibility of shaping the higher multipole moments through the introduction of an additional charge point (cf. Figure 2), the parametrization of polarizable four-point ID and DO models featuring reasonable bulk phase properties turned out to be impossible as long as α^g_{exp} was chosen.[45,46] For instance, Lamoureux et al.[46] concluded from a large number of parametrization attempts "that models with the experimental gas-phase polarizability systematically yield an overestimated dielectric constant, typically in the range of 150 to 200. Furthermore, none of these models could get both the correct density and enthalpy: liquid densities close to the experimental value always resulted in vaporization enthalpies that were too favorable by about 1 to 2 kcal/mol. The average dipole of those models is around 2.9 D, which is consistent with the overestimated dielectric constant." Therefore the authors deduced "that the value of α must be around 1.0 Å^3 to yield reasonable liquid properties."

The physical reasons for the necessity of a reduced α were eventually revealed[29] by the work of Schropp and Tavan.[26] These authors carried out hybrid calculations, which combined the density functional theory (DFT) description of a solute water molecule with a MM modeling of its aqueous environment.[24] They concluded that the external field $\mathbf{E}(\mathbf{r})$, which is generated by the surrounding MM water molecules and polarizes the electron density of the DFT solute, exhibits a substantial inhomogeneity within the volume v occupied by that density. Due to this inhomogeneity, the spot check $\mathbf{E}(\mathbf{r}_O)$ at the position $\mathbf{r}_O$ of the oxygen atom overestimates the polarizing volume average $\langle\mathbf{E}\rangle_v$ of the field by about 40%. Thus, ID and DO models, which compute the polarizing field as spot checks at or near $\mathbf{r}_O$, are necessarily plagued[38,45,46] by much too large induced dipole moments $\mu^i = \alpha^g_{exp}\mathbf{E}(\mathbf{r}_o)$.

To avoid this artifact, Schropp and Tavan[26] suggested to apply the mean field approximation and, hence, to employ the reduced polarizability α_{eff} = 1.005 Å^3. This value is close to the values previously suggested for DO four point models upon purely empirical reasoning by Lamoureux et al.[46,47] (SWM4-DP: 1.043 Å^3, SWM4-NDP: 0.978 Å^3), Yu et al.[45] (COS/G2: 1.255 Å^3, COS/G3: 1.250 Å^3), and Yu et al.[64] (SWM6: 0.88 Å^3). Furthermore, the more recent empirical DO model COS/D by Kunz and van Gunsteren[49] applies a field dependent polarizability, which is α = 1.49 Å^3 for fields smaller than 1.2 V/Å and is reduced to $\alpha \approx$ 0.9 Å^3 for fields of about 2.4 V/Å (the range [1.2, 2.4] V/Å approximately covers the distribution of field strengths occurring in the liquid; cf. Figure 6 in ref 26). Inspired by these results[26,49] also Baranyai and Kiss[51] employed a field-dependent polarizability, which converges from above to 1.0 Å^3 at large fields.

Another option, which does not resort to the mean field approximation, is to use α^g_{exp} and to combine Gaussian induced dipoles with static point charges.[77] Then the volume average $\langle\mathbf{E}\rangle_v$ required for the computation of the polarizing field is executed explicitly with the volume $v = (2\pi)^{3/2}\sigma^3$, where σ is the standard deviation of the employed Gaussian. Reference 26 predicts σ = 0.9 Å as the optimal value.[78] The use of Gaussian dipoles should entail a better transferability, because it makes no reference to the specific field inhomogeneity present in the bulk liquid at standard conditions. It is the method of choice, if one calculates properties of PMM models from DFT/MM simulations,[26,79] because in this setting a spacious electron density ρ_e is polarized by an environment of static MM partial charges and because the polarization of ρ_e can be easily emulated by assigning Gaussian inducible dipoles to the non-hydrogen DFT atoms.

Instead of assigning the Gaussian character to the induced dipoles, one can also employ α^g_{exp}, if one combines Gaussian partial charges with induced point dipoles[40,48] or Drude oscillators,[50] because this approach is equivalent to the combination of Gaussian dipoles with point charges as far as the dipole−charge interactions are concerned. It is however much more expensive, because for every pair of nearby water molecules nine (instead of one) Gaussian interactions have to be evaluated. Furthermore it is algorithmically less stable, because at short distances the mutual interaction of inducible point dipoles can lead to a diverging polarization,[77,80] which is avoided in our approach by the use of inducible Gaussian dipoles.

1.4. Choice of the Molecular Geometry. As is documented in Table 1, the molecular geometry $\mathbf{G}_m$ of a water molecule changes upon transfer into the liquid phase with a slight increase Δl_{OH} = 0.011 Å of l_{OH} and a slight widening $\Delta\varphi_{HOH}$ = 0.78° of φ_{HOH}. Because $\mathbf{G}^l_m$ is known, we will use it for the design of our liquid phase water models.

Table 1. Geometry $\mathbf{G}_m$ of a Water Molecule in the Gas and Liquid Phases, Respectively[20,81−83]

deg of freedom	$\mathbf{G}^g_m$	$\mathbf{G}^l_m$
l_{OH} (Å)	0.9572	0.9680
φ_{HOH} (deg)	104.52	105.30

Recently, the dependence of μ and α on $\mathbf{G}_m$ was thoroughly characterized by DFT calculations and by DFT/MM dynamics simulations of a flexible DFT water molecule embedded in a TIP4P solvent.[79] Interestingly, the effects of Δl_{OH} and $\Delta\varphi_{HOH}$ (i.e., of replacing $\mathbf{G}^g_m$ by $\mathbf{G}^l_m$) on the polarizability α turned out to compensate each other and the changes of μ were found to be small. Therefore, the authors suggested[79] to consider the experimental gas phase values μ^g_{exp} and α^g_{exp} as best estimates for their unknown liquid phase counterparts $\boldsymbol{\mu}^l$ and $\boldsymbol{\alpha}^l$. We will adopt this suggestion for our construction of polarizable four- and five-point models.

The presentation starts with an outline of the concepts and procedures guiding our parametrization effort. Subsequently we will describe the simulation systems and computational methods employed for the DFT/PMM derivation of model parameters, for the PMM-MD simulation of water dimers and bulk water systems, and for the calculation of observables characterizing bulk properties. The results will be presented and discussed.

dx.doi.org/10.1021/jp404548k | J. Phys. Chem. B 2013, 117, 9486−9500

2. GUIDELINES OF PMM WATER MODEL CONSTRUCTION

The degree of complexity of a PMM water model is given by the number ν of points at which forces have to be evaluated, by the shape (Gaussian vs point-like) assumed for the charges and induced dipoles, and by the number and structure of the terms employed for the description of the van der Waals interaction. Here, each increase of complexity necessarily entails an enhanced computational effort but does not lead with certainty to more accurate and better transferable descriptions. On the contrary, too many and physically poorly justified parameters may even lead to models that are overspecialized to some observables but fail for others (see refs 50, 51, 63, and 70 for examples). Therefore, models should be as simple as possible while rendering optimally accurate descriptions within the given class of complexity. The likelihood to achieve these conflicting aims can be heightened, if all model ingredients and parametrization procedures have clear physical motivations.

2.1. Physical Cornerstones. Physically well established are the geometry $\mathbf{G}^{\mathrm{g}}_{\mathrm{m}}$, the polarizability $\alpha^{\mathrm{g}}_{\mathrm{exp}}$, and the dipole moment $\mu^{\mathrm{g}}_{\mathrm{exp}}$ of the water molecule in the gas phase[20,67,68,81] as well as the geometry $\mathbf{G}^{l}_{\mathrm{m}}$ in the liquid phase[82,83] (cf. Table 1 and Figure 1). Furthermore, as explained in section 1.3, the use of a Gaussian inducible dipole at $\mathbf{r}_{\mathrm{O}}$ can (i) guarantee a correct volume average over the strongly inhomogeneous polarizing field[26] and can (ii) nicely model a DFT/MM setting. Moreover, DFT/MM calculations on flexible DFT water molecules have clearly shown that the deformation contributions to α and μ can be safely neglected.[79] Therefore, four- and five-point PMM models of liquid water should be constructed choosing the values

$$\mu^{l} \equiv \mu^{\mathrm{g}}_{\mathrm{exp}} = 1.855\ \mathrm{D} \quad \text{and} \quad \alpha^{l} \equiv \alpha^{\mathrm{g}}_{\mathrm{exp}} = 1.470\ \text{Å}^3 \tag{1}$$

for the static dipole moment and for the polarizability. While this choice of α^{l} should also apply to three-point PMM models, μ^{l} must be chosen differently in this case (cf. section 1.2).

Finally, a key motivation for our current effort is the fact that the DFT/MM method[24] used in refs 26 and 79 most recently has been extended toward the use of PMM force fields[84] and, therefore, now enables a new and self-consistent DFT/PMM strategy toward the parametrization of PMM models. In the new DFT/PMM method, the polarizable degrees of freedom within the PMM fragment are described by Gaussian induced dipoles located at the positions of the non-hydrogen atoms. In a joint iteration, the Kohn−Sham orbitals and the PMM dipoles are brought to self-consistent field (SCF) solutions. Using the linearly scaling "structure adapted, fourth order fast multipole method" called SAMM$_4$,[85] the electrostatic interactions between the (P)MM environment and the DFT fragment are accurately and efficiently calculated in a Hamiltonian DFT/(P)MM setting.

Like its predecessor, also the new approach models the partial charges of those PMM atoms, which occupy the immediate environment of a DFT atom, as Gaussian distributions of widths $\tilde{\sigma}_i$, because this choice can avoid artificial distortions of the DFT electron density. The $\tilde{\sigma}_i$ are important parameters of the DFT/PMM method,[84,86] which have to be optimized for the employed PMM atom types. Therefore, such an optimization of the $\tilde{\sigma}_i$ must be included into any attempt of constructing PMM models from DFT/PMM calculations. As described in section 3.1 we chose the isolated DFT-DFT, DFT-PMM, and PMM-DFT water dimers as our reference for iteratively fixing the $\tilde{\sigma}_i$ at all stages of the iterative parametrization procedure described below.

2.2. Optimal Widths σ of the Gaussian Dipoles. The new DFT/(P)MM method will be used by us to check for snapshots s from DFT/(P)MM structural ensembles $\mathcal{S}_\nu$ of ν-point models for bulk water, which are generated by *NVT* (P)MM-MD simulations as described in section 3.1 further below, how well the induced dipole moments

$$\mu^{\mathrm{i}}_{\mathrm{DFT/(P)MM}}(s) \equiv \mu^{l}_{\mathrm{DFT/(P)MM}}(s) - \mu^{l}_{\mathrm{DFT}} \tag{2}$$

which are calculated by DFT/(P)MM and DFT for water models rigidly fixed at the liquid phase geometry $\mathbf{G}^{l}_{\mathrm{m}}$, show the linear response

$$\mu^{\mathrm{i}}_{\sigma,\nu}(s) \equiv \alpha^{l}_{\mathrm{DFT/(P)MM}} \langle \mathbf{E}(s) \rangle_{\sigma,\nu} \tag{3}$$

to the polarizing electric field $\langle \mathbf{E}(s) \rangle_{\sigma,\nu}$ averaged over the Gaussian volume of a corresponding PMM dipole. Instead of $\alpha^{l}_{\mathrm{DFT/(P)MM}}$, which is the polarizability of a DFT water molecule at $\mathbf{G}^{l}_{\mathrm{m}}$ embedded in a (P)MM environment, one may equally well use[79] in eq 3 the polarizability $\alpha^{\mathrm{g}}_{\mathrm{DFT}}$ calculated[26] by DFT for an isolated water molecule exposed to homogeneous external fields at the experimental gas phase geometry $\mathbf{G}^{\mathrm{g}}_{\mathrm{m}}$; that is, we define

$$\alpha^{l}_{\mathrm{DFT/(P)MM}} \equiv \alpha^{\mathrm{g}}_{\mathrm{DFT}} \tag{4}$$

The width σ of the Gaussian PMM dipole $\mu^{\mathrm{i}}_{\sigma,\nu}$ will be varied until the correlations between $\mu^{\mathrm{i}}_{\mathrm{DFT/(P)MM}}(s)$ and $\mu^{\mathrm{i}}_{\sigma,\nu}(s)$ show minimal root-mean-square deviations $\chi(\sigma,\nu)$. It will be of interest to see whether the optimal value $\sigma = 0.9$ Å determined earlier[26,78] for TIP3P, TIP4P,[7] and SPC/E[9] environments also holds for PMM surroundings.

2.3. Electrostatic Geometries $\mathbf{G}^{l}_{\mathrm{e}}$ from DFT/(P)MM. The DFT/(P)MM method enables us to compute the electrostatic geometries $\mathbf{G}^{l}_{\mathrm{e}}$ of polarizable four- and five-point models by optimizing the match between the surface potentials of DFT water molecules, which are rigidly fixed at $\mathbf{G}^{l}_{\mathrm{m}}$ and are embedded in bulk (P)MM liquid structures $s \in \mathcal{S}_\nu$, with that of rigid PMM test molecules surrounded by the same structures. For the intended comparison of surface potentials, we choose μ^{l}_{DFT} and $\alpha^{\mathrm{g}}_{\mathrm{DFT}}$ for the dipole moment and polarizability of the PMM test molecules (as in eqs 2 and 3).

In each snapshot $s \in \mathcal{S}_\nu$ of a DFT/(P)MM hybrid system one can replace the DFT fragment by a ν-point PMM test molecule. Next one can compute the external field $\langle \mathbf{E}_s \rangle_{\sigma,\nu}$ polarizing this PMM model as an average over the volume $v(\sigma)$ of its Gaussian induced dipole, which according to eqs 3 and 4 has the value $\mu^{\mathrm{i}}_{\sigma,\nu}(s) = \alpha^{\mathrm{g}}_{\mathrm{DFT}} \langle \mathbf{E}_s \rangle_{\sigma,\nu}$. Subtracting the potential $\Phi[\mathbf{r} \mid \mu^{\mathrm{i}}_{\sigma,\nu}(s)]$, which is created by this dipole at points $\mathbf{r} \in \mathcal{P}$ randomly selected from a spherical surface $\mathcal{P}$ surrounding the PMM test molecule, from the potential $\Phi(\mathbf{r} \mid \rho_s)$, which stems from the distributed charge density ρ_s of the associated DFT fragment, yields a set of spot checks

$$\Phi(\mathbf{r} \mid \rho^{\mathrm{stat}}_{s}) = \Phi(\mathbf{r} \mid \rho_s) - \Phi[\mathbf{r} \mid \mu^{i}_{\sigma,\nu}(s)], \quad \forall\ \mathbf{r} \in \mathcal{P} \tag{5}$$

for the potential generated by the supposedly static part ρ^{stat}_{s} of ρ_s. The data set of surface potential spot checks $\Phi(\mathbf{r} \mid \rho^{\mathrm{stat}}_{s})$ belonging to all surface points $\mathbf{r} \in \mathcal{P}$ and all snapshots $s \in \mathcal{S}_\nu$ can now be used to optimize the parameters l_{OM} of four- and (l_{OL}, φ_{LOL}) of five-point models (cf. Figure 2) endowed with

dx.doi.org/10.1021/jp404548k | J. Phys. Chem. B 2013, 117, 9486−9500

the static dipole moment μ^l_{DFT}. This optimization is easily achieved by minimizing the root-mean-square deviation between $\Phi(\mathbf{r}\,|\,\rho_s^{stat})$ and the potential $\Phi(\mathbf{r}\,|\,q)$ generated by the static partial charges $q(l_{OM})$ at the three charge points or by the static partial charges $q(l_{OL}, \varphi_{LOL})$ at the four charge points, which are provided by these models.

2.4. van der Waals Parameters. Following the arguments in refs 10, 40, and 48 we decided to employ the Buckingham[69] instead of the usual Lennard-Jones[16,17] potential, because it leads to a softer onset of the RDF $g_{OO}(r)$. This advantage is bought at the expense of a somewhat larger computational effort that has to be spent on the computation of the short-range repulsion, of the need to determine three instead of only two parameters, and of a reduced compatibility with standard biomolecular force fields.

However, with a computational strategy such as SAMM$_4$, which employs distance classes for the evaluation of long-range forces,[85] the computation of the short-range repulsion can be confined to the innermost distance class containing at most a few hundred interaction partners for a given water molecule. The compatibility with biomolecular fields can be arranged and the interactions of the water with atoms of solute molecules can be covered by mapping the repulsive parts of the Buckingham potential

$$E_B(r) = A_1 \exp(-rA_2) - B/r^6 \qquad (6)$$

where r is the oxygen−oxygen distance, to an almost equivalent Lennard-Jones A/r^{12} repulsion. A corresponding fit determining A from (A_1, A_2) can be confined to $r \in [2.5, 6]$ Å for typical potentials.

In contrast to the electrostatics, the Buckingham parameters (A_1, A_2, B) must be determined by comparing results of PMM-MD simulations, which employ μ^l and α^l as given by eq 1, with experimental data on the bulk liquid. Reasonable choices for these reference data are the potential energy E_{pot} per molecule and the pressure p at the standard temperature T = 300 K and at the associated experimental density n.[87,88] The weak-coupling approach of Berweger et al.[89] provides suitable procedures for an automated parameter optimization. Because the Buckingham potential has three adjustable parameters, one can take the position of the first peak[11] in $g_{OO}(r)$ as an additional criterion. Here the size of the simulation system can influence the results of such a simulation-based parametrization. Therefore the stability toward finite size effects must be checked.

2.5. Three-Point PMM Model: Optimization of μ^l, A_1, A_2, B. As reviewed in section 1.2, the limited complexity prevents the use of μ^g in three-point PMM models, if one wants to obtain reasonable potential energies and RDFs.[73] Therefore, we choose μ^l as a fourth adjustable parameter to be optimized by PMM-MD simulations, which solely adopt α^l from eq 1. As a fourth optimization target we take the height of the first maximum of $g_{OO}(r)$.

2.6. Self-consistency Iteration. Sections 2.2−2.5 have outlined a series of steps for the parametrization of polarizable three-, four-, and five-point models, which utilize a new DFT/(P)MM technology[84] and weak-coupling[89] PMM-MD simulations. These steps combine into an iterative and self-consistent parameter optimization.

Figure 3 schematically characterizes the self-consistent parametrization for PMM four- and five-point water models. It employs the PMM models found at the end (top right corner of Figure 3) of the cycle to generate updated structural ensembles $\mathcal{S}$ representing the liquid phase from snapshots s of PMM-MD simulations (T = 300 K, n = 0.9965 g/cm^3). The iteration is initialized through MM-MD with standard non-polarizable models.[7,33,30] The snapshots $s \in \mathcal{S}$ are taken as representative (P)MM environments for the DFT fragments of DFT/(P)MM calculations, which yield dipole moments $\mu^l_{DFT/(P)MM}(s)$ and surface potentials $\Phi(\mathbf{r}\,|\,\rho_s)$ characterizing the polarized DFT fragments. According to eq 2 the associated induced dipoles $\mu^i_{DFT/(P)MM}(s)$ are defined relative to the dipole moment μ^g_{DFT} calculated for an isolated water molecule at the experimental liquid phase geometry $\mathbf{G}^l_m$ that is used in the DFT/(P)MM calculations. Minimizing the deviations between eqs 2 and 3 yields the width σ of the Gaussian PMM dipoles (lower left corner in Figure 3).

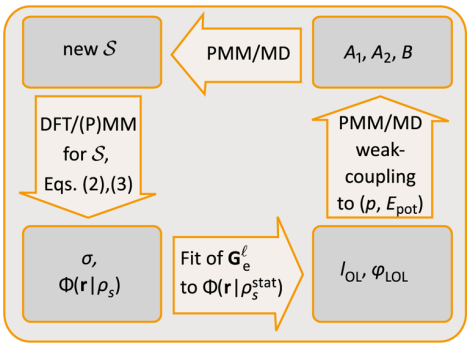


Figure 3. Scheme of the self-consistency iteration for parameter computation of PMM four- and five-point water models. Starting at top left: DFT/(P)MM for ensemble $\mathcal{S}$ and minimizing the deviation $\chi(\sigma)$ between eqs 2 and 3 ⇒ [Gaussian width σ of PMM dipole and surface potentials $\Phi(\mathbf{r}\,|\,\rho_s)$]. Bottom: Fit of $\mathbf{G}^l_e = \{l_{OL}, \varphi_{LOL}\}$ as given in Figure 2 to $\Phi(\mathbf{r}\,|\,\rho_s^{stat})$ as defined by eqs 3−5 ⇒ $\mathbf{G}^l_{e,opt}$. Right: Using (μ^l, α^l) from eq 1 in PMM-MD weakly coupled to (p, E_{pot}) ⇒ Buckingham parameters (A_1, A_2, B) (cf. eq 6). Top: PMM/MD ⇒ new ensemble $\mathcal{S}$.

Subtracting next the surface potentials $\Phi[\mathbf{r}\,|\,\mu^i_\sigma(s)]$ of the induced PMM dipoles $\mu^i_\sigma(s) = \alpha^g_{DFT}\langle \mathbf{E}_s\rangle_\sigma$ from the DFT/(P)MM potentials $\Phi(\mathbf{r}\,|\,\rho_s)$ yields according to eq 5 the static contributions $\Phi(\mathbf{r}\,|\,\rho_s^{stat})$, to which electrostatic geometries $\mathbf{G}^l_e$ of the four- and five-point PMM models are fitted under the constraint that the dipole moment is μ^g_{DFT} for $\langle \mathbf{E}_s\rangle_\sigma = 0$.

Up to this step (top left to lower right corner of Figure 3), the parametrization aimed at electrostatics properties $(\sigma, \mathbf{G}^l_e)$ that are accessible by DFT and DFT/(P)MM calculations. The remainder of the cycle serves to compute the parameters (A_1, A_2, B) of the Buckingham potential eq 6 applying the procedures described in section 2.4. The iteration is stopped as soon as the change of σ, and therefore the change of all further parameters, falls below a predefined threshold.

3. METHODS

Below we will first describe the methods employed in the parametrization cycle, by which our new water models are self-consistently derived (cf. Figure 3). Subsequently, we sketch the simulations and observables applied to their evaluation.

For all MD simulations executed during parametrization and evaluation we applied the parallelized program package IPHIGENIE,[85] which is a thorough revision of an earlier code called EGO.[90,91] IPHIGENIE implements the linearly scaling fast multipole method SAMM$_4$/RF,[85] which strongly

dx.doi.org/10.1021/jp404548k | J. Phys. Chem. B 2013, 117, 9486−9500

extends the preceding SAMM approach,[90,92,93] for the more accurate and efficient computation of the electrostatic interactions (cf. section 2.1). In periodic systems, which just can enclose a sphere of radius $L/2$, these interactions are, following the "minimum image convention (MIC)",[94] explicitly evaluated up to the distance $d_{MIC} = L/2$. At the associated moving boundary the description smoothly switches[91] to a Kirkwood continuum[95] with the dielectric constant $\varepsilon = 78$ of water at standard conditions.[96] For integration we applied the Verlet algorithm[17] with a time step of 1 fs. Molecular geometries were kept fixed using the M-SHAKE algorithm[97] with a relative tolerance of 10^{-10}. The PMM-SCF iterations were assumed to be converged as soon as the relative changes of the induced dipoles were all smaller than 5×10^{-5} D.

3.1. Parametrization. *Simulation Systems.* For parametrization three periodic cubic boxes of volume $V = L^3$ with the edge length $L = 35.6$ Å were filled with $N_m = 1500$ rigid and nonpolarizable TIP3P,[7] TIP4P/2005,[33] and TIP5P[30] water models, respectively, which all feature the experimental gas phase geometry $\mathbf{G}_m^g$. V was chosen such that the experimental density[87] $n_{exp} = 0.9965$ g/cm^3 of water at $T_0 = 300$ K and $p_0 = 1$ atm were reproduced. The boxes were equilibrated by 300 ps MM-MD simulations in the *NVT* ensemble. Here, the systems were steered toward T_0 by a Bussi thermostat[98] (coupling time: 0.1 ps). After the first parametrization cycle (cf. Figure 3) the MM models in the three boxes were replaced by the thus determined ν-point PMM models ($\nu = 3, 4, 5$), which all have the experimental liquid phase geometry $\mathbf{G}_m^l$. Also these boxes were then equilibrated for 300 ps by *NVT* PMM-MD.

Structural Ensembles $\mathcal{S}_\nu$. The equilibrated boxes were simulated for another 50 ps. From each of these *NVT* PMM-MD trajectories five snapshots were taken at temporal distances of 10 ps. From each snapshot 200 water molecules were randomly selected as DFT fragments for subsequent DFT/(P)MM calculations, within which the respective structures were kept fixed. Thus we obtained for each simulation box a structural ensemble $\mathcal{S}_\nu$ covering 1000 statistically independent solvation structures s.

DFT/(P)MM Calculations. Using the gradient corrected exchange functional of Becke,[99] the correlation functional of Perdew,[100] and the norm conserving pseudopotentials of Troullier and Martins,[101] as implemented in the grid based DFT program CPMD,[102] in combination with a plane wave basis set characterized by a 80 Ry cutoff, we calculated for each solvation structure s the charge distribution ρ_s and dipole moment $\mu^l_{DFT/(P)MM}(s)$. In each of these DFT/(P)MM calculations the respective DFT fragment was centered into a cubic box of edge length 9 Å, keeping the nuclei at least 3 Å away from its faces. This DFT approach was chosen identical to that in ref 79, to ensure comparability, and is called MT/BP. Our DFT/(P)MM approach, however, markedly differs, because ref 79 used the DFT/MM method of Eichinger et al.[24] whereas we took advantage of its recent extension[84] toward PMM force fields and toward a strongly enhanced accuracy and efficiency (cf. section 2.1).

We mentioned at the bottom of section 2.1 that the widths $\tilde{\sigma}_i$ of the Gaussian partial charge distributions assigned to atoms close to the DFT fragment have to be optimized for the employed PMM atom types and that we chose DFT/(P)MM hybrid descriptions of the water dimer to optimize the $\tilde{\sigma}_i$ at each stage of the iterative optimization. The dimer properties associated with the final values $\tilde{\sigma}_i$, which are required for the PMM models studied by us, are documented in section S1 of the Supporting Information (SI). We found the widths $\tilde{\sigma}_H = 0.24$ Å and $\tilde{\sigma}_{O/L/M} = 0.46$ Å applicable to the static partial charges at the hydrogens and oxygens (or lone-pair sites L/M), respectively. Hybrid DFT/MM water dimers, in contrast, were optimally described with the larger values $\tilde{\sigma}_H = 0.37$ Å and $\tilde{\sigma}_{O/L/M} = 0.62$ Å. The latter widths are close to the standard width of 0.57 Å employed in earlier studies.[24,26,79]

Width σ of the Gaussian PMM Dipoles. In the cycle $c = 1$ of the iteration sketched in Figure 3, the deviations $\chi(\sigma)$ introduced in section 2.2 were minimized for the structural ensembles $\mathcal{S}$ of the three MM water models by varying the widths σ of the Gaussian PMM dipoles in the range [0.7, 1.2] Å with steps of 0.01 Å. As a best guess for σ we took the arithmetic mean over the three ensembles. In the subsequent iterations $c = 2, 3, ...$, we scanned, for the ensembles $\mathcal{S}_\nu$, the deviations $\chi(\sigma_\nu)$ in the range [0.7, 0.9] Å using finer steps of 0.001 Å and determined optimal values $\sigma_{\nu,c}$. The iterations were stopped as soon as $|\sigma_{\nu,c+1} - \sigma_{\nu,c}| < 0.001$ Å.

Determination of G_e. The surface potential $\Phi(\mathbf{r}\,|\,\rho_s^{stat})$ defined by eq 5 was calculated on a spherical surface $\mathcal{P}$ of radius $R_S = 2.75$ Å around the center of mass of the DFT water molecule associated with snapshot $s \in \mathcal{S}_\nu$ at 500 equally distributed and randomly selected points. For $\nu \in \{4, 5\}$ the electrostatic geometries $\mathbf{G}_e$ were determined by minimizing the root-mean-square deviations

$$\zeta(l_{OL}, \varphi_{LOL}) \equiv \left[\langle\{\Phi(\mathbf{r}|\rho_s^{stat}) - \Phi[\mathbf{r}|Q(l_{OL}, \varphi_{LOL})]\}^2\rangle_{\mathcal{P},\mathcal{S}}\right]^{1/2} \quad (7)$$

where $\langle ... \rangle_{\mathcal{P},\mathcal{S}}$ denotes the arithmetic mean over the ensembles $\mathcal{P}$ and $\mathcal{S}_\nu$, between $\Phi(\mathbf{r}\,|\,\rho_s^{stat})$ and the surface potential $\Phi(\mathbf{r}\,|\,Q)$ of the static partial charges

$$Q \equiv \left\{q_i\left(l_{OL}, \varphi_{LOL}\right)\middle| i \in \{H1/2, L1/2\}\right\}$$

For $\nu = 4$ the parameter $l_{OL} = l_{OM}$ (cf. Figure 2) was determined by a line search over the range [0,0.5] Å with the accuracy $\Delta l_{OL} = 0.001$ Å. For $\nu = 5$ we scanned $l_{OL} \in [0,0.7]$ Å and $\varphi_{LOL} \in [0, 360]°$ and determined the optimum with the accuracies $\Delta l_{OL} = 0.001$ Å and $\Delta\varphi_{LOL} = 0.1°$. Variations of R_S left the results invariant within the indicated accuracy limits.

PMM/MD Weak-Coupling Simulations. For $\nu \in \{4, 5\}$ weak-coupling *NVT* simulations[89] were applied (at the experimental density[87] $n_{exp}(p_0,T_0) = 0.9965$ g/cm^3) to determine the Buckingham parameters A_2 and B. Here, A_1 was chosen from the range $[5,40] \times 10^4$ kcal/mol in steps of 10^4 kcal/mol. In these 20 ps MD simulations, A_2 was coupled to the pressure p with the target $p_0 = 1$ atm and a coupling time of 0.1 ps, and B was coupled to the potential energy E_{pot} per molecule with the experimental[88] target $E^0_{pot} = -9.92$ kcal/mol and a coupling constant of 1 ps. From the last 2 ps RDFs $g_{OO}(r)$ were derived. The value A_1 rendering the best match with the experimental position of the first peak at 2.76 Å as determined by Soper et al.[10] was chosen as the initial guess. Two further refinement steps scanning the vicinity of this guess and employing 10 times extended simulation spans and fifty times extended coupling times yielded the final values of A_2 and B as averages over the last 100 ps of MD simulation.

For $\nu = 3$ a similar procedure was applied coupling, however, the static dipole moment μ^l instead of the dispersion parameter B to the potential energy E_{pot} while scanning $B \in [200, 1200]\cdot$(kcal/mol)Å^6 and A_1 over the range given above. As

dx.doi.org/10.1021/jp404548k | J. Phys. Chem. B 2013, 117, 9486−9500

selection criteria for B and A_1, the experimental location (2.76 Å) and height[10,11] (2.50) of the first peak in $g_{OO}(r)$ were used. Note here that, after the parametrization of our TLνP models was complete, Skinner et al.[14] most recently derived an RDF from X-ray data, which locates the first peak at 2.80 ± 0.01 Å and assigns to it a height of 2.57 ± 0.05. Section S8 of the SI compares the quoted experimental functions $g_{OO}(r)$.

3.2. Evaluation. The resulting polarizable three-, four-, and five-point water models are denoted as TLνP (ν = 3, 4, 5) from now on. For each model the quadrupole moment of the monomer and the geometry, potential energy, and dipole moment of the dimer were calculated. Furthermore, we also calculated the cyclic TLνP trimer, because here oxygen−oxygen distances have been measured.[103] Bulk properties were calculated for three different periodic boxes containing N_s = 728, N_m = 1500, and N_l = 3374 molecular models, respectively, in the NVT and NpT ensembles, usually controlling T by a Bussi thermostat[98] (BU, coupling time: 1 ps) and p by a Berendsen barostat[104] (coupling time: 10 ps, compressibility: 0.46 GPa^{-1}). For the calculation of kinetic properties, i.e. the diffusion constant (eq 9) and the viscosity (eq 10), we replaced the BU thermostat by a minimally invasive (MI) Berendsen thermostat[105] with the target temperature T_t = 20 K. Every simulation system was pre-equilibrated by MD for 300 ps.

As documented by the top part of Table 2, NVT simulations were carried out for each of the three system sizes N_i at the experimental density $n = n_{exp}(p_0, T_0)$. Thus, the volumes were $V_i = N_i m/n_{exp}$, where m is the mass of a water molecule. Data for the computation of statistical observables were recorded every ps. The coupling times $\tau_{s,m,l}$ = [171, 426, 683] ps for the simulations $N_iV_iT_0$[MI] with the MI thermostat were determined from the relation[105] $\tau_i = k_B(T_0 - T_t)/(2\beta_i)$, where the heating rates $\beta_{s,m,l}$ = [3.254, 1.306, 0.815] kcal/(mol ns) of the three systems had been measured from a series of short 10 ps NVE simulations. In all cases the system temperature was reliably controlled at T_0 with deviations below 0.1 K/ns.

Table 2. NVT Simulations of Durations d at the Densities n and the Temperature $T = T_0$ Carried out with the Thermostats BU[98] or MI[105] for the Polarizable TLνP Water Models, ν = 3, 4, 5

name	thermostat	d (ns)	n (g/cm^3)
$N_sV_sT_0$	BU/MI	4.5	0.997
$N_mV_mT_0$	BU/MI	2.0	0.997
$N_lV_lT_0$	BU/MI	1.5	0.997
$N_sV_+T_0$	BU	1.0	1.047
$N_sV_-T_0$	BU	1.0	0.947

From the trajectories $N_iV_iT_0$[BU], the average pressure $\langle p\rangle$ and potential energy $\langle E_{pot}\rangle$ per molecule as well as the distribution $p(\mu^i)$ of the induced dipole moment μ^i were calculated to check the size dependence of these observables. For this purpose also the size dependence of the static dielectric constant was self-consistently determined through[106]

$$\varepsilon(N_i, V_i, T_0) = \frac{1 + \frac{\langle \mathbf{M}_i^2\rangle - \langle \mathbf{M}_i\rangle^2}{\langle \mathbf{M}_i^2\rangle - \langle \mathbf{M}_i\rangle^2}\frac{2\varepsilon_{RF}}{2\varepsilon_{RF}+1}}{1 - \frac{\langle \mathbf{M}_i^2\rangle - \langle \mathbf{M}_i\rangle^2}{\langle \mathbf{M}_i^2\rangle - \langle \mathbf{M}_i\rangle^2}\frac{1}{2\varepsilon_{RF}+1}} \quad (8)$$

from the fluctuations of the total dipole moment $\mathbf{M}_i$ of system i, where ε_{RF} = 78 characterizes the distant continuum[91] of SAMM$_4$/RF, and from the trajectories $E_{RF,i}(t)$ of the reaction field energies through an iterative correction,[107] which arises if $\varepsilon(N_i, V_i, T_0) \neq \varepsilon_{RF}$.

Using the $N_iV_iT_0$[MI] data, the self-diffusion coefficient

$$D(N_i, V_i, T_0) = \lim_{t\to\infty}\langle[\mathbf{r}(t) - \mathbf{r}(0)]^2\rangle/6t \quad (9)$$

of a water molecule in bulk water follows[94] from the long time limit of its ensemble average mean square displacement, where $\mathbf{r}(t)$ marks the position of a water molecule at time t. Adopting the arguments of Dünweg and Kremer,[108,109] we will analyze the size dependence of $D(N_i, V_i, T_0)$ and use the slope of the resulting data to calculate the diffusion constant D_0 of the infinite system and the viscosity η via the formula

$$D(N_i, V_i, T_0) = D_0 - \frac{k_BT\xi}{6\pi\eta V_i^{1/3}} \quad (10)$$

where ξ has the value 2.837297 for cubic periodic boxes.

The RDFs $g_{OO}(r \mid N_i, V_i, T_0)$ values should be independent of system size. They were calculated from all snapshots, which were taken every ps from the $N_iV_iT_0$[BU] trajectories. RDFs were binned with a 0.01 Å resolution in the range $r \in [2, 8]$ Å and were smoothed with a symmetrical three-point kernel.

The bottom of Table 2 lists two additional NVT simulations of the small system, which were executed at the densities $n_\pm = n_{exp} \pm \Delta n$ with Δn = 0.05 g/cm^3 with the aim to determine the isothermal compressibility

$$\kappa_T(n_{exp}, T) = \left(\frac{\partial \ln n}{\partial p}\right)_T \approx \left(\frac{\ln(n_+/n_-)}{\langle p_+\rangle - \langle p_-\rangle}\right)_T \quad (11)$$

at $T = T_0$ by numerical differentiation.[110]

The three $N_ip_0T_0$ simulations characterized in Table 3 served to check the size dependence of the average density $\langle n\rangle$. The two additional simulations $N_sp_0T_+$ and $N_sp_0T_-$ of the small system are used to determine the heat capacity at $T = T_0$ according to[73]

Table 3. NpT Simulations of Durations d at the Temperatures T and the Pressure $p = p_0$ Carried out for the Polarizable TLνP Water Models, ν = 3, 4, 5

name	d (ns)	T (K)
$N_sp_0T_0$	2.0	300
$N_mp_0T_0$	1.0	300
$N_lp_0T_0$	0.5	300
$N_sp_0T_+$	1.0	310
$N_sp_0T_-$	1.0	290

$$C_p \approx \frac{E_+ - E_-}{T_+ - T_-} + \Delta C_{QM} \quad (12)$$

where E is the total energy per molecule and $\Delta C_{QM} \approx -2.22$ cal/(mol K) is a quantum correction roughly accounting for the quantized character of the neglected intramolecular vibrations and for the classical description of the intermolecular librational modes in liquid water at standard conditions. Similarly, the thermal expansion coefficient is approximately given by[111]

$$\alpha_p \approx -\left(\frac{\ln(n_+/n_-)}{T_+ - T_-}\right)_p \quad (13)$$

dx.doi.org/10.1021/jp404548k | J. Phys. Chem. B 2013, 117, 9486−9500

4. RESULTS

We start with the results of the self-consistent parametrization sketched in Figure 3, which define the ν-point models TLνP (ν = 3, 4, 5) optimized by us.

4.1. The TLνP Models. For all models the parametrization cycle was converged after c = 3 iterations (cf. Sec. 3.1).

Gaussian Widths σ of the PMM Dipoles. Figure 4 shows the deviations $\chi(\sigma, \nu)$ between the induced DFT/PMM dipoles $\mu^{i}_{DFT/PMM}(s)$ and the induced Gaussian PMM dipoles $\mu^{i}_{\sigma,\nu}(s)$, which were obtained as functions of σ from the snapshot ensembles $\mathcal{S}_\nu$ of the resulting TLνP models. According to eqs 2−4 this comparison requires values for the DFT predictions μ^{l}_{DFT} and α^{g}_{DFT} on the dipole moment and polarizability of an isolated water molecule at the experimental liquid and gas phase geometries, respectively. In our MT/BP setting these values are[26,79] μ^{l}_{DFT} = 1.79 D and α^{g}_{DFT} = 1.58 Å^3.

All deviations $\chi(\sigma, \nu)$ displayed in Figure 4 are convex functions of σ, whose minima are found at σ^{opt}_{ν} = 0.770 Å, 0.842

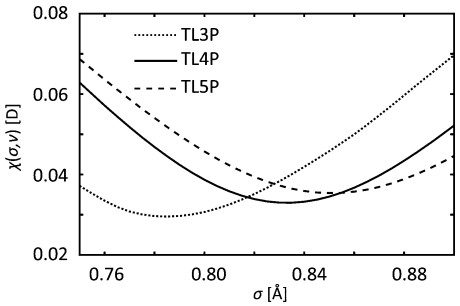


Figure 4. Root-mean-square deviations $\chi(\sigma, \nu)$ between $\mu^{i}_{DFT/PMM}(s)$ and $\mu^{i}_{\sigma,\nu}(s)$ obtained for the converged TLνP models: dotted, ν = 3; solid, ν = 4; dashed, ν = 5 (see sections 2.2 and 3.1 for explanations).

Å, and 0.853 Å for ν = 3, 4, and 5, respectively. These values are smaller than the width σ^{opt} = 0.9 Å expected[78] from the DFT/MM study in ref 26. However, this decrease compares well with the likewise smaller Gaussian width $\tilde{\sigma}_i$ of the static PMM partial charges, because the expected width belongs to a convolution of the Gaussian model for the DFT charge distribution (width 0.7 Å)[26] with the Gaussian attached to the partial charges. With $\tilde{\sigma}_O$ = 0.46 Å we obtain for the convolution the width 0.84 Å, which closely matches the σ^{opt}_{ν} calculated for ν = 4, 5.

As shown by Figure S10 in the SI, one obtains at σ^{opt}_{ν} for all TLνP models an excellent correlation between the DFT/PMM and PMM dipole moments, if the above MT/BP values μ^{l}_{DFT} and α^{g}_{DFT} are chosen instead of the experimental values given by eq 1 (or, in the case of TL3P, instead of the empirically optimized value μ^{l} = 2.091 D). Correspondingly, the respective liquid phase distributions $p(\mu)$ of the dipole moments are almost indistinguishable.

Electrostatic Geometries from DFT/PMM. As described in section 3.1, electrostatic geometries of the TLνP models were calculated for ν = 4, 5 by minimizing the root-mean-square deviation $\zeta(l_{OL}, \varphi_{LOL})$ defined by eq 7. Figures S11 and S12 in section S3 of the SI show the deviation landscapes ζ calculated for the ensembles $\mathcal{S}_\nu$ of the TLνP models for ν = 5 and 4, respectively, as functions of the parameters (l_{OL}, φ_{LOL}) and l_{OL}. The figures reveal distinct minima at the locations listed in Table 4.

For TL4P, the minimum of $\zeta(l_{OM})$ is found at l_{OM} = 0.242 Å, i.e. the charge q_M = −1.120 e. The nonpolarizable TIP4P/2005

Table 4. G_e and Charge q_H for ν = 4, 5

model	l_{OL} [a]	φ_{LOL} [b]	q_H [c]
TL4P	0.242	360.0	0.560
TL5P	0.323	268.2	0.533

[a] Å. [b] deg. [c] e.

model[33] (q_M = −1.11 e, l_{OM} = 0.1546 Å) depicted in Figure2B has a similar charge, but for TL4P the M site is located at a considerably larger distance l_{OM} from the oxygen. Correspondingly, the static dipole moment μ^{l} = μ^{g}_{exp} = 1.855 D is 20% smaller than that of the mean field model TIP4P/2005.

The TL4P value of l_{OM} is close to the values of 0.22 Å and 0.238 Å determined for the empirical DO four-point models COS/G2[45] and SWM4-NDP,[47] respectively. It is smaller than that of the empirically optimized DO four-point model COS/D (0.257 Å),[49] which is the most recent variant in a series of "charge on spring (COS)" models[45,73] developed by the group of van Gunsteren. Also, the ID four-point model of Paricaud et al.[48] and the DO four-point model of Baranyai and Kiss,[50] which represent the partial charges by Gaussian distributions, use larger values $l_{OM} \approx$ 0.27 Å.

Figure 5 visualizes the electrostatic geometry $\mathbf{G}_e$ of the TL5P model optimally matching the surface potential of a DFT water

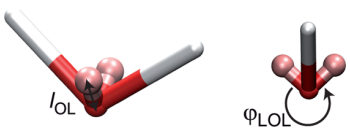


Figure 5. G_e of TL5P (cf. Table 4).

molecule embedded in a TL5P environment. A comparison with Figure 2 immediately reveals a distinct difference from the well-known TIP5P geometry. Whereas for TIP5P the projection of the lone-pair charges q_L on the molecular plane hits the HOH bisectrix outside the triangle of the water molecule, the projection is within that triangle for TL5P. Interestingly, the FQ six-point model,[63] which was derived from RI-MP2 ab initio calculations[112] on water clusters, and features two out-of-plane lone-pair charges q_L in addition to an in-plane lone-pair charge q_M on the bisectrix 0.29 Å distant from the oxygen, shows the q_L at locations similar to those of TL5P, i.e. with a projection on the molecular plane hitting the bisectrix inside the HOH triangle at a distance of 0.231 Å from r_O (TL5P: 0.225 Å).

Inspired by the recent interest in highly complex polarizable six-point models,[63,64] we have tentatively computed the electrostatic geometry of such a model (TL6P$_{ini}$) by the DFT/PMM procedure outlined in section 3.1 using the structural ensemble $\mathcal{S}_4$ for the PMM environment. The result is depicted in Figure S18 of the SI. Beyond a Gaussian dipole at the oxygen atom, whose width was adopted from TL4P, the model features positive charges q_H = 0.516 e at the hydrogens, a negative charge q_M = −0.570 e on the bisectrix at a rather large distance of 0.40 Å from r_O, and two lone-pair charges q_L = −0.231 e at a distance of 0.46 Å from r_O, whose projection on the molecular plane hits the bisectrix *outside* the HOH triangle at a distance of 0.02 Å. As a result, the center of the negative charge distribution is on the bisectrix at the quite small distance of 0.15 Å from r_O inside the molecular triangle. Thus, it is closer to the oxygen than for TL4P (0.242 Å) and TL5P (0.225 Å) and the electrostatic geometry of TL6P$_{ini}$ distinctly differs from

dx.doi.org/10.1021/jp404548k | J. Phys. Chem. B 2013, 117, 9486−9500

those of the other recent suggestions. The associated quadrupole moments match the experimental data quite well (cf. Table 6). A self-consistent optimization of such a highly complex model was, however, outside the scope of this work.

van der Waals Parameters from Empirical Optimization. Table 5 lists the parameters (A_1, A_2, B) of the Buckingham[69] potentials (eq 6) obtained by the weak-coupling PMM-MD simulations described in sections 2.4 and 3.1. For compatibility with biomolecular force fields, the table additionally gives the parameters (ε, σ) of almost equivalent Lennard-Jones potentials $E_{LJ}(r) = 4\varepsilon[(\sigma/r)^{12} - (\sigma/r)^6]$, which have the same dispersion attraction and closely fit the repulsive part of the associated Buckingham potentials in the range $r \in [2.5,6]$ Å.

Table 5. Buckingham and Associated Lennard-Jones Parameters

model	A_1[a]	A_2[a]	B[c]	ε[d]	σ[e]
TL3P	302100	4.17	486	0.11	3.23
TL4P	84120	3.55	992	0.29	3.12
TL5P	64300	3.40	1180	0.30	3.15
GCPM[48]	66948	3.46	1042	0.28	3.16

[a] Å^{12} kcal/mol. [b] Å^{-1}. [c] Å^6 kcal/mol. [d] kcal/mol. [e] Å.

Interestingly, Paricaud et al.[48] found for their "Gaussian charge polarizable model (GCPM)", which has the same static dipole moment μ^g_{exp} as TL4P, through an entirely empirical optimization, Buckingham parameters (cf. Table 5) similar to those of TL4P and TL5P. GCPM is very similar to TL4P, indeed. The main difference is that TL4P combines a single inducible Gaussian dipole (α = 1.47 Å^3) of width 0.842 Å with three static point charges, whereas GCPM employs three static Gaussian charges (q_M, q_H) of widths σ_M = 0.61 Å and σ_H = 0.46 Å with an inducible point dipole (α = 1.444 Å^3) located at the center of mass (i.e., GCPM is a much more complex five force point model).

Here the key difference of the parametrization strategies should be stressed once again. In the optimization of GCPM the Gaussian widths (σ_M, σ_H) are additional targets of an empirical search in a six-dimensional parameter space.[48] For TL4P, in contrast, we determine the Gaussian width σ of the inducible dipole and the distance l_{OM} from DFT/PMM calculations individually through separate procedures. Hence, the use of DFT/PMM and the choice of a single Gaussian distribution reduce the parameter space, which is to be scanned during the empirical optimization by three dimensions.

TL3P Dipole Moment from Empirical Optimization. As emphasized in sections 1.2, 2.5, and 3.1, the dipole moment μ^l of a polarizable three-point model cannot be chosen as μ^g_{exp} but must be empirically optimized instead. For TL3P we found μ^l = 2.091 D, which is by 13% larger than μ^g_{exp}. It is close to the value of 2.07 D empirically determined[73] for the polarizable DO three-point model COS/B2, which applies a reduced polarizability α_{eff} = 0.93 Å^3. Furthermore, also Straatsma et al.[72] found for their three-point DO models, which employ a large polarizability of 1.445 Å^3, a static dipole moment, which is larger (1.95 D) than μ^g_{exp}. Note here that Table S9 in section S4 of the SI collects all parameters of the TLνP models for a quick overview.

TLνP Quadrupole Moments. For TL3P we have used empirical liquid phase properties to compute the static dipole moment μ^l and, hence, the partial charges Q. For TL4P and TL5P, in contrast, we chose $\mu^l = \mu^g_{exp}$ = 1.855 D, and determined the associated electrostatic geometry $\mathbf{G}_e$ and, hence, Q from the supposedly static part $\Phi(\mathbf{r}\,|\,\rho_s^{stat})$ of the potential, which is calculated by DFT/PMM for an ensemble $\mathcal{S}_\nu$ of liquid structures s on a spherical surface surrounding the respective DFT fragment. Therefore, the quadrupole moments of TL4P and TL5P should be much closer to the corresponding experimental data[113] on water molecules isolated in the gas phase than that of TL3P.

Table 6 shows that this expectation is supported by the data. As demonstrated by the root-mean-square deviation R from the experimental data, TL5P is seen to yield an excellent match. While the quadrupole moments of TL4P are still very close, sizable deviations are apparent for TL3P.

Table 6. Quadrupole Moments in Units of DÅ and Their Root-Mean-Square Deviations R from the Experimental Gas Phase Data[113] for the TLνP and Various Other PMM Models[a]

model	ν_f	Q_{xx}[b]	Q_{yy}[b]	Q_{zz}[b]	R
exp[113]		−0.13	2.63	−2.50	
TL3P	3	−0.10	1.63	−1.53	0.80
TL4P	3	−0.29	2.53	−2.24	0.19
TL5P	4	−0.11	2.53	−2.42	0.07
TL6P$_{ini}$	5	−0.17	2.57	−2.40	0.07
COS/B2[73]	3	−0.27	1.93	−1.66	0.64
COS/G2[45]	3	−0.20	2.27	−2.07	0.33
COS/D[45]	3	−0.30	2.57	−2.27	0.17
SWM4-DP[46]	3	−0.24	2.41	−2.16	0.24
GCPM[48]	3	−0.37	2.69	−2.36	0.16
SWM6[64]	5	−0.30	2.39	−2.09	0.29

[a] ν_f is the number of force points. The preliminary attempt TL6P$_{ini}$ to develop a six-point PMM model is described in section 4.1 and in section S11 of the SI. [b] DÅ.

For comparison we have included data on the quadrupole moments of several empirically parametrized PMM water models. As a measure for the computational cost, we have indicated the number of force points ν_f.

The simple three-point DO model COS/B2[73] features a large deviation R comparable to that of the Gaussian ID model TL3P. This finding underlines the critique of three-point models in the discussion of Figure 1 concerning a supposedly poor performance on higher multipole moments. Among the more complex empirical models, GCPM,[48] with its three Gaussian charge distributions, provides the closest description of the gas phase quadrupole moments. It is comparably as good as that of TL4P.

The quality by which TL5P and TL4P describe the quadrupole moments of an isolated water molecule nourishes the hope that these PMM water models should also deliver reasonable descriptions of the local structures around a water molecule in the liquid phase. For TL3P, on the other hand, one expects certain errors.

4.2. Evaluation. The TLνP models have been constructed for liquid water at standard conditions. One cannot expect that they perform equally well at grossly different conditions such as those provided by the isolated dimer. To explore their transferability, we have, nevertheless, evaluated the dimer geometry, the total dipole moment, and the binding energy. These results are presented and discussed, using comparisons with experimental data[113−115] and with predictions of other PMM models,[45,47,48,64] in section S5 of the SI. The results

dx.doi.org/10.1021/jp404548k | J. Phys. Chem. B 2013, 117, 9486−9500

document a poor performance of TL3P on almost all dimer properties, particularly on the dipole moment, while those of TL4P and TL5P are mixed, i.e. quite good with respect to the dipole moment and binding energy and a little off concerning the geometry. The oxygen−oxygen distances calculated for the cyclic TLνP trimers are compared in section S6 of the SI with corresponding experimental data.[103] As compared to the dimers, the oxygen−oxygen distances of our cyclic PMM trimers show an improved performance, whose significance concerning issues of transferability is discussed in section S6 of the SI. However, much more important is the performance of the TLνP models on the properties of liquid water at standard conditions.

Distributions of the Dipole Moments. By presenting the distributions $p(\mu)$ of the dipole moment μ calculated from the $N_mV_mT_0$ trajectories, Figure 6 gives a first insight into the

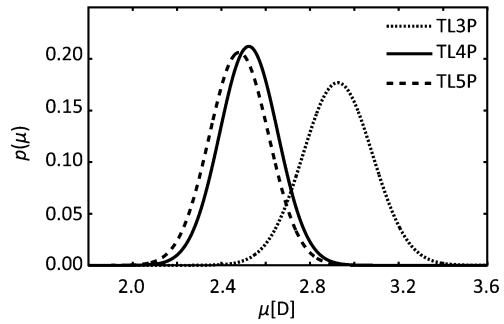


Figure 6. Dipole moment distributions $p(\mu)$ of the TLνP models from the $N_mV_mT_0$ trajectories.

properties of the TLνP models. For TL5P, TL4P, and TL3P, the $p(\mu)$ are normal distributions with the centers 2.480 D, 2.523 D, and 2.925 D and standard deviations 0.135 D, 0.131 D, and 0.136 D, respectively. Thus, the empirically optimized static dipole μ^l = 2.091 D of TL3P, which is 13% larger than that of TL4P and TL5P, leads to an average dipole moment $\langle\mu\rangle$ exceeding that of the more complex models by 16−18%. Recall here that TL4P and TL5P employ for μ^l and α^l the experimental gas phase values.

The TL4P and TL5P values of $\langle\mu\rangle$ are about as small as those of SWM6,[64] COS/D,[49] and BK,[51] which up to now were the PMM models with the smallest values of $\langle\mu\rangle$. GCPM,[48] for instance, has a larger $\langle\mu\rangle$ value of 2.72 D. The values $\langle\mu\rangle$ = 2.523 D and $\langle\mu\rangle$ = 2.480 D of TL4P and TL5P, respectively, are both within the "range of 2.4−2.6, which has been suggested to get the correct dielectric constant for simple water models[116,46]".[64] It will be interesting to check whether the quoted connection between $\langle\mu\rangle$ and the dielectric constant ε also holds for our TLνP models.

As shown and discussed in section S9 of the SI, the TL4P and TL5P distributions $p(\mu)$ comply quite well with corresponding data calculated by DFT/PMM for identical structural ensembles. Differences are explained by the facts that MT/BP overestimates the polarizability of a water molecule by 7.5% and underestimates its static dipole moment by 3.5%. For TL4P and TL5P the average dipole moments $\langle\mu\rangle$ and standard deviations σ_μ are 2% and 25%, respectively, smaller than the corresponding properties calculated for their DFT/PMM counterparts. For TL3P, in contrast, $\langle\mu\rangle$ overestimates the DFT/TL3P result of 2.675 D by 9.3% because of the large, empirically fitted dipole moment μ^l = 2.091 D.

Properties Targeted by Optimization. The empirical $N_mV_mT_0$ optimizations of the Buckingham parameters (A_1, A_2, B) and, for TL3P, also of the static dipole moment μ^l were executed as described in section 3.1. Table S12 in the SI demonstrates that all TLνP models closely match the experimental target $E_{pot}(p_0, T_0)$ in the $N_iV_iT_0$[BU], $i \in$ {s, m, l}, simulations executed at $n_{exp}(p_0, T_0)$ (cf. Table 2). Table S12 additionally shows that the average pressure $\langle p\rangle$ observed in these simulations is close to the target pressure p_0 = 1 atm. According to Table S13, the average densities $\langle n\rangle$ coincide with $n_{exp}(p_0, T_0)$ within the statistical errors, which were calculated by block-averaging.[117] Solely the small TL4P simulation system shows a slightly larger deviation.

As explained in section 3.1, the location of the first peak in Soper's[10] neutron diffraction RDF $g_{OO}(r)$ served as a further criterion in the empirical part of the parameter optimization. As expected, the RDFs $g_{OO}(r \mid N_iV_iT_0)$ values (cf. section 3.2) were independent of N.

Using the medium sized systems as representatives, Figure 7 now demonstrates that the first peak of the RDF $g_{OO}(r)$ is at

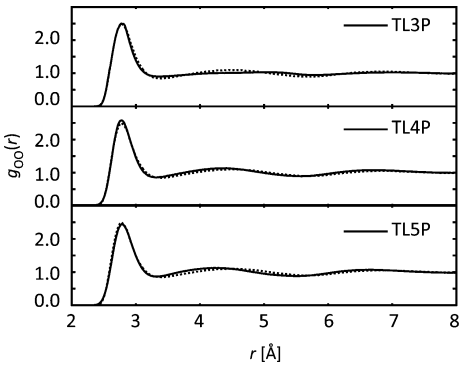


Figure 7. TLνP oxygen−oxygen RDFs $g_{OO}(r \mid N_mV_mT_0)$ (solid lines) are compared with the neutron diffraction result of Soper[10] (dotted lines), which has a peak of the height 2.50 at 2.76 Å.

the correct position for all TLνP models. In addition, the onset of Soper's RDF near 2.5 Å and the height of the first peak are nicely reproduced by all TLνP simulation results due to the use of a Buckingham potential. Furthermore, the TL4P and TL5P results follow the experimental reference also at distances beyond the first peak, if one neglects small differences concerning the locations of the first minimum and of the second peak. Larger deviations are present for TL3P. As is typical for three-point models (cf. section 1.2), the RDF of TL3P is almost flat after the first peak. The RDF of TL5P, in contrast, features a somewhat too pronounced structure in this region, because it erroneously shifts the second peak from 4.6 Å to 4.3 Å . The recent PMM model SWM6,[64] in contrast, while missing the shape of the first peak provides a good description of Soper's RDF at distances beyond 3.5 Å.

Resulting Bulk Properties. For the properties of the TLνP bulk liquid discussed up to now, the observed good match with the corresponding experimental data is an immediate consequence of the empirical parameter optimization, which targeted the parameters of the Buckingham potential (eq 6) and, for TL3P, also μ^l. Comparisons of further properties with experimental data, however, can shed light on the predictive power of the models.

dx.doi.org/10.1021/jp404548k | J. Phys. Chem. B 2013, 117, 9486−9500

A most important property of liquid water is its large dielectric constant ε, which has the value[96] 78 at T_0 and p_0. In our simulation setting, which employs a Kirkwood reaction field characterized by the dielectric constant ε_{RF} = 78 to cover the long-range part of the electrostatics,[91] initial guesses ε_i^i for the dielectric constant of the considered water model are calculated through 8 from the fluctuations of the total dipole $\mathbf{M}_i$ observed in the simulation systems $i \in \{s, m, l\}$ during the $N_iV_iT_0$[BU] simulations (cf. Table 2). Applying the iterative correction, which is described in ref 107 and is based on perturbation theory, until self-consistency is reached yields the final guesses ε_i^f.

In Figure 8 the values ε_i^i and ε_i^f, which resulted from these simulations, are plotted over the inverse edge lengths L_i^{-1} of the

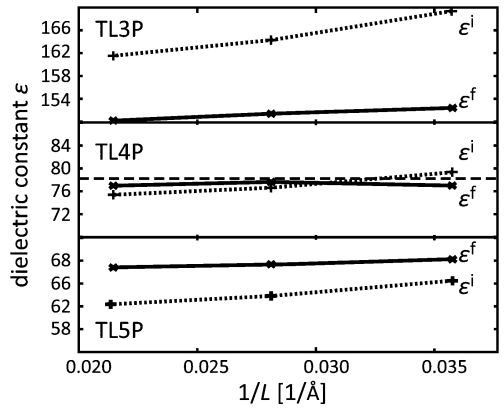


Figure 8. Size dependences of the initial (ε^i, dotted lines) and final (ε^f, solid lines) guesses for the dielectric constants of the TLνP models obtained from the $N_iV_iT_0$[BU] simulations.

simulation systems. Whereas the initial guesses ε^i decrease with increasing system size for all models, the final values ε^f show a much weaker size dependence, which should disappear, if ε^f is calculated by resampling with a self-consistently chosen ε_{RF}. The case of TL4P, in which the calculated dielectric constant is always very close to ε_{RF}, provides evidence for this claim, because here the corrected values ε^f do not show (up to remaining statistical errors) any size dependence.

To roughly estimate the statistical errors of the data displayed by Figure 8, we have partitioned all $N_iV_iT_0$[BU] trajectories, which were used for the evaluation of the average values $\varepsilon_i^{i/f}$, into two parts and have taken the deviations from the averages as measures for the associated standard deviations $\sigma_\varepsilon^{i/f}$. Interestingly, the average statistical errors $\langle\sigma_\varepsilon^i\rangle_{i,\nu}$ of the initial estimates ε^i obtained for the three system sizes i and the three TLνP models were about 3, whereas for the final values ε^f we found much smaller errors $\langle\sigma_\varepsilon^f\rangle_{i,\nu} \approx 1$. The errors were particularly small for the final values ε^f of the TL4P model, where we found $\langle\sigma_\varepsilon^f\rangle_i = 0.5$. Hence, the lacking size dependence of ε^f revealed by Figure 8 is statistically well founded.

In all cases the size dependence of the corrected values ε^f is weak enough to estimate the dielectric constants ε_0 of the models at standard conditions for infinitely large systems. Linear extrapolation yields ε_0(TL3P) = 147, ε_0(TL4P) = 77, and ε_0(TL5P) = 67. The latter two values are lower bounds, because the iterative correction leads to larger values of ε^f for large systems. In contrast, for TL3P it is an upper bound, because here the correction reduces the estimates of ε. For TL4P the lines associated to ε^i and ε^f intersect at a certain system size. Here the correction vanishes and, therefore, the corresponding value ε = 77 = ε_0 is a good estimate for the dielectric constant of the TL4P model.

Thus, one may state that the dielectric constant of TL4P is very close to that of liquid water, whereas that of TL5P is a little smaller. As is common for polarizable three-point models,[73] TL3P grossly overestimates the dielectric constant of water. We would like to stress that we are unaware of any previous attempt to determine the size dependence of ε from simulations. It may well be that there is no such size dependence under Ewald boundary conditions. But the dependence is pronounced for RF methods with an inconsistently chosen ε_{RF}. Note furthermore that the values of ε_0 determined here agree with the suggested connection[116] between ε and the average dipole moment $\langle\mu\rangle$ (cf. section 4.2), which is for the TL4P and TL5P models in the required range of 2.4−2.6 D. Correspondingly, their dielectric constants ε_0 are reasonably close to the experimental value.

Figure 9 shows how the diffusion coefficients $D(N_i,V_i,T_0)$ of the TLνP models vary with the inverse edge lengths L_i^{-1} of the

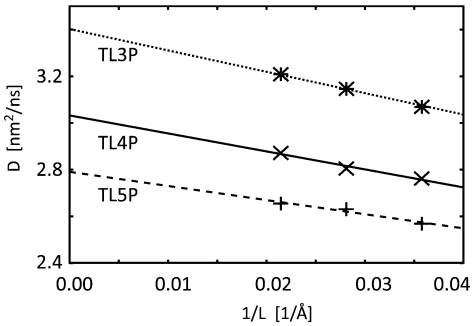


Figure 9. Size dependence of the diffusion coefficients obtained through eq 9 from the $N_iV_iT_0$[MI] simulations.

simulation systems. As expected from formula 10 of Dünweg and Kremer,[108] which Yeh and Hummer[109] exemplified for the standard nonpolarizable TIP3P[104] model, the diffusion coefficients linearly increase toward larger systems, where they approach the limiting values D_0 of the infinite systems. The viscosities η then follow from the slopes of the straight lines.

One finds for TL5P the slight overestimate D_0 = 2.8 nm^2/ns of the experimental value[118] (2.4 nm^2/ns at T_0 = 300 K) and for TL4P and TL3P increasing overestimates of 3.0 and 3.4 nm^2/ns, respectively (cf. Table 7). In line with the increasing diffusion coefficients, the viscosities η decrease from 0.98 mPa s (TL5P) over 0.80 mPa s (TL4P) down to 0.68 mPa s (TL3P), which have to be compared with the experimental value[119] of 0.81 mPa s (at 300 K). Thus, TL5P and TL4P generate viscosities close to that of liquid water at (T_0, p_0, n_{exp}), whereas the viscosity of TL3P is considerably too small.

The increasing diffusivity and decreasing viscosity in the sequence TL5P, TL4P, and TL3P agrees with the decreasing structure in the corresponding RDFs (cf. Figure 7). Similarly, Yeh and Hummer[109] determined for TIP3P, whose RDF $g_{OO}(r)$ has almost no structure beyond the first peak, for D_0 the large value of 6.1 nm^2/ns and for η the small value of 0.31 mPa s.

Exploiting and confirming the size dependence[109] of D, Tazi et al.[120] recently determined D_0 and η for the well-known nonpolarizable water models SPC/E[9] and TIP4P/2005.[33] For

dx.doi.org/10.1021/jp404548k | J. Phys. Chem. B 2013, 117, 9486−9500

Table 7. Bulk Properties of the TLνP Models Determined from the MD Simulations Listed in Tables 2 and 3 by the Procedures Described in the Text[a]

	TL3P	TL4P	TL5P	exp
ε_0	147	77	67	78[96]
D_0[b]	3.4	3.0	2.8	2.4[118]
η[c]	0.61	0.80	0.98	0.81[119]
κ_T[d]	48.0	37.4	36.4	45.6[87]
$\sigma(\kappa_T)$[d]	0.5	0.3	0.2	
C_p[e]	18.5	18.7	19.6	18.0[121]
$\sigma(C_p)$[e]	0.4	0.2	0.2	
α_p[f]	8.7	5.9	7.2	2.8[87]
$\sigma(\alpha_p)$[f]	0.7	0.5	0.4	

[a]For κ_T, C_p, and α_p, estimates of the statistical errors were calculated by block-averaging. [b]nm^2/ns. [c]mPa s. [d]10^{-6}/atm. [e]cal/(mol K). [f]10^{-4}/K.

(D_0, η) they found (3.0 nm^2/ns, 0.64 mPa s) (SPC/E) and (2.5 nm^2/ns, 0.83 mPa s) (TIP4P/2005), implying that TIP4P/2005 matches the experimental diffusivity and viscosity very well. Tazi et al.[120] additionally studied the polarizable DC[41] model and found (D_0, η) = (2.7 nm^2/ns, 0.78 mPa s), i.e. values close to those of our TL4P model.

Due to the neglected size dependence and the use of quite small simulation systems, diffusion constants published for other models[15,45−47,49,73] are of limited value. For instance, using the TIP3P result of Yeh and Hummer,[109] Vega and Abascal[15] estimated the error induced by the neglected size dependence into D to be about 10% for a system with 500 water models. However, a comparison of their finite size diffusion coefficients with the more recent results of Tazi et al.[120] on D_0 for SPC/E and TIP4P/2005 demonstrates that the errors are closer to 20%.

The isothermal compressibility κ_T was calculated through eq 11 from the simulations $N_s V_\pm T_0$ (cf. Table 2). Comparing the results listed in Table 7 with the given experimental reference, we recognize small underestimates for the more structured TL4P and TL5P liquids and an overestimate for TL3P. The polarizable COS models[45,49,73] show a similar behavior. While the three-point model COS/B2[73] overestimates κ_T just like TL3P, the more complex four-point models COS/G2[45] and COS/D[49] underestimate κ_T just like TL4P and TL5P.

Because we use stiff models for the water molecules, which, at T_0, are essentially frozen in their vibrational ground states, the heat capacities C_p calculated through eq 12 from the classical $N_s p_0 T_\pm$ simulations (cf. Table 3) can be in the correct range. According to Table 7, TL4P and TL3P predict C_p within 3% whereas the more structured liquid TL5P yields a 9% overestimate. With the exception of TIP3P, whose C_p is equal to that of TL4P,[15] standard nonpolarizable water models as well as earlier polarizable models[45,73] of the COS series generally overestimate[15] C_p by 10%−60%. In contrast, the advanced COS/D[49] four-point model yields a sizable underestimate of 17%.

Whereas all predictions of bulk properties obtained so far with the TL4P and TL5P models were in reasonable or even in very good agreement with observations on the bulk liquid at (T_0, p_0), we finally turn to one quantity, which is missed by the TLνP models. This quantity is the thermal expansion coefficient α_p at constant pressure. It was derived by eq 13 from the same $N_s p_0 T_\pm$ simulations used to compute C_p. A glance at Table 7 demonstrates that the TLνP models overestimate α_p by factors 1.5−3.1, indicating that the decrease of the density $n(T)$ with increasing temperature T is much too steep at T_0. Section S7 of the SI documents our preliminary TLνP results on $n(T)$ for $T \in [230\ \text{K}, 310\ \text{K}]$, compared with the experimental profile $n_{\text{exp}}(T)$,[87] and it argues that the apparent deviations are a common feature of polarizable models. Possible sources[66] of these deficiencies are discussed, and a thorough study aiming at their removal is envisaged.

5. SUMMARY AND OUTLOOK

The empirical parametrization of complex PMM models for complicated liquids such as water requires parameter searches in high-dimensional spaces and diligent choices of reliably measured liquid phase target properties. As we have shown above, the dimension of the parameter space, which has to be empirically scanned by sample simulations, can be considerably reduced by adding information on the liquid phase properties of the molecules through the use of DFT/MM and DFT/PMM hybrid calculations.

Previous DFT/MM studies of water molecules in the liquid phase had already demonstrated that the polarizing local fields should be calculated as molecular volume averages[26] and that the polarizability and static dipole moment are essentially invariant in the transition from the gas phase to the liquid phase geometry.[79] Consequently, Gaussian distributions of an appropriate width σ and the experimental gas-phase polarizability $\alpha^{\text{g}}_{\text{exp}}$ = 1.47 Å^3 should be used to model the induced dipoles. In addition, for sufficiently complex models featuring more than three force points, the gas phase value $\mu^{\text{g}}_{\text{exp}}$ = 1.855 D should be used for the static dipole moment in the liquid (i.e., $\mu^l = \mu^{\text{g}}_{\text{exp}}$).

Adopting these results and applying a newly developed DFT/PMM technology,[84] we have now demonstrated that the Gaussian widths σ of the induced dipoles and the electrostatic geometries of sufficiently complex PMM models can be self-consistently adjusted to the properties of DFT models in PMM environments. The thus achieved dimension reduction simplifies and stabilizes the empirical optimization of the remaining model parameters.

As an example we have developed and evaluated the TLνP models, which have ν = 3, 4, and 5 force points, a Gaussian induced dipole and a Buckingham potential at the oxygen, and the experimental liquid phase geometry $\mathbf{G}_{\text{m}}{}^l$ (Table 2). Results of MD simulations show that PMM models determined by such a mixed empirical and DFT/PMM computational parametrization approach can feature bulk phase properties comparable to or even better than those of similarly complex models, which were obtained by purely empirical optimization.

The TLνP models accurately reproduce the properties (E_{pot}, p_0) targeted by the empirical part of the optimization (cf. Table S12 in the SI) and their RDF's look reasonable (cf. Figure 7) at each degree of model complexity. Np_0T_0 simulations, in particular, quite precisely yield the experimental value n_{exp} = 0.997 g/cm^3 for the average density $\langle n \rangle$ (see Table S13 in the SI). Moreover, as demonstrated by Table 7, TL4P and TL5P predict quite reasonable values for the dielectric constant ε, the diffusion coefficient D_0, the viscosity η, the isothermal compressibility κ_T, and the isobaric heat capacity C_p of water at (T_0, p_0). Concerning these properties, the TL3P model shares the well-studied deficiencies[72−76] of previous PMM three-point models, which arise from their too limited degree of model complexity.

dx.doi.org/10.1021/jp404548k | J. Phys. Chem. B 2013, 117, 9486−9500

Furthermore the electrostatic properties of TL4P and TL5P, such as the arrangements of the static partial charges and the resulting quadrupole moments, which were derived from self-consistent DFT/PMM calculations, appear to be definitely better than those of comparably (or even of more) complex empirical models. According to Figure 6, for instance, the highly complex empirical SWM6 model has the root-mean-square deviation R = 0.29 DÅ from the experimental quadrupole moments, which served as targets during the joint optimization of its eight model parameters, whereas the less complex TL4P and TL5P models (as well as the preliminary $TL6P_{ini}$ attempt) show much smaller deviations.

The surprising electrostatic geometry of the tentative six-point model $TL6P_{ini}$ described in section 4.1 and depicted in Figure S18 of the SI bears the potential of offering an improved transferability to conditions other than the bulk liquid at (T_0, p_0). This hope is nourished by the improved geometry of the $TL6P_{ini}$ water dimer as compared to the other TLνP dimers (see sections S5 and S11 in the SI for discussions and graphical representations).

The enhanced computational effort, which is associated for such a six-point model with the empirical part of the parameter optimization and with the evaluation of the liquid phase properties, requires a highly efficient and parallelized MD code. Such a code is provided by the PMM-MD program package IPHIGENIE[85] used in the present study, such that large scale MD simulations of highly complex PMM models have become computationally feasible. In addition IPHIGENIE offers a Hamiltonian DFT/PMM interface.[84] Correspondingly, we now have the means to optimize also a TL6P model by our self-consistent DFT/PMM parametrization procedure.

The computational efficiency of IPHIGENIE additionally opens the chance to systematically tackle the questions associated with the density−temperature profile of PMM models by large scale simulation studies. Here, the failure of the TLνP models to predict reasonable temperature profiles $n(T)$ or values for the thermal expansion coefficient α_p at T_0 = 300 K has already indicated that the model design still lacks a key ingredient, which can enable the PMM models to reproduce the observed density maximum.[87] Meanwhile this lacking ingredient may have been identified (cf. section S7 in the SI), if the recent suggestion of Kiss and Baranyai[66] of using polarizable van der Waals potentials finds further support.

■ ASSOCIATED CONTENT

Supporting Information

Nine figures (S10−S18), six tables (S8−S13), and various pieces of text explaining and documenting various issues. First it sketches the optimization of the widths $\tilde{\sigma}_i$ of the Gaussian partial charge distributions used in the DFT/PMM interface, which is based on DFT/PMM calculations of the water dimer. Next it presents the liquid-phase correlations of the induced DFT/PMM and PMM dipole moments and the optimization of the TL4P and TL5P electrostatic geometries. Also given are the parameters, the dimer and trimer properties, and the temperature density profiles of the TLνP models. Furthermore, experimental RDFs are presented and discussed, the DFT/PMM and PMM dipole distributions of the TLνP models are given together with tables of targeted observables and an initial guess of a TL6P model. This material is available free of charge via the Internet at http://pubs.acs.org/.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: tavan@physik.uni-muenchen.de.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work has been supported by the Deutsche Forschungsgemeinschaft (SFB 749/C4).

■ REFERENCES

(1) Guillot, B. *J. Mol. Liq.* **2002**, *101*, 219−260.
(2) van Gunsteren, W. F.; et al. *Angew. Chem., Int. Ed.* **2006**, *45*, 4064−4092.
(3) Tavan, P.; Carstens, H.; Mathias, G. In *Protein Folding Handbook*; Buchner, J., Kiefhaber, T., Eds.; Wiley-VCH: Weinheim, 2005; Vol. *1*; pp 1170−1195.
(4) MacKerell, A. D.; et al. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.
(5) Pearlman, D.; Case, D.; Caldwell, J.; Ross, W.; Cheatham, T., III; DeBolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. *Comput. Phys. Commun.* **1995**, *91*, 1−41.
(6) Christen, M.; Hünenberger, P.; Bakowies, D.; Baron, R.; Bürgi, R.; Geerke, D.; Heinz, T.; Kastenholz, M.; Kräutler, V.; Oostenbrink, C.; Peter, C.; Trzesniak, D.; van Gunsteren, W. *J. Comput. Chem.* **2005**, *26*, 1719−1751.
(7) Jorgensen, W.; Chandrasekhar, J.; Madura, J.; Impey, R.; Klein, M. *J. Phys. Chem.* **1983**, *79*, 926−935.
(8) Berendsen, H.; Postma, J.; Van Gunsteren, W.; Hermans, J. *Intermol. Forces* **1981**, *11*, 331−342.
(9) Berendsen, H.; Grigera, J.; Straatsma, T. *J. Phys. Chem.* **1987**, *91*, 6269−6271.
(10) Soper, A. *Chem. Phys.* **2000**, *258*, 121−137.
(11) Soper, A. *J. Phys.: Condens. Matter* **2007**, *19*, 335206.
(12) Hura, G.; Sorenson, J.; Glaeser, R.; Head-Gordon, T. *J. Chem. Phys.* **2000**, *113*, 9140−9148.
(13) Hura, G.; Russo, D.; Glaeser, R.; Head-Gordon, T.; Krack, M.; Parrinello, M. *Phys. Chem. Chem. Phys.* **2003**, *5*, 1981−1991.
(14) Skinner, L. B.; Huang, C.; Schlesinger, D.; Pettersson, L. G.; Nilsson, A.; Benmore, C. J. *J. Chem. Phys.* **2013**, *138*, 074506.
(15) Vega, C.; Abascal, J. *Phys. Chem. Chem. Phys.* **2011**, *13*, 19663−19688.
(16) Lennard-Jones, J. E. *Proc. Phys. Soc.* **1931**, *43*, 461−482.
(17) Verlet, L. *Phys. Rev.* **1967**, *159*, 98−103.
(18) VandeVondele, J.; Tröster, P.; Tavan, P.; Mathias, G. *J. Phys. Chem. A* **2012**, *116*, 2466−2474.
(19) Sorenson, J.; Hura, G.; Glaeser, R.; Head-Gordon, T. *J. Chem. Phys.* **2000**, *113*, 9149−9161.
(20) Clough, S.; Beers, Y.; Klein, G.; Rothman, L. *J. Chem. Phys.* **1973**, *59*, 2254−2259.
(21) Sprik, M. *J. Phys. Chem.* **1991**, *95*, 2283−2291.
(22) Silvestrelli, P.; Bernasconi, M.; Parrinello, M. *Chem. Phys. Lett.* **1997**, *277*, 478−482.
(23) Silvestrelli, P.; Parrinello, M. *J. Chem. Phys.* **1999**, *111*, 3572−3570.
(24) Eichinger, M.; Tavan, P.; Hutter, J.; Parrinello, M. *J. Chem. Phys.* **1999**, *110*, 10452−10467.
(25) Gubskaya, A.; Kusalik, P. *J. Chem. Phys.* **2002**, *117*, 5290−5302.
(26) Schropp, B.; Tavan, P. *J. Phys. Chem. B* **2008**, *112*, 6233−6240.
(27) Badyal, Y.; Saboungi, M.; Price, D.; Shastri, S.; Haeffner, D.; Soper, A. *J. Chem. Phys.* **2000**, *112*, 9206−9208.
(28) Babitzki, G.; Denschlag, R.; Tavan, P. *J. Phys. Chem. B* **2009**, *113*, 10483−10495.
(29) Lopes, P.; Roux, B.; MacKerell, A. *Theor. Chem. Acc.* **2009**, *124*, 11−28.
(30) Mahoney, M.; Jorgensen, W. *J. Chem. Phys.* **2000**, *112*, 8910−8922.

(31) Horn, H.; Swope, W.; Pitera, J.; Madura, J.; Dick, T.; Hura, G.; Head-Gordon, T. *J. Chem. Phys.* **2004**, *120*, 9665−9678.
(32) Sanz, E.; Vega, C.; Abascal, J.; MacDowell, L. *J. Chem. Phys.* **2004**, *121*, 1165−1166.
(33) Abascal, J.; Vega, C. *J. Chem. Phys.* **2005**, *123*, 234505.
(34) Huggins, D. J. *J. Chem. Phys.* **2012**, *136*, 064518.
(35) Niesar, U.; Corongiu, G.; Clementi, E.; Kneller, G.; Bhattacharya, D. *J. Phys. Chem.* **1990**, *94*, 7949−7956.
(36) Kozack, R.; Jordan, P. *J. Chem. Phys.* **1992**, *96*, 3120−3130.
(37) Corongiu, G. *Int. J. Quantum Chem.* **1992**, *42*, 1209−1235.
(38) Brodholt, J.; Sampoli, M.; Vallauri, R. *Mol. Phys.* **1995**, *86*, 149−158.
(39) Chialvo, A.; Cummings, P. *J. Chem. Phys.* **1996**, *105*, 8274−8281.
(40) Chialvo, A.; Cummings, P. *Fluid Phase Equilib.* **1998**, *150*, 73−81.
(41) Dang, L.; Chang, T. *J. Chem. Phys.* **1997**, *106*, 8149−8159.
(42) Kiyohara, K.; Gubbins, K. E.; Panagiotopoulos, A. *Z. Mol. Phys.* **1998**, *94*, 803−808.
(43) Svishchev, I.; Kusalik, P.; Wang, J.; Boyd, R. *J. Chem. Phys.* **1996**, *105*, 4742−4750.
(44) van Maaren, P. J.; van der Spoel, D. *J. Phys. Chem. B* **2001**, *105*, 2618−2626.
(45) Yu, H.; van Gunsteren, W. *J. Chem. Phys.* **2004**, *121*, 9549−9564.
(46) Lamoureux, G.; MacKerell, A., Jr.; Roux, B. *J. Chem. Phys.* **2003**, *119*, 5185−5197.
(47) Lamoureux, G.; Harder, E.; Vorobyov, I.; Roux, B.; MacKerell, A. *Chem. Phys. Lett.* **2006**, *418*, 245−249.
(48) Paricaud, P.; Předota, M.; Chialvo, A.; Cummings, P. *J. Chem. Phys.* **2005**, *122*, 244511.
(49) Kunz, A.; van Gunsteren, W. *J. Phys. Chem. A* **2009**, *113*, 11570−11579.
(50) Baranyai, A.; Kiss, P. *J. Chem. Phys.* **2010**, *133*, 144109.
(51) Baranyai, A.; Kiss, P. *J. Chem. Phys.* **2011**, *135*, 234110.
(52) Sprik, M.; Klein, M. *J. Chem. Phys.* **1988**, *89*, 7556−7554.
(53) Rick, S.; Stuart, S.; Berne, B. *J. Chem. Phys.* **1994**, *101*, 6141−6167.
(54) Chen, B.; Xing, J.; Siepmann, J. *J. Phys. Chem. B* **2000**, *104*, 2391−2401.
(55) Cieplak, P.; Kollman, P.; Lybrand, T. *J. Chem. Phys.* **1990**, *92*, 6755−6760.
(56) Burnham, C. J.; Li, J.; Xantheas, S. S.; Leslie, M. *J. Chem. Phys.* **1999**, *110*, 4566−4581.
(57) Burnham, C. J.; Xantheas, S. S. *J. Chem. Phys.* **2002**, *116*, 1500−1510.
(58) Ishiyama, T.; Morita, A. *J. Phys. Chem. C* **2007**, *111*, 721−737.
(59) Stern, H.; Rittner, F.; Berne, B.; Friesner, R. *J. Chem. Phys.* **2001**, *115*, 2237−2251.
(60) Viererblová, L.; Kolafa, J. *Phys. Chem. Chem. Phys.* **2011**, *13*, 19925−19935.
(61) Zhu, S.; Singh, S.; Robinson, G. *J. Chem. Phys.* **1991**, *95*, 2791−2799.
(62) Nada, H.; van der Eerden, J. *J. Chem. Phys.* **2003**, *118*, 7401−7413.
(63) Wang, L.-P.; Chen, J.; Van Voorhis, T. *J. Chem. Theory Comput.* **2012**, *9*, 452−460.
(64) Yu, W.; Lopes, P. E. M.; Roux, B.; MacKerell, A. D., Jr. *J. Chem. Phys.* **2013**, *138*, 034508.
(65) Ren, P.; Ponder, J. W. *J. Phys. Chem. B* **2003**, *107*, 5933−5947.
(66) Kiss, P. T.; Baranyai, A. *J. Chem. Phys.* **2012**, *137*, 084506.
(67) Dyke, T.; Muenter, J. *J. Chem. Phys.* **1973**, *59*, 3125−3127.
(68) Murphy, W. *J. Chem. Phys.* **1977**, *67*, 5877−5882.
(69) Buckingham, R. A.; Corner, J. *Proc. R. Soc. London, Ser. A* **1947**, *189*, 118−129.
(70) Wallqvist, A.; Berne, B. *J. Phys. Chem.* **1993**, *97*, 13841−13851.
(71) de Pablo, J. J.; Prausnitz, J. M.; Strauch, H. J.; Cummings, P. T. *J. Chem. Phys.* **1990**, *93*, 7355−7359.
(72) Straatsma, T. P.; McCammon, J. A. *Mol. Simul.* **1990**, *5*, 181−192.
(73) Yu, H.; Hansson, T.; Van Gunsteren, W. *J. Chem. Phys.* **2003**, *118*, 221−234.
(74) Caldwell, J.; Dang, L.; Kollman, P. *J. Am. Chem. Soc.* **1990**, *112*, 9144−9147.
(75) Dang, L. *J. Chem. Phys.* **1992**, *97*, 2659−2667.
(76) Caldwell, J. W.; Kollman, P. A. *J. Phys. Chem.* **1995**, *99*, 6208−6219.
(77) Elking, D.; Darden, T.; Woods, R. J. *J. Comput. Chem.* **2007**, *28*, 1261−1274.
(78) In ref 26 the value σ_{opt} = 0.7 Å is given for the optimal volume of Gaussian averaging. However, in the applied DFT/MM approach[24] the MM partial charges polarizing the DFT fragment are represented as Gaussians of widths σ_{MM} = 0.57 Å. Thus, to equivalently model a Gaussian dipole interacting with point charges, the effective width $\sigma = (\sigma_{opt}^2 + \sigma_{MM}^2)^{1/2}$ = 0.9 Å has to be employed.
(79) Schropp, B.; Tavan, P. *J. Phys. Chem. B* **2010**, *114*, 2051−2057.
(80) Thole, B. *Chem. Phys.* **1981**, *59*, 341−350.
(81) Benedict, W. S.; Gailar, N.; Plyler, E. K. *J. Chem. Phys.* **1956**, *24*, 1139−1165.
(82) Ichikawa, K.; Kameda, Y.; Yamaguchi, T.; Wakita, H.; Misawa, M. *Mol. Phys.* **1991**, *73*, 79−86.
(83) Thiessen, W. E.; Narten, A. H. *J. Chem. Phys.* **1982**, *77*, 2656−2662.
(84) Schwörer, M.; Breitenfeld, B.; Tröster, P.; Lorenzen, K.; Tavan, P.; Mathias, G. *J. Chem. Phys.* **2013**, *138*, 244103.
(85) Lorenzen, K.; Schwörer, M.; Tröster, P.; Mates, S.; Tavan, P. *J. Chem. Theory Comput.* **2012**, *8*, 3628−3636.
(86) Laio, A.; VandeVondele, J.; Rothlisberger, U. *J. Chem. Phys.* **2002**, *116*, 6941−6947.
(87) Kell, G. *J. Chem. Eng. Data* **1967**, *12*, 66−69.
(88) Jancso, G.; Van Hook, W. A. *Chem. Rev.* **1974**, *74*, 689−750.
(89) Berweger, C.; van Gunsteren, W.; Müller-Plathe, F. *Chem. Phys. Lett.* **1995**, *232*, 429−436.
(90) Eichinger, M.; Grubmüller, H.; Heller, H.; Tavan, P. *J. Comput. Chem.* **1997**, *18*, 1729−1749.
(91) Mathias, G.; Egwolf, B.; Nonella, M.; Tavan, P. *J. Chem. Phys.* **2003**, *118*, 10847−10860.
(92) Niedermeier, C.; Tavan, P. *J. Chem. Phys.* **1994**, *101*, 734−748.
(93) Niedermeier, C.; Tavan, P. *Mol. Simul.* **1996**, *17*, 57−66.
(94) Allen, M. P.; Tildesley, D. *Computer Simulations of Liquids*; Clarendon: Oxford, 1987.
(95) Kirkwood, J. G. *J. Chem. Phys.* **1934**, *2*, 351−361.
(96) Kaatze, U. *J. Chem. Eng. Data* **1989**, *34*, 371−374.
(97) Kräutler, V.; Van Gunsteren, W.; Hünenberger, P. *J. Chem. Phys.* **2001**, *22*, 501−508.
(98) Bussi, G.; Parrinello, M. *Comput. Phys. Commun.* **2008**, *179*, 26−29.
(99) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098−3100.
(100) Perdew, J. P. *Phys. Rev. B* **1986**, *33*, 8822−8824.
(101) Troullier, N.; Martins, J. *Phys. Rev. B* **1991**, *43*, 1993−2006.
(102) *CPMD V3.9*; Copyright IBM Corp 1990−2008, Copyright MPI für Festkörperforschung Stuttgart 1997−2001; see also www.cpmd.org.
(103) Pugliano, N.; Saykally, R. *Science* **1992**, *257*, 1937−1940.
(104) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(105) Lingenheil, M.; Denschlag, R.; Reichold, R.; Tavan, P. *J. Chem. Theory Comput.* **2008**, *4*, 1293−1306.
(106) Neumann, M. *Mol. Phys.* **1983**, *50*, 841−858.
(107) Smith, P. E.; van Gunsteren, W. F. *J. Chem. Phys.* **1994**, *100*, 3169−3174.
(108) Dünweg, B.; Kremer, K. *J. Chem. Phys.* **1993**, *99*, 6983−6997.
(109) Yeh, I.-C.; Hummer, G. *J. Phys. Chem. B* **2004**, *108*, 15873−15879.
(110) Motakabbir, K. A.; Berkowitz, M. *J. Phys. Chem.* **1990**, *94*, 8359−8362.

(111) Tironi, I. G.; Van Gunsteren, W. F. *Mol. Phys.* **1994**, *83*, 381−403.
(112) Weigend, F.; Köhn, A.; Hättig, C. *J. Chem. Phys.* **2002**, *116*, 3175−3183.
(113) Verhoeven, J.; Dymanus, A. *J. Chem. Phys.* **1970**, *52*, 3222−3233.
(114) Dyke, T. R.; Mack, K. M.; Muenter, J. S. *J. Chem. Phys.* **1977**, *66*, 498−510.
(115) Odutola, J. A.; Dyke, T. R. *J. Chem. Phys.* **1980**, *72*, 5062−5070.
(116) Sprik, M. *J. Chem. Phys.* **1991**, *95*, 6762−6769.
(117) Rapaport, D. C. *The art of molecular dynamics simulation*; Cambridge University Press: Cambridge, U.K., 2004.
(118) Krynicki, K.; Green, C.; Sawyer, D. *Faraday Discuss. Chem. Soc.* **1978**, *66*, 199−208.
(119) Harris, K.; Woolf, L. *J. Chem. Eng. Data* **2004**, *49*, 1064−1069.
(120) Tazi, S.; Botan, A.; Salanne, M.; Marry, V.; Turq, P.; Rotenberg, B. *J. Phys.: Condens. Matter* **2012**, *24*, 284117.
(121) Weast, R. C. *Handbook of Chemistry and Physics*; CRC: Boca Raton, 1983.

Supporting Information to: Polarizable water models from mixed computational and empirical optimization

Philipp Tröster, Konstantin Lorenzen, Magnus Schwörer, and Paul Tavan*

Lehrstuhl für Biomolekulare Optik, Fakultät für Physik,

Ludwig-Maximilians-Universität München,

Oettingenstr. 67, D-80538 München, Germany

E-mail: tavan@physik.uni-muenchen.de

*To whom correspondence should be addressed

S1 DFT/PMM Gaussian widths $\tilde{\sigma}_i$ for the TLνP models

As indicated in Section 3.1, the Gaussian widths $\tilde{\sigma}_i$ of PMM atoms, which are located close to a DFT atom in a DFT/PMM hybrid calculation, have to be carefully chosen. For this purpose we considered two possible hybrid models of the water dimer, in which either the H-bond donor (QP) or the acceptor (PQ) is described by DFT and its H-bonded partner by TLνP (see also Section S5 for the resulting PMM dimers and experimental data). The widths $\tilde{\sigma}_\mathrm{H}$ and $\tilde{\sigma}_\mathrm{O/L/M}$ of the partial charges at the positions of the PMM hydrogen and oxygen atoms (or massless charge sites L/M) were varied as to achieve an optimal symmetry of the QP and PQ dimer properties and a reasonable interpolation between experimental data and complete DFT descriptions (for details on the applied DFT and DFT/PMM methods see Section 3.1).

Table S8: DFT/PMM hybrid results on water dimers obtained at the optimal Gaussian widths $\tilde{\sigma}_\mathrm{H} = 0.24$ Å and $\tilde{\sigma}_\mathrm{O/L/M} = 0.46$ Å of the TLνP partial charges. Optimized were the geometries $\mathrm{G_d}$ [specified by the parameters $(d_\mathrm{OO}, \beta_1, \beta_2)$ explained in Figure S13.a], and binding energies E_pot. Also listed are experimental data[1–3] and results of a complete DFT description.

	d_OO^a	β_1^b	β_2^c	$E_\mathrm{pot}^\mathrm{d}$
exp[1–3]	2.98	51	123	−5.4
	±0.03	±10	±10	±0.6
DFT/TL5P QP	2.85	56	100	−5.1
TL5P/DFT PQ	2.87	57	112	−5.1
DFT/TL4P QP	2.83	55	115	−5.1
TL4P/DFT PQ	2.85	58	112	−5.3
DFT/TL3P QP	2.75	52	168	−6.5
TL3P/DFT PQ	2.83	52	113	−4.9
DFT QQ	2.98	56	123	−4.3

[a][Å] [b][deg] [c][deg] [d][kcal/mol]

Table S8 compares the equilibrium geometries and binding energies obtained for the DFT/PMM dimers in the PQ and QP settings with data from experiments and from pure DFT. With the exception of the DFT/TL3P dimer in the QP arrangement, which features a much too short O-O distance, a too negative binding energy, and a too large angle β_2, all other hybrid dimers actually do represent reasonable interpolations between the experimental and DFT data. Particularly the TL5P and

TL4P dimers exhibit a very nice PQ vs. QP symmetry. The final Gaussian widths $\tilde{\sigma}_{\mathrm{H}} = 0.24\,\text{Å}$ and $\tilde{\sigma}_{\mathrm{O/L/M}} = 0.46\,\text{Å}$ were quite clearly singled out by our parameter search and were, therefore, employed in all subsequent DFT/PMM calculations.

Note here that we also carried out an analogous optimization of Gaussian widths for DFT/MM water dimers, in which the MM fragment was described by TIP3P,[4] TIP4P/2005,[5] and TIP5P.[6] We found the somewhat larger widths $\tilde{\sigma}_{\mathrm{H}} = 0.37\,\text{Å}$ and $\tilde{\sigma}_{\mathrm{O/L/M}} = 0.62\,\text{Å}$, which are similar to the single width $\tilde{\sigma} = 0.57\,\text{Å}$ employed in previous studies.[7–9]

S2 Correlation of induced dipole moments

Figure 4 in the main text displays the root mean square deviations $\chi(\sigma, \nu)$ between the DFT/PMM and PMM induced dipole moments calculated with the PMM parameters $\mu^{\mathrm{l}}_{\mathrm{DFT}} = 1.79$ D and $\alpha^{\mathrm{g}}_{\mathrm{DFT}} = 1.58\ \text{Å}^3$. The minima of the curves identify the optimal widths $\sigma_\nu^{\mathrm{opt}}$ of the induced PMM dipoles.

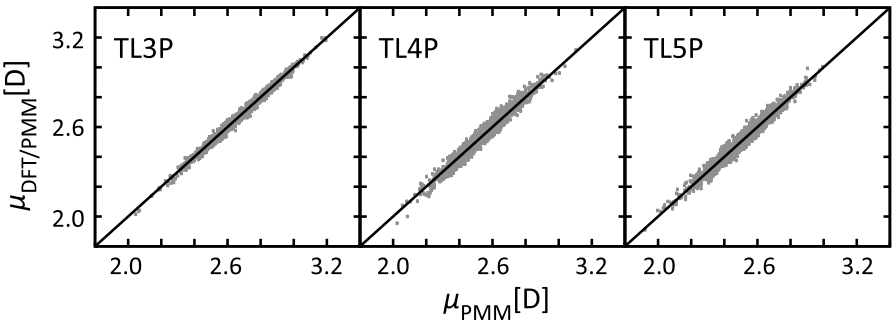


Figure S10: Correlation between the dipole $\mu^{\ell}_{\mathrm{DFT/PMM}}$ of the DFT fragment and the corresponding dipole $\mu_{\sigma,\nu} = \mu^{\ell}_{\mathrm{DFT}} + \mu^{\mathrm{i}}_{\sigma,\nu}$ of the PMM test molecule.

Figure S10 shows the correlations between the total DFT/PMM dipole moments and PMM dipole moments. Up to the constant offset of $\mu^{\mathrm{l}}_{\mathrm{DFT}}$ these correlations are identical to those of the associated induced dipole moments given by Eq. (2) and Eq. (3). The extraordinarily good correlations show that the dipole distributions calculated by DFT/PMM and PMM with the indicated parameters are essentially identical.

S3 Determination of $\mathbf{G}_e$ for $\nu = 4, 5$

The electrostatic geometries $\mathbf{G}_e$ of our polarizable ν-point models are characterized by the parameters (l_{OL}, φ_{LOL}) defined in Figure 2. For $\nu = 4$ the two lone-pair charges q_L of the five point model are degenerate, i.e. $q_M \equiv 2q_L$ for $\varphi_{LOL} = 360°$.

As explained in Sec. Section 3.1, the parameters (l_{OL}, φ_{LOL}) of $\mathbf{G}_e$ are optimized by fitting the static part $\Phi(\mathbf{r}\,|\,\rho_s^{stat})$ of the potential, which is generated by a DFT water molecule polarized by a surrounding ensemble $\mathscr{S}_\nu$ of TLνP models, on a spherical surface $\mathscr{P}$ to the potential $\Phi[\mathbf{r}\,|\,Q(l_{OL}, \varphi_{LOL})]$, which originates from the static partial charges $Q(l_{OL}, \varphi_{LOL})$ of a five-point model replacing the DFT fragment. Here the static partial charges $Q(l_{OL}, \varphi_{LOL})$ are subject to the constraint that (for $\langle \mathbf{E}_s \rangle_\sigma = 0$) the dipole moment of the five-point model is $\mu^{\ell}_{DFT} = 1.79\,D$.

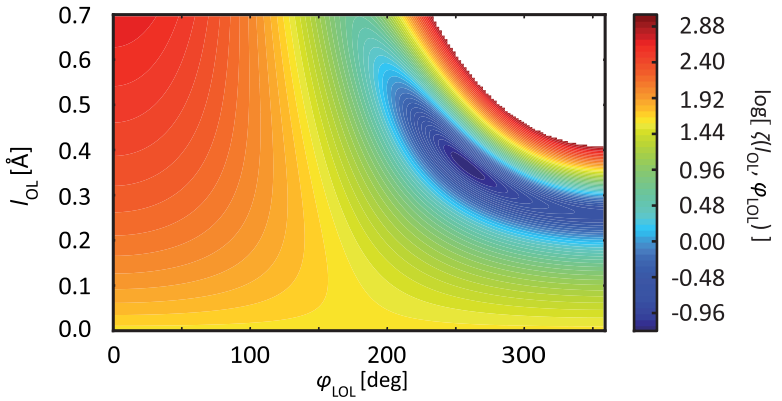


Figure S11: Landscape $\zeta(l_{OL}, \varphi_{LOL})$ of the deviation (7) between the static part $\Phi(\mathbf{r}\,|\,\rho_s^{stat})$ of the potential $\Phi(\mathbf{r}\,|\,\rho_s)$, which is generated by a DFT water molecule in a bulk environment of TL5P water models, and the potential $\Phi[\mathbf{r}\,|\,Q(l_{OL}, \varphi_{LOL})]$, which is caused by the four static partial charges Q of a five-point model with the electrostatic geometry $\mathbf{G}_e = (l_{OL}, \varphi_{LOL})$ replacing the DFT fragment, on a surrounding spherical surface $\mathscr{P}$ (see Secs. Section 2.3 and Section 3.1).

Figure S11 shows for TL5P the landscape of the root mean square deviations $\zeta(l_{OL}, \varphi_{LOL})$ defined by Eq. (7). The landscape is seen to be a convex function of its arguments featuring a single minimum at $l_{OL} = 0.3231$ Å and $\varphi_{LOL} = 268.2°$. The associated geometrical arrangement of the four static partial charges is depicted in Figure 5.

In Figure S11 the one-dimensional cut through $\zeta(l_{OL}, \varphi_{LOL})$ at $\varphi_{LOL} = 360°$ determines the optimal location $l_{OM} \equiv l_{OL}$ of the degenerate lone-pair charges $q_M = 2q_L$ on the HOH bisec-

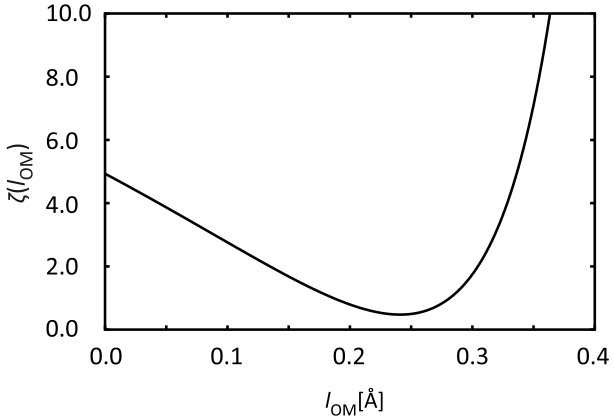


Figure S12: One-dimensional landscape $\zeta(l_{\mathrm{OM}})$ of the deviation (7) between the static part $\Phi(\mathbf{r}\,|\,\rho_s^{\mathrm{stat}})$ of the potential $\Phi(\mathbf{r}\,|\,\rho_s)$, which is generated by a DFT water molecule in a bulk environment of TL4P water models, and the potential $\Phi[\mathbf{r}\,|\,Q(l_{\mathrm{OM}})]$, which is caused by the three static partial charges Q of a four-point model with a given parameter l_{OM} replacing the DFT fragment, on a surrounding spherical surface $\mathscr{P}$ (see Secs. Section 2.3 and Section 3.1).

trix. This cut is plotted in Figure S12, which reveals a well-defined minimum of $\zeta(l_{\mathrm{OM}}, 360^\circ)$ at $l_{\mathrm{OM}} = 0.242$ Å. Thus, the charge q_{M} of TL4P is located by 0.017 Å more distant from $\mathbf{r}_{\mathrm{O}}$ than the projection of the two lone pair charges q_{L} of TL5P on the molecular plane, which hits that plane at the distance of 0.225 Å.

S4 The parameters of the TLνP models

To enable a quick overview, the TLνP parameters are summarized in Table S9.

Table S9: Parameters of the TLνP models.

parameter	unit	TL3P	TL4P	TL5P
A_1	$\mathrm{Å}^{12}\,\mathrm{kcal/mol}$	302100	84120	64300
A_2	$\mathrm{Å}^{-1}$	4.17	3.55	3.40
B	$\mathrm{Å}^{6}\,\mathrm{kcal/mol}$	486	992	1180
σ	Å	0.780	0.842	0.852
α	$\mathrm{Å}^3$	1.47	1.47	1.47
μ^ℓ	D	2.091	1.855	1.855
q_{H}	e	0.3717	0.5577	0.5328
l_{OL}	Å	-	0.2419	0.3231
φ_{LOL}	deg	-	-	268.2

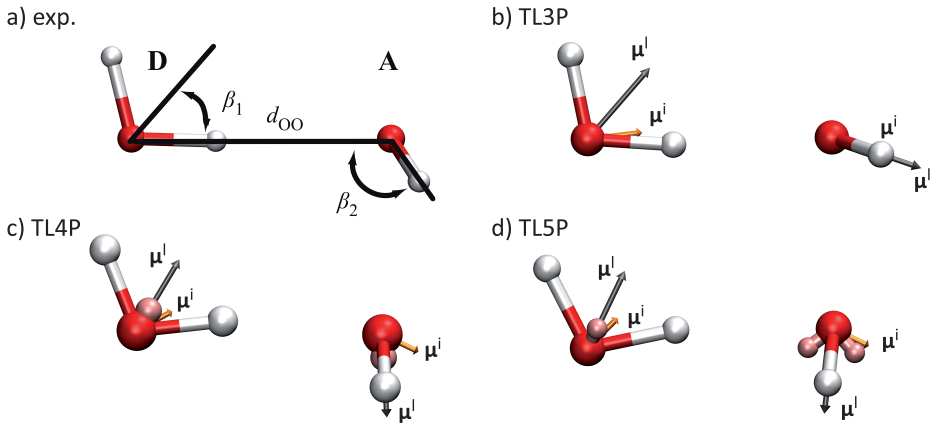


Figure S13: Geometry $\mathbf{G}_\mathrm{d}$ of the water dimer as determined by (a) experimental data[3] and (b-d) the TLνP models, $\nu = 3,4,5$. Also drawn are the static dipole moments μ^{ℓ} and their induced counterparts μ^{i} of the TLνP water models making up the PMM dimers. Drawing a) highlights the hydrogen bond donor (**D**) and acceptor (**A**), respectively.

S5 Properties of TLνP water dimers

The geometry $\mathbf{G}_\mathrm{d}$ of the water dimer, its binding energy E_pot and dipole moment μ_Dimer are experimentally quite well known.[1–3] The corresponding data and their associated uncertainties σ_exp are listed in Table S10. $\mathbf{G}_\mathrm{d}$ is characterized by the oxygen-oxygen distance d_OO and by two angles (β_1, β_2). Figure S13a illustrates these parameters for the experimental model. Table S10 additionally provides experimental data on the binding energy[1] E_pot and the total dipole moment[2] μ_d of the isolated dimer. The experimental data on $\mathbf{G}_\mathrm{d}$, E_pot, and μ_d are compared with descriptions obtained with the TLνP and four other[10–13] PMM models.

Looking first in Table S10 at the inclination β_1 of the hydrogen bond donor **D** with respect to the line connecting the two oxygens (cf. Figure S13a) one recognizes that almost all PMM values reproduce the experimental angle of 51° within the 10° error bound σ_exp. Solely SWM4-DP clearly overestimates the inclination β_1 by at least 10°.

β_2 measures the orientation of the bisectrix of the acceptor **A** with respect to the O-O connection. Here, the most recent polarizable six-point DO model SWM6[13] (which actually employs seven force points due to the use of a Drude charge) and GCPM[13] (which employs five force points

and three Gaussian charges) perfectly match the experimental orientation, while the four-point DO models COS/G2[11] and SWM4-DP[12] yield slight underestimates. Much larger ($\geq 20°$) are the underestimates predicted by TL4P and TL5P, which assign to **A** an almost perpendicular orientation. This underestimate of β_2 is visible in Figure S13, if one compares c) and d) with a). TL3P, on the other hand, overestimates β_2 and, correspondingly, **A** features an almost collinear orientation with respect to the O-O connection. Such large angles β_2 are a hallmark of three point models, as one can see from a glance into the corresponding literature.[11]

Table S10: Geometries G_d [specified by the parameters $(d_{OO}, \beta_1, \beta_2)$ explained in Figure S13.a], binding energies E_{pot}, and dipole moments μ_d of the water dimer as given by experimental data,[1–3] the TLνP and various other[10–12] PMM models.

	d_{OO}^{a}	β_1^{b}	β_2^{c}	E_{pot}^{d}	μ_d^{e}
exp[1–3]	2.98	51	123	−5.4	2.6
σ_{exp}	±0.03	±10	±10	±0.6	±0.05
TL3P	2.78	48	161	−5.42	6.07
TL4P	2.78	59	102	−5.17	3.36
TL5P	2.71	59	85	−5.50	2.67
GCPM[10]	2.88	56	125	-4.95	-
COS/G2[11]	2.81	56	106	-5.00	2.08
SWM4-DP[12]	2.82	70	106	-5.18	2.09
SWM6[13]	2.79	56	123	-5.27	2.48

[a][Å] [b][deg] [c][deg] [d][kcal/mol] [e][D]

The almost collinear orientation β_2 of **A** in three-point PMM models leads to a huge overestimate ($> 200\%$) of the total dipole moment μ_d, which is caused by the comparably small angle between the mutually inducing dipole moments of **D** and **A** (cf. Figure S13b). Such a small angle is absent in all other polarizable models, because of the prevalent underestimate of β_2 (cf. Figs. S13b-c for examples). Therefore, these model tend to underestimate μ_d with the exception of TL4P and TL5P, which overestimate or almost reproduce $\mu_{d,exp}$, respectively.

As we have seen, the predictions on β_2 and μ_d show some correlation. A similar correlation exists between the predictions for d_{OO} and E_{pot}. According to Table S10 all models significantly underestimate the experimental distance $d_{OO} = 2.98$ Å by 3-10%. GCPM yields the smallest error

due to its soft Gaussian distributions of the static partial charges. Correspondingly it predicts a rather weakly bound dimer. TL5P in contrast, shows the strongest underestimate of d_{OO} and, therefore, predicts the most strongly bound dimer. Note, however, that all predictions of E_{pot} are within the range of experimental uncertainty.

As a final remark we would like to add that the TL4P and TL5P dimer geometries move closer to the experimental data if the experimental gas phase geometry $\mathbf{G}_m^g$ is chosen instead of $\mathbf{G}_m^\ell$ (cf. Table 1) while retaining the zero-field dipole moment μ_{exp}^g.

S6 Properties of TLνP water trimers

As a further test for the gas phase properties of the TLνP models and, hence, for the transferability of the 300 K bulk phase models to other environments, we have calculated the geometry of the cyclic trimer. The results are compared with the distances of the oxygen atoms d_{OO} measured by Pugliano et. al.[14] by far infrared absorption spectroscopy.

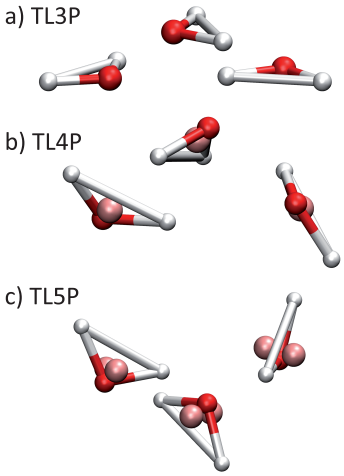


Figure S14: Geometries of the cyclic TLνP trimers .

Figure S14 shows the minimum energy geometries of the cyclic TLνP trimers. TL3P model is seen to have an almost planar geometry, as one might expect for a three-point model. The TL4P and TL5P models, in contrast, exhibit non-planar structures, which agree much better with the structural model derived by Pugliano et. al.[14] from their data and depicted in Fig. 1 of their paper.

Table S11: Distance of the oxygen atoms d_{OO} and binding energy E^{b}_{pot} of the cyclic water trimers of the TLνP models.

	d^{a}_{OO}	E^{b}_{pot}
exp[14]	2.96	-
TL3P	2.76	−15.39
TL4P	2.80	−14.75
TL5P	2.78	−15.40

[a][Å] [b][kcal/mol]

As far as the distances d_{OO} of the oxygen atoms are concerned, Table S11 reveals that all TLνP models underestimate the experimental value, just like in the case of the dimer (cf. Table S10). However, these underestimates are smaller in the trimer than in the dimer (for TL5P, e.g. it shrinks from 9.1% to 6.1%). Thus with increasing cluster size our bulk phase models show an improved performance as one might expect for such models. However, when using cluster data one must bear in mind that they refer to the temperature $T = 0\,\mathrm{K}$ with enthalpically driven most compact structures, whereas entropy guarantees at $T_0 = 300\,\mathrm{K}$ strongly loosened molecular configurations.

For the additionally listed binding energies we unfortunately found no experimental value.

S7 Temperature density profile of the TLνP water models

The temperature density profile $n(T)$ and the temperature of maximum density (T_{MD}), which have first been calculated for the TIP5P model,[6] are important tests for the transferability of water models to other densities and temperatures than the usual target temperature $T_0 = 300\,\mathrm{K}$ of the parameterization. Interestingly, polarizable water models (COS/B2,[15] COS/G2, COS/G3,[11] SWM4-DP[16] and SWM4-NDP[12]) generally could not reproduce[17] the experimental[18] profile $n_{exp}(T)$. For $T < T_0$ these models overestimate $n_{exp}(T)$, whereas they underestimate it for $T > T_0$. The non-polarizable TIP4P/2005 model,[5] in contrast, performs very well on $n(T)$.

In an attempt to explain the poor performance of polarizable models on $n(T)$ Kiss and Baranyai[17] argued that the polarizability should decrease and the van der Waals repulsion should increase with

decreasing temperature and correspondingly increasing density. A correspondingly designed polarizable water model called BKd3 could then actually reproduce $n_{\mathrm{exp}}(T)$. For our TLνP models we plan to pursue this interesting suggestion further.

Therefore, in the current context of a parameterization targeted to T_0, we did not spend much effort to characterize the expectedly poor performance of the TLνP models on $n(T)$. For a first estimate we solely carried out temporally restricted $N_{\mathrm{s}}pT$ trajectories at the temperatures $T \in [230-310]$ K in steps of 20 K, whose durations of 2 ns do not suffice for statistically reliable sampling at temperatures of 270 K and below. For an improved characterization one should resort to extended ensemble techniques such as replica exchange.[19,20] Nevertheless, in Figure S15 we

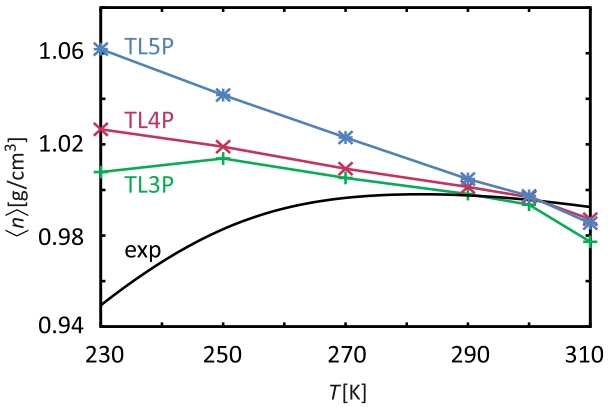


Figure S15: Temperature density profile of the TLνP water models. The more structured the RDF is, the steeper the $n(T)$ curve is. TL4P and TL5P show an almost linear behaviour and especially TL5P has a way too high density at low temperatures. TL3P shows a weaker $n(T)$ dependence.

document for the TLνP models the results of our preliminary $N_{\mathrm{s}}pT$ simulations, in which the coupling constants of the Bussi thermostat and the Berendsen barostat were chosen 1 ps and 10 ps, respectively. The depicted density profiles show the deficiencies, which were to be expected for such polarizable models (see above). Particularly, for TL4P and TL5P no T_{MD} shows up in the examined temperature range. However, error estimates such as the so-called block-averaging[21] clearly demonstrated that the data points for $n(T)$ were far from convergence particularly at $T \leq 270$ K. Thus, the absence of a T_{MD} in the depicted temperature range cannot be ascertained for TL4P and TL5P. Furthermore, the value $T_{\mathrm{MD}} = 250$ K suggested for the TL3P model does not

represent a reliable result. Note that the slopes $-dn(T)/dT$ at T_0 =300 K are essentially the overestimated thermal expansion coefficients α_p discussed in Section 4.2 and listed in Table 7.

S8 Experimental RDFs

During the evaluation of the parameterized TLνP models Skinner et al.[22] published most recently an O-O RDF $g_{OO}(r)$, which was calculated from X-ray diffraction data. This RDF places the first maximum to 2.80±0.01 Å, i.e. to a distance which is by 1.4 % larger than Soper's[23] value of 2.76 Å, which we used for the empirical optimization of the Buckingham parameter A_1 [cf. Eq. (6) and Section 3.1] of the TLνP models. Furthermore, the most recent RDF also assigns a slightly larger value of 2.57±0.05 to the height of the first peak.

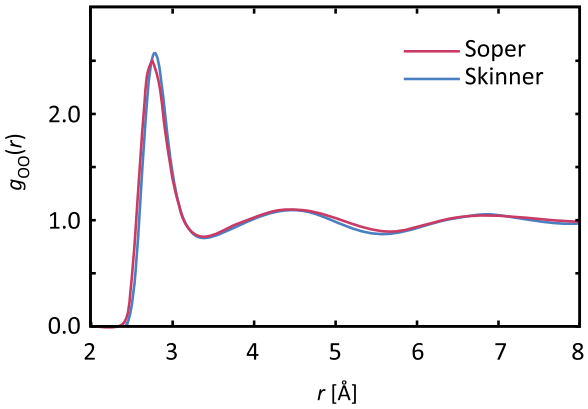


Figure S16: Comparison of the RDFs $g_{OO}(r)$ of Skinner[22] (blue) with Soper's[23] experimental reference used by us (red).

Beyond the slight changes of the position and height of the first peak Skinner's RDF reveals only one significant difference with respect to Soper's result, i.e. the position of the second minimum is shifted by about 0.1 Å toward smaller distances. A comparison with Figure 7 shows that the TL4P and TL5P models both predict similar shifts of this minimum into the same direction. In future refinements of the TLνP models or in the development of extensions (e.g. $\nu = 6$) it will be interesting to check to what extent the use of Skinner's O-O distance of 2.80 Å in the optimization of A_1 modifies the properties of the various models.

S9 DFT/PMM and PMM dipole distributions

Figure S17 shows the distributions $p(\mu)$ of the DFT dipole moment (solid lines) calculated by DFT/PMM from extended TLνP ensembles $\mathscr{S}_\nu$ covering 5000 statistically independent solvation structures of the DFT fragment. These structures were generated from the $N_\mathrm{m}V_\mathrm{m}T_0$ trajectories by taking every 5 ps a snapshot and selecting 25 water molecules from each snapshot. These DFT/PMM distributions are compared with their PMM counterparts (dotted lines), which were taken from the same trajectories. In both cases the $p(\mu)$ are represented as normal distributions. However, histograms are additionally drawn for the DFT/PMM data to illustrate the quality of the Gaussian models.

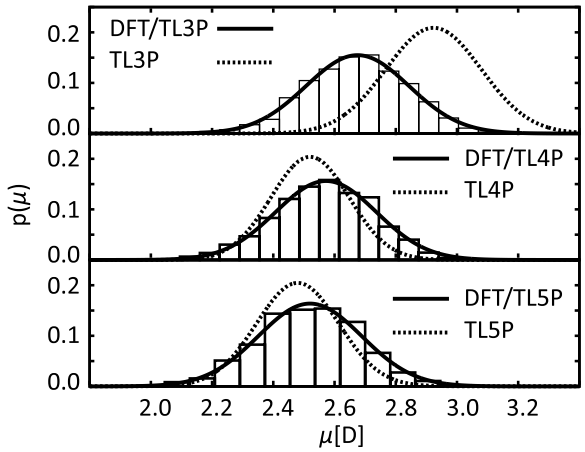


Figure S17: The distributions $p(\mu)$ of the DFT dipole moment, which were calculated by DFT/PMM from extended TLνP ensembles $\mathscr{S}_\nu$, are represented as histograms and normal distributions (solid lines). They are compared with normal distributions (dotted lines) representing the PMM data for the same ensembles.

For all TLνP models the DFT/PMM dipole distributions $p(\mu)$ in Figure S17 have almost identical widths as is apparent from the almost identical heights. But these widths are by 25% larger than those of the three PMM distributions, whose widths are similarly small among each other. The larger widths of the DFT/PMM dipole distributions follow from the fact that MT/BP overestimates the polarizability of a water molecule by 7.5%. Correspondingly, the fluctuations of the local polarizing fields are exaggerated by the MT/BP description. On the other hand, the thus expected 7.5% overestimate of the average induced dipole moment $\langle\mu^\mathrm{i}\rangle$ by DFT/PMM is partially compen-

sated (for TL4P and TL5P) by the smaller value of the static dipole moment ($\mu^{\ell}_{\mathrm{DFT}} = 1.79\,\mathrm{D} < \mu^{\ell} = 1.855\,\mathrm{D}$). Hence, the suggested[9] use of $\mu^{\ell} = 1.855\,\mathrm{D}$ and $\alpha^{\ell} = 1.47\,\text{Å}^3$ in TL4P and TL5P in combination with the self-consistent DFT/PMM optimization of the electrostatic geometry $\mathbf{G}_{\mathrm{e}}$ actually yield PMM models, which comply reasonably well with the DFT/PMM descriptions. For TL3P, in contrast, the empirical optimization of μ^{ℓ} ($= 2.091\,\mathrm{D}$) leads to much larger dipole moments ($\langle\mu\rangle = 2.925\,\mathrm{D}$) than suggested by DFT/TL3P ($\langle\mu\rangle = 2.675\,\mathrm{D}$).

S10 Tables of observables targeted by optimization

For completeness we document in Table S12 the results of the $N_iV_iT_0$, $i \in \{\mathrm{s},\mathrm{m},\mathrm{l}\}$, simulations (cf. Table 2) on the potential energy per molecule and the average pressure.

Table S12: Average potential energies per molecule and average pressures at T_0 and n_{exp}.

observable	TL3P	TL4P	TL5P
$\langle E_{\mathrm{pot}}\rangle(N_{\mathrm{s}}V_{\mathrm{s}})/(\mathrm{kcal/mol})$	-9.889 ± 0.003	-9.918 ± 0.003	-9.921 ± 0.004
$\langle E_{\mathrm{pot}}\rangle(N_{\mathrm{m}}V_{\mathrm{m}})/(\mathrm{kcal/mol})$	-9.916 ± 0.004	-9.924 ± 0.004	-9.924 ± 0.004
$\langle E_{\mathrm{pot}}\rangle(N_{\mathrm{l}}V_{\mathrm{l}})/(\mathrm{kcal/mol})$	-9.921 ± 0.004	-9.932 ± 0.003	-9.938 ± 0.006
$\langle p\rangle(N_{\mathrm{s}}V_{\mathrm{s}})/(\mathrm{atm})$	-50 ± 6	25 ± 4	18 ± 6
$\langle p\rangle(N_{\mathrm{m}}V_{\mathrm{m}})/(\mathrm{atm})$	-32 ± 4	18 ± 5	-34 ± 5
$\langle p\rangle(N_{\mathrm{l}}V_{\mathrm{l}})/(\mathrm{atm})$	-46 ± 3	-25 ± 4	-52 ± 4

Furthermore Table S13 lists the average densities observed in the $N_ip_0T_0$ simulations (cf. Table 3) on the three systems. The statistical errors of the listed mean values were estimated using the block-averaging method described in Chap. 4 of Ref. 21.

Table S13: Average densities $\langle n\rangle$ at p_0 and T_0 from NpT simulations.

model	TL3P	TL4P	TL5P
$\langle n\rangle(N_{\mathrm{s}})/(\mathrm{g/cm^3})$	0.996 ± 0.004	0.994 ± 0.001	0.997 ± 0.003
$\langle n\rangle(N_{\mathrm{m}})/(\mathrm{g/cm^3})$	0.997 ± 0.002	0.997 ± 0.001	0.998 ± 0.003
$\langle n\rangle(N_{\mathrm{l}})/(\mathrm{g/cm^3})$	0.998 ± 0.002	0.998 ± 0.002	0.998 ± 0.001

The potential energies per molecule and the average pressure are almost independent of the system size, if one disregards the very small decreases of the two observables with increasing N.

Correspondingly the average densities $\langle n \rangle$ obtained by the Np_0T_0 simulations also increase only slightly with N. This increase is largest for TL3P and smallest for TL5P, in which case it cannot be ascertained with statistical significance.

S11 Electrostatic geometry of a six-point PMM model

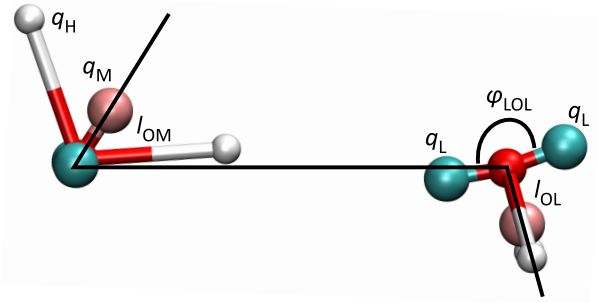


Figure S18: The electrostatic geometry $\mathbf{G}_e$ of TL6P$_{ini}$ and the associated dimer structure. $\mathbf{G}_e$ is given by the parameters $l_{OM} = 0.40$ Å, $l_{OL} = 0.46$ Å, $\varphi_{LOL} = 175°$, $q_M = -0.57$ e, $q_L = -0.23$ e, $q_H = 0.52$ e, and the experimental liquid phase molecular geometry $\mathbf{G}_m^\ell$ specified in Table 1. TL6P$_{ini}$ has the static gas phase dipole moment of 1.855 D and the experimental gas phase polarizability of 1.47 Å^3.

Figure S18 shows an initial guess for the electrostatic geometry $\mathbf{G}_e$ of a polarizable six point model (TL6P$_{ini}$), whose DFT/PMM computation is described in Section 4.1. Adopting the Buckingham parameters and the width of the Gaussian induced dipole from TL4P we have additionally calculated the equilibrium geometry $\mathbf{G}_d = (d_{OO}, \beta_1, \beta_2)$ (cf. Figure S13a) of a corresponding dimer, which is also depicted in the figure. We found the values $d_{OO} = 2.81$ Å, $\beta_1 = 58°$, and $\beta_2 = 106°$ associated with the binding energy $E_{pot}^d = -5.76$ kcal/mol. The observables are a little closer to the experimental findings than for TL4P or TL5P and may even move significantly closer to their experimental counterparts as soon as the Buckingham potentials are optimized through weak-coupling MD simulations.

References

(1) Verhoeven, J.; Dymanus, A. *J. Chem. Phys.* **1970**, *52*, 3222–3233.

(2) Dyke, T. R.; Mack, K. M.; Muenter, J. S. *J. Chem. Phys.* **1977**, *66*, 498–510.

(3) Odutola, J. A.; Dyke, T. R. *J. Chem. Phys.* **1980**, *72*, 5062–5070.

(4) Jorgensen, W.; Chandrasekhar, J.; Madura, J.; Impey, R.; Klein, M. *J. Phys. Chem.* **1983**, *79*, 926–935.

(5) Abascal, J.; Vega, C. *J. Chem. Phys.* **2005**, *123*, 234505.

(6) Mahoney, M.; Jorgensen, W. *J. Chem. Phys.* **2000**, *112*, 8910–8922.

(7) Eichinger, M.; Tavan, P.; Hutter, J.; Parrinello, M. *J. Chem. Phys.* **1999**, *110*, 10452–10467.

(8) Schropp, B.; Tavan, P. *J. Phys. Chem. B* **2008**, *112*, 6233–6240.

(9) Schropp, B.; Tavan, P. *J. Phys. Chem. B* **2010**, *114*, 2051–2057.

(10) Paricaud, P.; Předota, M.; Chialvo, A.; Cummings, P. *J. Chem. Phys.* **2005**, *122*, 244511.

(11) Yu, H.; van Gunsteren, W. *J. Chem. Phys.* **2004**, *121*, 9549–9564.

(12) Lamoureux, G.; Harder, E.; Vorobyov, I.; Roux, B.; MacKerell, A. *Chem. Phys. Lett.* **2006**, *418*, 245–249.

(13) Yu, W.; Lopes, P. E. M.; Roux, B.; A. D. MacKerell, J. *J. Chem. Phys.* **2013**, *138*, 034508.

(14) Pugliano, N.; Saykally, R. *Science* **1992**, *257*, 1937–1940.

(15) Yu, H.; Hansson, T.; Van Gunsteren, W. *J. Chem. Phys.* **2003**, *118*, 221–234.

(16) Lamoureux, G.; MacKerell Jr, A.; Roux, B. *J. Chem. Phys.* **2003**, *119*, 5185–5197.

(17) Kiss, P. T.; Baranyai, A. *J. Chem. Phys.* **2012**, *137*, 084506.

(18) Kell, G. *J. Chem. Eng. Data* **1967**, *12*, 66–69.

(19) Hukushima, K.; Nemoto, K. *J. Phys. Soc. Jpn.* **1996**, *65*, 1604–1608.

(20) Hansmann, U. H. *Chem. Phys. Lett.* **1997**, *281*, 140–150.

(21) Rapaport, D. C. *The art of molecular dynamics simulation*; Cambridge university press: Cambridge, UK, 2004.

(22) Skinner, L. B.; Huang, C.; Schlesinger, D.; Pettersson, L. G.; Nilsson, A.; Benmore, C. J. *J. Chem. Phys.* **2013**, *138*, 074506.

(23) Soper, A. *Chem. Phys.* **2000**, *258*, 121–137.

2.2 Polarisierbare Sechspunktmodelle

Die nachfolgende Publikation[2]

> „Polarizable Six-Point Water Models from Computational and Empirical Optimization“, Philipp Tröster, Konstantin Lorenzen, and Paul Tavan, *J. Phys. Chem. B*, **118**, 1589-1602, (2014)

wurde von mir zusammen mit Konstantin Lorenzen und Paul Tavan verfasst. Darin wird gezeigt dass ein PMM 6-Punktmodell eine signifikante Verbesserung gegenüber 4- und 5-Punktmodellen darstellt. Da für die Optimierung der Modellparameter die gleiche DFT/PMM Optimierungsmethode verwendet wird, ist die Vergleichbarkeit der verschieden komplexen Modelle garantiert.

[2]Reproduced with permission from the Journal of Physical Chemistry, **118**, 1589-1602, 2014.

THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

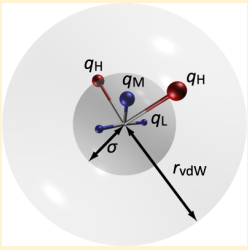
Polarizable Six-Point Water Models from Computational and Empirical Optimization

Philipp Tröster, Konstantin Lorenzen, and Paul Tavan*

Lehrstuhl für Biomolekulare Optik, Fakultät für Physik, Ludwig-Maximilians-Universität München, Oettingenstrasse 67, D-80538 Müunchen, Germany

Supporting Information

ABSTRACT: Tröster et al. (*J. Phys. Chem B* **2013**, *117*, 9486−9500) recently suggested a mixed computational and empirical approach to the optimization of polarizable molecular mechanics (PMM) water models. In the empirical part the parameters of Buckingham potentials are optimized by PMM molecular dynamics (MD) simulations. The computational part applies hybrid calculations, which combine the quantum mechanical description of a H_2O molecule by density functional theory (DFT) with a PMM model of its liquid phase environment generated by MD. While the static dipole moments and polarizabilities of the PMM water models are fixed at the experimental gas phase values, the DFT/PMM calculations are employed to optimize the remaining electrostatic properties. These properties cover the width of a Gaussian inducible dipole positioned at the oxygen and the locations of massless negative charge points within the molecule (the positive charges are attached to the hydrogens). The authors considered the cases of one and two negative charges rendering the PMM four- and five-point models TL4P and TL5P. Here we extend their approach to three negative charges, thus suggesting the PMM six-point model TL6P. As compared to the predecessors and to other PMM models, which also exhibit partial charges at fixed positions, TL6P turned out to predict all studied properties of liquid water at p_0 = 1 bar and T_0 = 300 K with a remarkable accuracy. These properties cover, for instance, the diffusion constant, viscosity, isobaric heat capacity, isothermal compressibility, dielectric constant, density, and the isobaric thermal expansion coefficient. This success concurrently provides a microscopic physical explanation of corresponding shortcomings of previous models. It uniquely assigns the failures of previous models to substantial inaccuracies in the description of the higher electrostatic multipole moments of liquid phase water molecules. Resulting favorable properties concerning the transferability to other temperatures and conditions like the melting of ice are also discussed.



1. INTRODUCTION

Liquid water features many unusual properties and is the biologically most important solvent, because life originates from water.[1] Driven by the hope to understand, in most simple terms, the microscopic causes of the bulk liquid's very special properties, the development of simplified molecular energy functions, which can reproduce these properties as closely as possible in molecular dynamics (MD) simulations, is a long lasting and ongoing international effort.[2,3]

Accurate theoretical descriptions of molecular properties require quantum mechanical methods like density functional theory (DFT) or other approaches of quantum chemistry. Simulation systems addressing the bulk liquid should comprise several thousand water molecules[4−8] and corresponding MD simulations should cover several nanoseconds[9,10] for well-sampled computations of properties such as the dielectric constant. Here, DFT and other methods of quantum chemistry are excluded for reasons of computational manageability[11,12] and one must resort to less accurate but computationally much more efficient and preferentially polarizable[3,13] molecular mechanics (PMM) models.

The sketched dilemma between accuracy and computational efficiency can be partially circumvented, for instance, by the recent hybrid method of Schwörer et al.,[14] because it combines a PMM-MD treatment of almost all molecules in a liquid phase simulation system with a DFT treatment of a few. In a first application this novel technology has been exploited for the mixed computational and empirical optimization of three PMM model potentials featuring ν = 3, 4, and 5 points of force action.[8]

The construction of these so-called[15] TLνP model potentials was driven by the conviction that quantitatively accurate descriptions of water in the liquid phase and in inhomogeneous condensed phase environments such as protein−solvent mixtures can only be achieved, if the physics of the individual molecules is represented as correctly as possible. For reasons of computational manageability, corresponding models should also be as simple as possible, of course. The question then is how complex a model must be made until it can start to cover the essential physics of a water molecule in a condensed phase environment with sufficient accuracy.

The TLνP models describe the electrostatic signature of the H_2O molecule by three (ν = 3, 4) or four (ν = 5) static partial charges, where two positive charges $q_H > 0$ are located at the hydrogens and the remaining negative charges either at the oxygen or in its vicinity, such that the static dipole moment has

Received: September 16, 2013
Revised: January 16, 2014
Published: January 17, 2014

   dx.doi.org/10.1021/jp4125765 | *J. Phys. Chem. B* 2014, 118, 1589−1602

the experimental gas phase value[16] μ^{g}_{exp} = 1.855 D for $v \geq 4$ and is 2.091 D for $v = 3$. They account for the polarizability α by a Gaussian inducible dipole $\boldsymbol{\mu}^{i}(r)$ of width σ, apply the experimental gas phase value[17] α^{g}_{exp} = 1.47 Å^3, and model the van der Waals interactions by a Buckingham potential[18]

$$E_B(r) = A_1 \exp(-rA_2) - B/r^6 \quad (1)$$

where r is the distance from the oxygen atom and (A_1, A_2, B) are positive parameters. Thus, $E_B(r)$ and the Gaussian inducible dipole $\boldsymbol{\mu}^{i}(r)$ are both centered at the oxygen. The three atomic masses of H_2O are arranged in the experimental liquid phase molecular geometry $\mathbf{G}^{l}_{m}$, which is given by[19,20] the bond angle φ_{HOH} = 105.3° and the bond length l_{OH} = 0.968 Å.

Just like other PMM three-point models,[21−25] the resulting TL3P potential did not render reasonable descriptions[8] of liquid water at the standard temperature T_0 = 300 K and experimental density[26] $n_{exp}(T_0,p_0)$ = 0.9965 g/cm^3 assumed at the standard pressure p_0 = 1 bar. For instance, TL3P and related three-point PMM models overestimate the dielectric constant[27] $\epsilon(T_0,p_0)$ = 78 by about a factor 2, which is why we will exclude these models from our further discussions.

The TL4P and TL5P models, however, not only reproduced the few properties targeted by the empirical optimization but also rendered excellent predictions for a number of other bulk phase properties[8] at T_0 and $n_{exp}(T_0,p_0)$, which include ϵ, the diffusion coefficient D, the viscosity η, the heat capacity C_p at constant pressure, and the isothermal compressibility κ_T. A notable exception was the thermal expansion coefficient $\alpha_p(T_0,p_0) = -(\partial \ln n/\partial T)_p$, which is the negative temperature derivative of the density at constant pressure, has the experimental value[26] of 2.75 × 10^{-4} K^{-1}, and was overestimated by at least a factor of 2.

Concerning α_p, the performance of the TL4P and TL5P models is comparable to that of other four- and five-point PMM models,[28−34] which were optimized by exclusively empirical parameter searches. As a possible remedy Kiss and Baranyai[34] suggested to choose also the Lennard-Jones[35] parameters A and B as polarizable in a four-point "one-charge on spring" PMM model (called BKd3) featuring Gaussian partial charges. In a second attempt[36] they combined a nonpolarizable Buckingham potential with a four-point "three Gaussian charges on springs" PMM model (called BK3) representing a polarizability distributed on the three charges.

In both cases the parameter sets are larger than the eight numbers required to specify conventional PMM four-point models, because they additionally contain two response parameters and two Gaussian widths. Furthermore, the use of Gaussian partial charges implies an enhanced computational effort. After empirical optimizations, which included the dielectric constant and the temperature density function of liquid water (BKd3) or the densities of the liquid and of ice (BK3) as targets, both models managed to reproduce α_p at T_0 much better. BKd3 missed α_p by[37] only −11% and BK3 by +17%.

The BKd3 and BK3 deviations for α_p are larger than the corresponding 3% underestimate provided by the recent partially polarizable[38] and very complex model iAMOEBA,[39] whose 19 independent parameters were optimized using large numbers of experimental data on liquid water [including $\alpha_p(T_0,p_0)$ and $\epsilon(T_0,p_0)$] and of quantum chemical results on small icy clusters as targets. These deviations are similar to the 9% overestimate[37] characterizing the nonpolarizable[38] TIP4P/2005 model,[40] whose parameters (just like those of BKd3 and iAMOEBA) were optimized using $\alpha_p(T_0,p_0)$ as one of the targets. On the other hand, TIP4P/2005 underestimates the dielectric constant[27] $\epsilon(T_0,p_0)$ by about 25% whereas the two PMM models BKd3 and BK3 as well as iAMOEBA yield close (±3%) matches also in this respect. Note that the nonpolarizable TIP4Q model,[41] which features four partial point charges and has been parametrized with $\epsilon(T_0,p_0)$ and with the temperature-density profile $n(T,p_0)$ as optimization targets, manages to reproduce $\alpha_p(T_0,p_0)$ quite accurately and overestimates $\epsilon(T_0,p_0)$ by only about 3%.

Because the quoted parametric PMM approaches[34,36] are based on different physical pictures with BKd3 emphasizing a field-dependence of the van der Waals interactions and with BK3 putting forward three distributed polarizabilities of the "Gaussian charges on springs"-type, the microscopic physical reason for the concurrent reproduction of $\epsilon(T_0,p_0)$ and $\alpha_p(T_0,p_0)$ remains unclear. Similarly, the complexity of iAMOEBA, which includes a molecular flexibility, a distributed but incomplete polarizability,[38] and atomic static multipoles as described by 19 adjustable parameters, prevents any identification of those microscopic physical properties, which lead to reasonable liquid phase descriptions in some respects and to suboptimal ones in others. The quadrupole moments of an iAMOEBA molecular model, for instance, deviate by an average of 12% and the polarizability even by 24% from the respective experimental values,[39] demonstrating that iAMOEBA is an effective[38] but not a physical model for water. Similar considerations apply to other nonpolarizable and effective model potentials like TIP4P/2005[40] or TIP4Q[41] (despite their much smaller complexity).

In contrast, the mixed computational and empirical approach toward the optimization of water models suggested in ref 8 opens the chance to identify the microscopic physical causes for the joint match of $\epsilon(T_0,p_0)$ and $\alpha_p(T_0,p_0)$, because it is based on experimentally or theoretically well-established physical properties of individual water molecules, because it clearly distinguishes between electrostatic and van der Waals interactions, and because it applies separate and conceptually different methods to the choice or optimization of the associated model parameters. Concerning the electrostatics, the well-known[16,17] dipole moment μ^{g}_{exp} and polarizability α^{g}_{exp} of an isolated water molecule as well as the liquid phase geometry[19,20] $\mathbf{G}^{l}_{m}$ are employed as invariant cornerstones of the model construction. The widths σ of the Gaussian inducible dipole distribution $\boldsymbol{\mu}^{i}(r)$ centered at the oxygen and the spatial distribution of the negative partial charges within the molecule, which characterize the electrostatic signature of the respective model, are determined by DFT/PMM calculations. Solely the three van der Waals parameters (A_1, A_2, B) are eventually obtained by empirical optimization.

1.1. Significance of a TL6P Model. Suppose now that a slightly more complex PMM model, which solely extends the TL5P electrostatics by one additional negative charge, is parametrized by the same approach,[8] and therefore, will be called TL6P, could concurrently reproduce ϵ and α_p. Then such a success would prove that the incorrect electrostatic signatures of the predecessor models TL4P and TL5P are the cause for the noted failures of these and related[28−34] PMM models. In particular, it would demonstrate that neither polarizable van der Waals potentials[34] nor distributed polarizabilities[36,39] are necessary for remedying the shortcomings of PMM four- and five-point models concerning α_p. The check of this possibility is the key objective of this paper.

dx.doi.org/10.1021/jp4125765 | J. Phys. Chem. B 2014, 118, 1589−1602

The successful outcome of this check, which will be described below, has inspired us to a follow-up study,[42] which by some strange fortune happened to be published before the current manuscript was accepted for publication. Based on 20 ns replica exchange molecular dynamics simulations, this follow-up study demonstrates that the accurate TL6P value calculated for the negative density derivative $\alpha_p(T_0,p_0)$ even translates into a remarkably accurate prediction of the whole temperature density profile $n(T,p_0)$ in the range $T \in [250, 320]$ K by TL6P, whereas TL4P and TL5P grossly fail in this respect. As explained in ref 42, this progress sheds light on the microscopic physical cause for the density maximum of liquid water. Thus, the construction and properties of the polarizable six-point model TL6P and the differences to the less complex predecessor models are now even of an enhanced interest.

As follows from Figure 1, the electrostatic signature of the six-point PMM model TL6P is given by the parameter set $\Lambda_{e,6}$

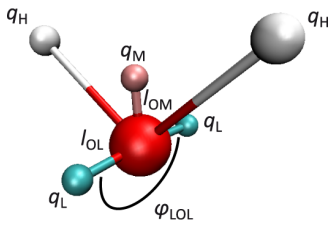


Figure 1. Electrostatic signature of the TL6P model defined by the liquid phase molecular geometry $\mathbf{G}^l_m$, by the distance l_{OM} between the red oxygen atom O and the pink massless point M, whose charge $q_M < 0$ is located on the bisectrix of the HOH triangle, by the distance l_{OL} between O and each of the cyan massless points L, which carry charges $q_L < 0$ and form a triangle LOL perpendicular to the molecular plane, whose bisectrix coincides with that of the HOH triangle, and by the angle φ_{LOL}.

$= \{q_M, l_{OM}, l_{OL}, \varphi_{LOL}\}$. TL6P carries, beyond the partial charges $q_H > 0$ at the hydrogen atoms, negative partial charges $q_{M/L}$ at the massless points M and L. The electrostatic neutrality of the molecule dictates that

$$q_M + 2q_L + 2q_H = 0 \tag{2}$$

Matching one of the corner stones of the parametrization, the charges q_H, q_M, and q_L must additionally reproduce the experimental gas phase dipole moment

$$\mu^g_{exp} = 2q_H h_{HOH} + q_M l_{OM} - 2q_L h_{LOL} \tag{3}$$

where $h_{HOH} = l_{OH}\cos(\varphi_{HOH}/2)$ and $h_{LOL} = l_{OL}\cos(\varphi_{LOL}/2)$ are the heights of the HOH- and LOL-triangles, respectively. According to eqs 2 and 3 only one of the charges, e.g., q_M, can be independently chosen, if the geometric quantities l_{OM}, l_{OL}, and φ_{LOL} are considered as further independent parameters. For $q_M = 0$ the six-point model TL6P reduces to a five-point model with the electrostatic parameters $\Lambda_{e,5} \equiv \{l_{OL}, \varphi_{LOL}\}$ and for $q_L = 0$ to a four-point model with the single parameter $\Lambda_{e,4} \equiv \{l_{OM}\}$.

1.2. DFT/PMM Scheme of TL*v*P Electrostatics Optimization. According to the scheme of parameter optimization, which is explained in great detail in ref 8, the electrostatic parameters $\Lambda_{e,\nu}$ and the Gaussian widths σ of such ν-point PMM models are determined by DFT/PMM calculations. Here for each ν-point PMM potential an ensemble $\mathcal{S}$ consisting of liquid phase structural snapshots s is generated by PMM-MD simulations, which are executed at T_0 and at the experimental density $n_{exp}(T_0,p_0)$ for a constant number N of TLνP models enclosed by a periodic volume V. In subsequent one-point DFT/PMM hybrid calculations on the snapshots $s \in \mathcal{S}$, one of the water molecules is selected as the DFT fragment of the hybrid system such that the remaining PMM fragment represents a spot check of a liquid phase environment.

In this way, DFT/PMM predictions on electrostatic properties of water molecules are determined, which should be representative for aqueous environments at the given conditions. These properties cover the dipole moments $\boldsymbol{\mu}_{DFT/PMM}(s)$ of the DFT fragments and the electrostatic potentials $\Phi(\mathbf{r}|\rho_s)$, which are caused by their electron densities ρ_s at points $\mathbf{r} \in \mathcal{P}$ randomly selected from surrounding spherical surfaces $\mathcal{P}$.

First, the induced dipole moments

$$\boldsymbol{\mu}^i_{DFT/PMM}(s) \equiv \boldsymbol{\mu}_{DFT/PMM}(s) - \boldsymbol{\mu}_{DFT} \tag{4}$$

where $\boldsymbol{\mu}_{DFT}$ is the dipole moment of an isolated DFT water molecule at the liquid phase geometry ($|\boldsymbol{\mu}_{DFT}| = 1.79$ D), are employed to optimize the width σ of the induced Gaussian PMM dipole. To this aim the $\boldsymbol{\mu}^i_{DFT/PMM}(s)$ are compared with the linear responses[8,43]

$$\boldsymbol{\mu}^i_\sigma(s) \equiv \alpha^g_{DFT}\langle \mathbf{E}(s)\rangle_\sigma \tag{5}$$

of Gaussian inducible dipoles of the polarizability $\alpha^g_{DFT} = 1.58$ Å^3, which are placed at the positions of the oxygen atoms within the respective DFT fragments of the snapshots $s \in \mathcal{S}$. These oxygen atoms are located at the origin of the coordinate system and the brackets $\langle ... \rangle_\sigma$ in eq 5 denote the Gaussian average

$$\langle f\rangle_\sigma = \frac{1}{(2\pi\sigma^2)^{3/2}} \int d^3r' f(r') \exp[-r'^2/2\sigma^2]$$

of a function $f(\mathbf{r})$ around that origin. In eq 5 this function is the electric field $\mathbf{E}(\mathbf{r}|s)$, which is generated by the PMM water models surrounding in the chosen snapshot s the selected DFT fragment. The optimal width σ of the Gaussian dipole then follows from optimizing the correlations between the compared data $\boldsymbol{\mu}^i_{DFT/PMM}(s)$ and $\boldsymbol{\mu}^i_\sigma(s)$.

Next, the values $\Phi(\mathbf{r}|\rho_s)$ of the surface potential serve to compute optimized electrostatic signatures. For this purpose the potential $\Phi[\mathbf{r}|\boldsymbol{\mu}^i_\sigma(s)]$, which is generated by the induced Gaussian dipole of a PMM model replacing the DFT fragment in each snapshot $s \in \mathcal{S}$, is subtracted from the DFT/PMM potential $\Phi(\mathbf{r}|\rho_s)$ of the DFT fragment at all points $\mathbf{r} \in \mathcal{P}$, yielding a set of values $\Phi(\mathbf{r}|\rho^{stat}_s)$. This set serves as a mean field template for the static part of the PMM surface potential $\Phi(\mathbf{r}|\Lambda_{e,\nu})$, which is determined by the electrostatic signature $\Lambda_{e,\nu}$. The optimal parameters $\Lambda^{opt}_{e,\nu}$ then follow by minimizing for all $s \in \mathcal{S}$ and all $\mathbf{r} \in \mathcal{P}$ the root-mean-square deviation (RMSD)

$$\zeta(\Lambda_{e,\nu}) \equiv \sqrt{\langle[\Phi(\mathbf{r}|\rho^{stat}_s) - \Phi(\mathbf{r}|\Lambda_{e,\nu})]^2\rangle_{S,\mathcal{P}}} \tag{6}$$

with respect to the parameters $\Lambda_{e,\nu}$.

1.3. Empirical Optimization of the van der Waals Interactions. As soon as the electrostatic parameters σ and $\Lambda_{e,\nu}$ are determined, the three parameters of the Buckingham potential eq 1 are empirically optimized by weak-coupling[44] *NVT* MD simulations executed at at T_0 and $n_{exp}(T_0,p_0)$. Here the targets for A_1, A_2, and B are the position r_1 of the first peak

dx.doi.org/10.1021/jp4125765 | J. Phys. Chem. B 2014, 118, 1589−1602

of an experimental[45−48] oxygen−oxygen radial distribution function (RDF) $g_{OO}(r)$, the standard pressure of $p_0 = 1$ bar, and the experimental[49,50] potential energy $E_{pot}(T_0,p_0) = -9.92$ kcal/mol per molecule, respectively. The sketched computational and empirical steps of parameter optimization are repeated until self-consistency is reached.

The Buckingham parameters A_1 of the TL4P and TL5P models were optimized[8] to match the value $r_1 = 2.76$ Å given by the RDF of Soper.[46] The recent data of Skinner et al.,[48] however, suggest the 1.5% larger value $r_1 = 2.80$ Å. It is unknown how strongly the properties of resulting TLνP models are affected by such a slight change of the target value for A_1. This is one of the minor questions, which we will also address below, by additionally adjusting the Buckingham parameter A_1 of an alternative TL6P model to the most recent[48] RDF. Correspondingly, we will call this alternative model TL6P^{Sk}.

2. METHODS AND SIMULATIONS

The self-consistent scheme of mixed computational and empirical parameter optimization and the various methods and simulations applied to the construction and evaluation of PMM water models are motivated and explained in sections 2 and 3 of ref 8. Apart from a few slight modifications we have exactly copied all these procedures for the construction of the six-point models TL6P and TL6P^{Sk} introduced above. Therefore, the given reference and a short sketch of the copied procedures must suffice in combination with an outline of the few modifications.

2.1. Parameter Optimization. Just like for TL4P and TL5P also the TL6P and TL6P^{Sk} simulation systems, which were employed during the parametrization cycles, contained $N_m = 1500$ water models. Here, N_mVT PMM-MD simulations were executed at $T = T_0$ and $n_{exp}(T_0,p_0)$ with the parallelized program package IPHIGENIE[51] taking advantage of its efficient fast multipole treatment of the electrostatics,[51−54] of its toroidal boundary conditions,[9] and of its moving-boundary reaction field correction[55] for the long-range electrostatics (for further details on the electrostatics computation see section S1 in the Supporting Information). For each of the two TL6P models they served to generate the ensembles $\mathcal{S}$ covering 1000 liquid phase snapshots s for the PMM/DFT calculations. The same conditions were also applied to the weak-coupling simulations[44] aiming at the Buckingham parameters. The PMM/DFT calculations were again executed with the interface[14,56] between IPHIGENIE and the grid-based DFT program CPMD[57] using the same functionals,[58,59] pseudopotential,[60] and 80 Ry plane-wave cutoff as for TL4P and TL5P.[8]

The dipole moments $\boldsymbol{\mu}_{DFT/PMM}(s)$ and the electrostatic surface potentials $\Phi(\mathbf{r}|\rho_s)$ of the DFT fragments belonging to the ensembles $\mathcal{S}$ were then employed to optimize the widths σ and parameter sets $\Lambda_{e,6}$ characterizing the electrostatic signatures of the two TL6P models by the procedures developed in ref 8. As compared to TL5P, each parameter set $\Lambda_{e,6}$ contains four instead of only two parameters, such that a four-dimensional space had to be scanned for the minimization of the deviation functional $\zeta(\Lambda_{e,6})$ defined by eq 6. Here q_M was scanned in the range $[-0.65, -0.35]$ e with steps of 0.0001 e, the distances l_{OM} and l_{OM} in the range [0.0, 0.7] Å with the steps $\Delta l = 0.001$ Å, and the angle φ_{LOL} in the range $[0, 360]^\circ$ with the steps $\Delta\varphi_{LOL} = 1^\circ$.

2.2. New Observables for Quality Control. In the presentation of the quality, by which the surface potential $\Phi(\mathbf{r}|\Lambda_{e,\nu})$ of the respective ν-point model matches the associated ensemble average DFT/PMM reference potential $\langle\Phi(\mathbf{r}|\rho_s^{stat})\rangle_S$ on all points $\mathbf{r} \in \mathcal{P}$, it is advantageous to employ instead of the RMSD $\zeta(\Lambda_{e,\nu})$ given by eq 6 a more direct dimensionless measure. If the values of the potentials are given with respect to the molecular frame of reference, then $\zeta(\Lambda_{e,\nu})$ is related through

$$\xi^2(\Lambda_{e,\nu}) = \zeta^2(\Lambda_{e,\nu}) - {\sigma_S}^2 \tag{7}$$

to the RMSD

$$\xi(\Lambda_{e,\nu}) \equiv \sqrt{\langle[\langle\Phi(\mathbf{r}|\rho_s^{stat})\rangle_S - \Phi(\mathbf{r}|\Lambda_{e,\nu})]^2\rangle_{\mathcal{P}}} \tag{8}$$

of the values of the model potential $\Phi(\mathbf{r}|\Lambda_{e,\nu})$ from the values of the ensemble average surface potential $\langle\Phi(\mathbf{r}|\rho_s^{stat})\rangle_S$ measured at all points $\mathbf{r} \in \mathcal{P}$ and to the average variance

$${\sigma_S}^2 \equiv \langle[\Phi(\mathbf{r}|\rho_s^{stat}) - \langle\Phi(\mathbf{r}|\rho_s^{stat})\rangle_S]^2\rangle_{S,\mathcal{P}} \tag{9}$$

of the DFT/PMM potential values $\langle\Phi(\mathbf{r}|\rho_s^{stat})\rangle$ within the snapshot ensemble $\mathcal{S}$, where the average extends over all points $\mathbf{r} \in \mathcal{P}$. The standard deviation

$$\sigma_{\mathcal{P}} \equiv \sqrt{\langle[\langle\Phi(\mathbf{r}|\rho_s^{stat})\rangle_S - \langle\langle\Phi(\mathbf{r}|\rho_s^{stat})\rangle_S\rangle_{\mathcal{P}}]^2\rangle_{\mathcal{P}}} \tag{10}$$

of the ensemble average values $\langle\Phi(\mathbf{r}|\rho_s^{stat})\rangle_S$ of the DFT/PMM surface potential with respect to all points $\mathbf{r} \in \mathcal{P}$ then characterizes for $\langle\Phi(\mathbf{r}|\rho_s^{stat})\rangle_S$ the dynamical range of variations on the surrounding surface and, therefore, defines a scale on which the standard deviations σ_S and the RMSD $\xi(\Lambda_{e,\nu})$ can be measured by the dimensionless quantities $\sigma_S/\sigma_{\mathcal{P}}$ and $\xi(\Lambda_{e,\nu})/\sigma_{\mathcal{P}}$, respectively.

2.3. Evaluation of the TL6P Models. To evaluate the quality of the resulting TL6P and TL6P^{Sk} models by comparisons with relevant experimental data, we applied the methods described in ref 8 and calculated the same molecular and bulk phase observables that were previously evaluated for the TLνP models with $\nu \leq 5$.

The bulk properties at T_0 and $n_{exp}(T_0,p_0)$ were obtained from PMM-MD simulations with durations and settings identical to the ones enumerated in Tables 2 and 3 of ref 8 for three different system sizes N_s, N_m, and N_l. While the numbers $N_s = 728$ and $N_m = 1500$ of molecular models in the small and intermediately sized periodic boxes were chosen identical, the large boxes contained this time $N_l = 5300$ instead of only 3374 molecular models. Thus, we carried out $N_iV_iT_0$ simulations, i $\in$ {s, m, l}, with durations of 1.5−4.5 ns at the experimental density $n_{exp}(T_0,p_0)$ and at slightly varied densities $n_\pm = (1 \pm 0.05)n_{exp}$ applying Bussi[61] and minimally invasive (MI) Berendsen thermostats.[62] Furthermore, we carried out $N_ip_0T_0$ and $N_sp_0T_\pm$ simulations ($T_\pm = T_0 \pm 10$ K) for similar time spans controlling the pressure by a slowly coupled (10 ps) Berendsen barostat.[63]

The evaluated molecular properties cover the quadrupole moments of the monomers and the geometries $\mathbf{G}_d$, potential energies $E_{pot,d}$, and dipole moments μ_d of the dimers. Among the bulk phase observables are those that were targets of the empirical optimization, i.e., the average potential energy per molecule $\langle E_{pot}\rangle$ and the average pressure $\langle p\rangle$, and a series of other observables,[8] for which the computational results represent predictions. These are the average density $\langle n\rangle$ (from the $N_ip_0T_0$ simulations), the dielectric constant ϵ (from the fluctuations of the total dipole moment in the $N_iV_iT_0$ simulations), the self-diffusion coefficient D and viscosity ν (from the size dependence[4] of D observed by means of the

dx.doi.org/10.1021/jp4125765 | J. Phys. Chem. B 2014, 118, 1589−1602

$N_iV_iT_0$[MI] simulations), the isothermal compressibility κ_T (from the $N_sV_{\pm}T_0$ simulations), and the heat capacity C_p and the thermal expansion coefficient α_p (from the $N_sp_0T_{\pm}$ simulations).

According to our original design of this study, the evaluation of the TL6P models should be restricted to the liquid phase standard conditions T_0 and p_0. The capability to predict also liquid phase properties at other conditions, e.g., the temperature profiles $n(T,p_0)$ of the density, $\alpha_p(T,p_0)$ of the thermal expansion coefficient, $\Delta H_{vap}(T,p_0)$ of the heat of vaporization, and $C_p(T,p_0)$ of the heat capacity, was planned for a possible follow-up study, which we wanted to tackle in the case of a favorable outcome of the present study. These data are meanwhile presented and discussed in ref 42 and in the associated Supporting Information. Following reviewer requests we have, however, included in the current study further information on gas phase and solid state properties like the second virial coefficient $B_2(T)$ and the melting temperature T^m of ice Ih. The computational methods applied in these additional simulation studies are described in sections S7 (B_2) and S10 (T^m) of the Supporting Information, which present and discuss also the associated results. The following text will take up these issues of transferability to other than standard liquid phase conditions solely in a condensed and summarizing form in section 4.3.

3. RESULTS

Before presenting the molecular and bulk phase properties of the TL6P and TL6P^{Sk} models resulting from the mixed computational and empirical optimization, we address a remarkable peculiarity of these six-point models, which became apparent through the DFT/PMM optimization of their electrostatic signatures.

3.1. Quality of the Electrostatic DFT/PMM Refinement. First the correlations between the DFT/PMM and PMM induced dipole moments $\boldsymbol{\mu}^i_{DFT/PMM}(s)$ (eq 4) and $\boldsymbol{\mu}^i_{\sigma}(s)$ (eq 5) were employed to optimize the widths σ resulting in the values 0.806 Å for TL6P and 0.802 Å for TL6P^{Sk}. The resulting optimal correlations are depicted by Figure S4 in section S2 of the Supporting Information. The access to optimized widths σ is a prerequisite for the DFT/PMM calculation of the electrostatic signatures $\Lambda_{e,6}$.

As explained above in connection with eq 6, the parameters $\Lambda_{e,6}$ characterizing the electrostatic signatures of the TL6P models are optimized by comparing the zero-field potential $\Phi(\mathbf{r}|\Lambda_{e,6})$ of the respective model with the supposedly static parts $\Phi(\mathbf{r}|\rho_s^{stat})$ of the potentials $\Phi(\mathbf{r}|\rho_s)$, which are generated by the electron densities ρ_s of the DFT fragments in liquid phase snapshots $s \in \mathcal{S}$ at points $\mathbf{r} \in \mathcal{P}$ on spherical surfaces surrounding the fragments. Here, the supposedly static parts are calculated for all $s \in \mathcal{S}$ as differences

$$\Phi(\mathbf{r}|\rho_s^{stat}) = \Phi(\mathbf{r}|\rho_s) - \Phi[\mathbf{r}|\mu^i_{\sigma}(s)]$$

between the potentials $\Phi[\mathbf{r}|\boldsymbol{\mu}^i_{\sigma}(s)]$ of induced Gaussian dipoles $\boldsymbol{\mu}^i_{\sigma}(s)$ at the positions of the respective DFT fragment's oxygen atom and the DFT surface potentials $\Phi(\mathbf{r}|\rho_s)$. The addressed comparison is effected by means of the RMSD $\zeta(\Lambda_{e,6})$ defined by eq 6.

Now it is interesting to check, whether the supposedly static parts $\Phi(\mathbf{r}|\rho_s^{stat})$ of the DFT/PMM potentials $\Phi(\mathbf{r}|\rho_s)$ are actually static. Using the DFT fragment's molecular frame of reference, we have therefore calculated the average values $\langle\Phi(\mathbf{r}|\rho_s^{stat})\rangle_{\mathcal{S}}$ and the variances $\sigma_{\mathcal{S}}{}^2(\mathbf{r})$ of the values $\Phi(\mathbf{r}|\rho_s^{stat})$ within $\mathcal{S}$ not only for the two TL6P potentials but also for their TL5P and TL4P predecessors. Averaging over all points $\mathbf{r} \in \mathcal{P}$ and taking the square root then yielded the standard deviations $\sigma_{\mathcal{S}}$ defined by eq 9, which measure the average sizes of the fluctuations in the ensembles $\mathcal{S}$ of potential values $\Phi(\mathbf{r}|\rho_s^{stat})$. If these potentials should be actually static, as assumed by the construction of the TLνP models (ν = 4, 5, 6), then the $\sigma_{\mathcal{S}}$ should be very small as compared to the variations of the ensemble average potentials $\langle\Phi(\mathbf{r}|\rho_s^{stat})\rangle_{\mathcal{S}}$ on the spherical surfaces. The sizes of these variations are measured by the quantities $\sigma_{\mathcal{P}}$ defined by eq 10.

Table 1 lists the resulting standard deviations $\sigma_{\mathcal{S}}$ in units of $\sigma_{\mathcal{P}}$. Apparently, the fluctuations $\sigma_{\mathcal{S}}$ of the snapshot potentials

Table 1. Standard Deviations $\sigma_{\mathcal{S}}$ and RMSDs $\xi(\Lambda^{opt}_{e,\nu})$ Measured on the Scales $\sigma_{\mathcal{P}}$ As Defined by Eqs 8−10 of the Optimal TLνP Models

model	TL6P	TL6P^{Sk}	TL5P	TL4P
$\sigma_{\mathcal{S}}/\sigma_{\mathcal{P}}$ [%]	0.2	0.2	0.3	0.3
$\xi(\Lambda^{opt}_{e,\nu})/\sigma_{\mathcal{P}}$ [%]	2.1	2.0	5.7	7.9

$\Phi(\mathbf{r}|\rho_s^{stat})$ are smaller than 0.3% of the total variation $\sigma_{\mathcal{P}}$ of the respective average surface potentials $\langle\Phi(\mathbf{r}|\rho_s^{stat})\rangle_{\mathcal{S}}$. Thus, the supposedly static parts $\Phi(\mathbf{r}|\rho_s^{stat})$ of the DFT/PMM potentials $\Phi(\mathbf{r}|\rho_s)$ are nearly static, proving that one of the key assumptions of our procedure for model constructions is valid.

Table 1 additionally lists the RMSDs $\xi(\Lambda^{opt}_{e,\nu})$ between the surface potential values $\Phi(\mathbf{r}|\Lambda^{optt}_{e,\nu})$, which are generated by the optimized electrostatic signatures $\Lambda^{opt}_{e,\nu}$ of the TLνP models, and the ensemble average potential values $\langle\Phi(\mathbf{r}|\rho_s^{stat})\rangle_{\mathcal{S}}$. Because of eq 7 and because the $\sigma_{\mathcal{S}}$ are small, these RMSDs $\xi(\Lambda^{opt}_{e,\nu})$ are almost identical to the optimization functionals $\zeta(\Lambda^{opt}_{e,\nu})$, which additionally shows that our optimization actually aims at the average DFT/PMM surface potential $\langle\Phi(\mathbf{r}|\rho_s^{stat})\rangle_{\mathcal{S}}$.

Interestingly, the RMSDs $\xi(\Lambda^{opt}_{e,6})$ of the six-point models are by factors of about 3 and 4 smaller than those of the five- and four-point models, respectively. Hence, the six-point models represent the static part $\langle\Phi(\mathbf{r}|\rho_s^{stat})\rangle_{\mathcal{S}}$ of the DFT/PMM mean field surface potential substantially better than their electrostatically less complex predecessors.

As shown above, in our construction of the TLνP models we have taken the DFT/PMM liquid phase quantity $\langle\Phi(\mathbf{r}|\rho_s^{stat})\rangle_{\mathcal{S}}$ as the reference for the optimization of the electrostatic signatures $\Lambda_{e,\nu}$. Now one may ask how much this quantity differs from the surface potential $\Phi(\mathbf{r}|\rho^{vac})$, which is generated by the electron density ρ^{vac} of an isolated water molecule and is calculated, using the liquid phase geometry $\mathbf{G}^i_m$, by the same DFT approach. This question can be answered by replacing in eq 8 the PMM potential $\Phi(\mathbf{r}|\Lambda_{e,\nu})$ by the DFT potential $\Phi(\mathbf{r}|\rho^{vac})$ and by evaluating the new RMSD $\xi(\rho^{vac})$ between $\Phi(\mathbf{r}|\rho^{vac})$ and $\langle\Phi(\mathbf{r}|\rho_s^{stat})\rangle_{\mathcal{S}}$ at the surface points $\mathbf{r} \in \mathcal{P}$.

We found that $\xi(\rho^{vac})$ is only 0.4% of the total variation $\sigma_{\mathcal{P}}$ of $\langle\Phi(\mathbf{r}|\rho_s^{stat})\rangle_{\mathcal{S}}$; i.e., the differences between this average liquid phase potential and the vacuum result $\Phi(\mathbf{r}|\rho^{vac})$ are very small. Nevertheless, such differences exist and can be characterized by taking $\Phi(\mathbf{r}|\rho^{vac})$ as a reference for the optimization of $\Lambda_{e,\nu}$. Some of the resulting vacuum model parameters (q_M, φ_{LOL}) turn out to be identical to the liquid phase result, whereas others (l_{OM}, l_{OL}) increase by 2−3% (data not shown). Thus, the enhanced computational effort of employing DFT/PMM liquid phase ensembles has not been completely superfluous.

dx.doi.org/10.1021/jp4125765 | J. Phys. Chem. B 2014, 118, 1589−1602

Table 2. TL6P Parameters Compared to Those of Their Predecessors[8]

model	l_{OM} [a]	l_{OL} [a]	φ_{LOL} [b]	q_H [c]	q_M [c]	q_L [c]	σ [a]	A_1 [d]	A_2 [e]	B [f]
TL6P	0.38	0.46	175	0.51	−0.57	−0.22	0.80	2.9	4.0	6.6
TL6P^{Sk}	0.39	0.46	173	0.51	−0.57	−0.22	0.80	3.0	4.0	6.3
TL5P		0.32	268	0.53		−0.53	0.85	0.6	3.4	11.8
TL4P	0.24			0.56	−1.12		0.84	0.8	3.6	9.9

[a] Å. [b] deg. [c] e. [d] 10^5 kcal/mol. [e] Å^{-1}. [f] 10^2 (Å^6 kcal)/mol.

Having ascertained the strongly enhanced quality, by which the electrostatic signatures $\Lambda_{e,6}$ match the DFT/PMM reference, we can now turn to the resulting molecular and bulk phase properties. Assuming that the DFT/PMM approach yields accurate descriptions, the properties obtained with TL6P should represent the experimental findings much better than the ones found[8] for the less complex predecessors.

3.2. Properties of the TL6P Potentials. Table 2 lists the final parameters $\lambda \in \Lambda_{e,6} \cup \{\sigma\} \cup \{A_1, A_2, B\}$ obtained for the two TL6P models by the iterative optimization scheme described in detail in ref 8 and sketched further above. Additionally, these parameters are compared with those of the preceding TL5P and TL4P models. The shown values are rounded. A table containing all relevant digits is provided for the TL6P models by section S3 of the Supporting Information. This section additionally gives parameters of Lennard-Jones potentials, which are almost equivalent to the TL6P Buckingham potentials.

3.2.1. Electrostatic Signatures from DFT/PMM. As explained by Figure 1, the electrostatic geometry of the TL6P models is defined by the distances l_{OM} and l_{OL} of the massless charge sites M and L from the oxygen atom O and by the angle φ_{LOL} spanned by the two massless "lone pair" sites L with respect to the center O. Figure 1 actually represents the optimized TL6P values specified in Table 2. Because the electrostatic parameters of TL6P and TL6P^{Sk} are almost identical, Figure 1 equally represents the latter model.

The angle φ_{LOL} of the TL6P models is smaller than 180° such that the projection of the points $\mathbf{r}_L$ onto the molecular plane falls slightly outside the molecular HOH-triangle on its bisectrix at a distance of 0.02 Å from $\mathbf{r}_O$. For TL5P, in contrast, the projection is 0.22 Å inside. Furthermore, for TL6P the distance l_{OL} is larger and the absolute value $|q_L|$ of the charge smaller than for TL5P.

An interesting feature of each of the two TL6P models is the very large distance l_{OM} of the site M from the oxygen atom (cf. Table 2). It is by about 60% larger than the corresponding distance in TL4P, which is moreover quite typical for PMM four-point models.[28−32,64] On the other hand, the absolute value $|q_M|$ of the associated TL6P charge is by about a factor 2 smaller than its four-point counterpart such that in TL6P the total negative charge is about evenly distributed among the M and combined L sites. Because q_M is small and l_{OM} large, the M site contributes only 30% to the 1.855 D static dipole moment of TL6P. Here we stress that it is this unexpected distribution of static partial charges within TL6P, that generates the favorable match of surface potentials discussed above in section 3.1.

In light of other nonpolarizable[66,67] and polarizable[65,68] six-point models, the electrostatic TL6P geometry represents a surprise. As is illustrated in Figure 2, the empirically optimized SWM6 model,[65] for instance, features a short l_{OM} distance (0.247 Å) and a large negative q_M charge (−1.133 e), which are both typical for PMM four-point models, and a very small angle

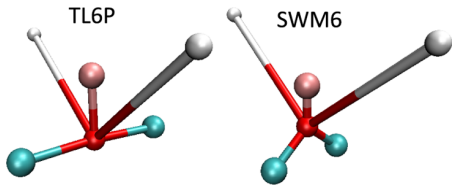


Figure 2. Electrostatic geometries of the two PMM six-point models TL6P and SWM6.[65]

φ_{LOL} = 101° combined with small negative charges q_L = −0.11 e and small distances l_{OL} = 0.315 Å. In the zero-field case SWM6 additionally has a net positive charge of 0.288 e at O, because the negative charge of the Drude particle attached to O is overcompensated by the positive charge of O. As a result SWM6 has seven charge centers, for which Coulomb interactions have to be calculated. The complex PMM six-point model suggested by Wang et al.,[68] on the other hand, features an angle φ_{LOL}, which resembles that of TL5P and yields a projection of the L sites inside the HOH triangle on the bisectrix 0.231 Å distant from $\mathbf{r}_O$. Furthermore, the distance l_{OM} is by 24% smaller than that of TL6P. Thus, one of these models resembles TL4P and the other a mixture of TL4P and TL5P but none TL6P.

3.2.2. Widths of the Gaussian Inducible Dipoles from DFT/PMM. The Gaussian widths σ for the TL6P models, which are listed in Table 2, were given above in section 3.1 and derived by optimizing the PMM and DFT/PMM dipole correlations in section S2 of the Supporting Information. They are a little smaller than those of TL4P and TL5P. This decrease should cause slightly enhanced induced dipole interactions at small distances.

3.3.3. Buckingham Parameters from Empirical Optimizations. Much more significant are the differences among the listed Buckingham parameters. By construction also the two TL6P models differ in this respect. Recall that the target r_1 = 2.76 Å, which is the position of the first peak in Soper's[46] RDF $g_{OO}(r)$ and was weakly coupled to A_1 during the optimization of TL6P, was exchanged by the value[48] r_1 = 2.80 Å to yield TL6P^{Sk}. Correspondingly, the TL6P^{Sk} value of A_1 is about 5% larger and the value of B by as much smaller than the corresponding TL6P values. Thus, the 1.5% difference of the targets r_1 has translated into 5% differences of A_1 and B. Remarkably, the RDFs of the two TL6P models, which are depicted in Figure S8 of the Supporting Information, are almost identical at distances larger than about 3.0 Å and show differences only well within the first peak, indicating that the noted differences of A_1 and B solely affect the first solvation shells of the simulated liquids.

Much larger are the differences among the Buckingham potentials for TLnP models of different complexity. According to Table 2 the TL4P and TL5P potentials show some similarity but both strongly differ from the TL6P potential. The latter exhibit a much steeper and more short-range repulsion than the

dx.doi.org/10.1021/jp4125765 | J. Phys. Chem. B 2014, 118, 1589−1602

former, as is witnessed by the much larger values of A_1 and A_2. The dispersion attraction, on the other hand, whose strength is measured by B, becomes smaller by a factor of almost 2 in the transition from TL4P and TL5P to the six-point case.

3.2.4. Quadrupole Moments from DFT/PMM. In section 3.1 we have seen that the potential $\Phi(\mathbf{r}|\rho^{vac})$ generated by an isolated DFT water molecule on a surrounding spherical surface deviates from its DFT/PMM liquid phase relative $\langle\Phi(\mathbf{r}|\rho_s^{stat})\rangle_S$ by an RMSD of only 0.4% as measured by the variation $\sigma_{\mathcal{P}}$ of $\langle\Phi(\mathbf{r}|\rho_s^{stat})\rangle_S$ (cf. section 2.2). The main contributions to these surface potentials should be due to the quadrupole moments of the DFT and average DFT/PMM charge densities ρ^{vac} and $\langle\rho^{stat}\rangle$, respectively. Thus, these quadrupole moments should be quite similar. However, because the DFT/PMM surface potentials $\langle\Phi(\mathbf{r}|\rho_s^{stat})\rangle_S$ served as the references in the optimizations of TL6P electrostatic signatures $\Lambda_{e,6}$, the quadrupole moments of an isolated DFT water molecule should also closely resemble those of the TL6P models.

Table 3 confirms this expectation. Here, the comparison of the DFT vacuum result with the TL6P quadrupole moments is

Table 3. Quadrupole Moments of TL6P and of Other PMM Water Models, Which Are Given with Respect to the Center of Mass, Compared with Results of a DFT Vacuum Description and with Experimental Gas Phase Data through the Associated RMSDs R_{DFT} and R_{exp} As Measured by the Standard Deviation σ_{exp} = 2.1 DÅ of the Gas Phase Data

	Q_{xx}/DÅ	Q_{yy}/DÅ	Q_{zz}/DÅ	R_{DFT}/σ_{exp}	R_{exp}/σ_{exp}
exp[69]	−0.13	2.63	−2.50	3.6%	
DFT^{vac}	−0.19	2.59	−2.39		3.6%
TL6P	−0.18	2.59	−2.41	0.6%	3.0%
$TL6P^{Sk}$	−0.19	2.59	−2.40	0.2%	3.4%
$TL5P^{8}$	−0.11	2.53	−2.42	2.7%	3.4%
$TL4P^{8}$	−0.29	2.53	−2.24	5.4%	9.0%
SWM6[65]	−0.30	2.39	−2.09	10.5%	14.0%

handily expressed by the associated RMSD R_{DFT} measured on the scale σ_{exp}, which is given by the variations among the values of the elements of the experimental[69] quadrupole tensor.

For the TL6P models R_{DFT} is less than 1% of σ_{exp} and, thus, is of the same order of magnitude as the 0.4% deviation of the surface potentials $\Phi(\mathbf{r}|\rho^{vac})$ and $\langle\Phi(\mathbf{r}|\rho_s^{stat})\rangle_S$ mentioned above. With decreasing complexity of the model R_{DFT} strongly increases measuring 2.7% for TL5P and 5.4% for TL4P. Because of the similarity between the surface potentials $\Phi(\mathbf{r}|\rho^{vac})$ and $\langle\Phi(\mathbf{r}|\rho_s^{stat})\rangle_S$, this increase is related to the one observed earlier in Table 1 for the match of the TLνP surface potentials with the DFT/PMM reference.

On the other hand, the quadrupole moments calculated by our DFT approach[8,58−60] (cf. section 2.1) for an isolated water molecule deviate from the experimental data by R_{DFT}/σ_{exp} = 3.6% indicating that this DFT method describes the electrostatic signature of an isolated water molecule quite well but not perfectly. This comparison requires the caveat, however, that the experimental liquid phase geometry had been assumed in the DFT calculation, whereas the experimental data pertain to the gas phase geometry. Note in this context that the dipole moment calculated with this DFT approach[43] for an isolated water molecule at the optimized geometry underestimates the experimental gas phase value by 2.4%.

The RMSDs R_{exp} between the quadrupole moments of the various models and the gas phase data (see the last column of Table 3) demonstrate that DFT^{vac} represents the gas phase data about as well as the TL6P models. Surprisingly also the TL5P model happens to perform well in this respect, whereas for TL4P, in contrast, R_{exp} is by about a factor of 3 larger than for the TL6Ps. For the empirical SWM6 model R_{exp} is even by a factor of about 5 larger despite the fact that, here, the experimental quadrupole moments were among the optimization targets.[65] Note that for the effective model iAMOEBA[39] R_{exp} is similarly large (12%).

3.3. Evaluation. The construction of the two TL6P potentials aimed at water molecules embedded in liquid phase environments at T_0 and $n_{exp}(T_0,p_0)$. Thus, the quality at which the TL6P liquids resemble water at these conditions are of primary interest and, therefore, we will start with these properties.

Only after a close similarity between liquid water and its computer model at T_0 and p_0 has been demonstrated, issues of transferability gain interest. Examples are the performance at other temperatures,[42] or with respect to the experimentally well-known properties of the isolated water dimer[69−71] and of vapor.[72] Also the TL6P prediction on the melting temperature[73] T^m_{exp} = 273.15 K of ice Ih can add insights here. Correspondingly, these themes will be addressed later in section 4.3.

3.3.1. Targeted Bulk Liquid Properties. Section S7 in the Supporting Information presents and discusses those properties of the liquid at T_0 and $n_{exp}(T_0,p_0)$, which were targeted during the empirical optimization of the Buckingham parameters A_1, A_2, and B. Here, the $N_iV_iT_0$ test simulations, which were executed for the three system sizes i ∈ {s, m, l}, have demonstrated for the two TL6P models that the absolute values of the binding energies $\langle E_{pot}\rangle$ are by only 0.1−0.4% larger than the target value of −9.92 kcal/mol, that the average pressures $\langle p\rangle$ fluctuate in a small range of ±30 bar around the target value p_0 = 1 bar, and that the position r_1 of the first maximum in the RDF $g_{OO}(r)$ of TL6P is at 2.76 Å as given by Soper's neutron diffraction data[46] whereas r_1 has for $TL6P^{Sk}$ the required[48] value of 2.80 Å.

3.3.2. Radial Distribution Functions. Interestingly, the RDFs of the two TL6P models, which are depicted in Figure S8 of the Supporting Information, are almost identical for distances larger than 3 Å and match the recent RDF of Skinner et al.[48] very well for distances beyond 4.5 Å. Thus, the differences of the TL6P and $TL6P^{Sk}$ liquid structures at T_0 are confined to the inner parts of the respective first solvation shells. One might have expected such short ranged differences from the fact that the target r_1 had been coupled to the parameter A_1, which steers the strength of the repulsive part of the Buckingham potential [cf. Equation 1].

3.3.3. Predicted Bulk Liquid Properties. We start the presentation of the TL6P predictions on properties of the liquid water at p_0 and T_0 by shortly sketching the results for the diffusion constant D_0 and viscosity η, for the isothermal compressibility κ_T, and for the isobaric heat capacity C_p. By presenting results on the average dipole moment and on the dielectric constant ϵ_0, we subsequently address in some more detail the important issues of polarity, before we finally turn to the average density $\langle n\rangle$ and its temperature derivative, the thermal expansion coefficient α_p.

Table 4 compares the TL6P predictions enumerated above in the given sequence with corresponding experimental

dx.doi.org/10.1021/jp4125765 | J. Phys. Chem. B 2014, 118, 1589−1602

Table 4. TL6P Bulk Properties Compared with TL5P/TL4P[8] and with Experimental Data

model	D_0 [a]	η [b]	κ_T [c]	C_p [d]	ϵ_0	α_p [e]
TL6P	2.3	0.86	40.1	18.3	77	2.9
TL6P^{Sk}	2.2	0.92	39.7	18.6	79	2.8
TL5P	2.8	0.98	36.4	19.6	67	7.2
TL4P	3.0	0.80	37.4	18.7	77	5.9
exp	2.4[74]	0.81[75]	45.6[26]	18.0[76]	78[27]	2.8[26]

[a] nm^2/ns. [b] mPa s. [c] 10^{-6}/atm. [d] cal/(mol K). [e] 10^{-4} K^{-1}.

data[26,27,74−76] and with properties[8] of the TL5P and TL4P predecessors (an exception is the density $\langle n \rangle$, which will be separately documented). The statistical errors associated with the TL6P data in Table 4 are listed by Table S10 in section S8.1 of the Supporting Information. Other PMM models are hardly ever mentioned in the following discussion, because extensive and very detailed comparisons of this kind led in ref 8 to the conclusion that TL5P and TL4P feature "bulk phase properties comparable to or even better than those of similarly complex models". Therefore, the TL5P and TL4P descriptions can be safely taken as very good representatives for the large class of similarly complex PMM water models.

3.3.3.1. Self-diffusion Coefficients and Viscosities. Section S8.2 in the Supporting Information explains how the self-diffusion constant D_0 and viscosity η of the infinite system were derived by applying the Dünweg−Kremer formula (eq S12, Supporting Information) to the results of the finite size $N_iV_iT_0$[MI] simulations described in section 2.3. The results of this analysis are collected in the first two data columns and lines of Table 4.

Inspecting the data immediately demonstrates that the two TL6P models describe the experimental diffusivity of liquid water at T_0 and p_0 much better than their predecessors. Instead of considerable overestimates in the range of 17−25%, the more complex TL6P potentials now yield slight underestimates in the range 4−8%. Hence, the TL6P models, whereas approaching the mobility of real water molecules, are considerably less mobile than the TL5P and TL4P models.

A reduced mobility should be concurrently visible in a larger viscosity, which is the case for the TL6P models in comparison with the real liquid but is only partially the case as compared to the earlier models. Compared to TL4P, the TL6P values of η do show the increase expected from the smaller TL6P values for D_0. Compared to TL5P, however, the TL6P values of η are smaller despite the smaller TL6P diffusion constants.

The noted inconsistency is most likely caused by the quite substantial statistical errors of 10−20%, which are according to Table S10 in section S8.1 of the Supporting Information attached to the values of η. These errors are large, because the η-values derive by eq S12 (Supporting Information) from slopes of regression lines, such as those shown by Supporting Information Figure S9 in section S8.2 for the TL6P models, and because the values of such slopes are very sensitive to a small statistical scatter of the underlying data. In contrast, the extrapolated diffusion coefficients show errors of only 1−2% and, hence, are much more reliable.

3.3.3.2. Isothermal Compressibility. A glance at the third data column in Table 4 immediately reveals that also the isothermal compressibility $\kappa_T = (\partial \ln n/\partial p)_T$ at n_{exp} and T_0 is better described by the TL6P than by the less complex models. The shown model data were derived from the respective $N_sV_\pm T_0$ simulations by numerical differentiation.[77] In the transition from TL4P/TL5P to the TL6P models the underestimates of κ_T are reduced from about 20% to about 12%. Note that according to Table S10 in the Supporting Information the statistical uncertainties $\sigma(\kappa_T)$ are smaller than 1%.

3.3.3.3. Heat Capacity. Additionally including a required quantum correction,[25] the isobaric heat capacity C_p is calculated as a numerical temperature derivative of the total energy per molecule from the $N_sp_0T_\pm$ simulations. The TL6P data on C_p are listed in the fourth data column of Table 4 and feature the same pattern, which we already observed for the diffusivity D_0 and isothermal compressibility κ_T. According to this pattern, TL6P generally performs a little better than TL6P^{Sk} and these two six-point PMM models perform much better than TL5P/TL4P. For $C_p(T_0,p_0)$, in particular, the six-point models thus overestimate the experimental value only by 1−3%, whereas the overestimates provided by the predecessors were larger (4−9%). Because the statistical uncertainties of the C_p values obtained by simulations are in the range of 2% (cf. Supporting Information Table S10), the heat capacity predicted by TL6P, in particular, is statistically indistinguishable from the experimental value.

3.3.3.4. Dipole Distributions. Section S8.3 in the Supporting Information presents and discusses the PMM dipole distributions resulting from the $N_mV_mT_0$ simulations of the TL6P models. The resulting average dipole moments $\langle \mu \rangle$ are almost equal at about 2.53 D.

With this value the average dipole moments $\langle \mu \rangle$ are very close to the averages $\langle \mu_{DFT/PMM} \rangle \approx 2.54$ D of the DFT/PMM dipole moments, which were employed to optimize the Gaussian widths σ of the inducible TL6P dipoles (cf. section S2 in the Supporting Information). Thus, the TL6P potentials not only represent the static part $\langle \Phi(\mathbf{r}|\rho_s^{stat}) \rangle_S$ of the DFT/PMM mean field surface potential very well (cf. section 3.1) but also excellently mimic the dipolar response properties of their DFT/PMM prototypes.

Furthermore, the TL6P averages $\langle \mu \rangle$ are in the range 2.4−2.6 D. It has been previously suggested that only those PMM models can have a dielectric constant ϵ close to the experimental value $\epsilon(T_0,p_0)$, whose average dipole moments $\langle \mu \rangle$ are in the given range.[29,78] Thus one may expect that the dielectric constants of the two TL6P potentials match $\epsilon(T_0,p_0)$ quite well.

3.3.3.5. Dielectric Constant. In computations of the dielectric constant ϵ special care has to be taken, if one employs, like our MD program IPHIGENIE,[51] a moving-boundary reaction field approach[55] for the treatment of the long-range electrostatics. In this computational scenario, the simulated system is enclosed by toroidal boundary conditions[9] and is, beyond the distance dictated by the minimum image convention,[9] embedded in a dielectric continuum with the dielectric constant ϵ_{RF}. If the dielectric constant ϵ of the simulated model differs from ϵ_{RF}, then the value calculated for ϵ by Neumann's[79] formula (cf. eq 8 in ref 8) from the fluctuations of the total dipole moment in a $N_iV_iT_0$ simulation depends on V_i. This size dependence can be diminished by applying an iterative correction,[80] which is based on perturbation theory. But nevertheless, the finite size results $\epsilon(N_iV_iT_0)$ must be eventually extrapolated to the value ϵ_0 applicable to the infinitely large system.

The data resulting from the $N_iV_iT_0$ simulations [i ∈ {s, m, l}] are represented in Figure 3 as functions of the inverse box lengths L^{-1} together with regression lines, which enable

dx.doi.org/10.1021/jp4125765 | J. Phys. Chem. B 2014, 118, 1589−1602

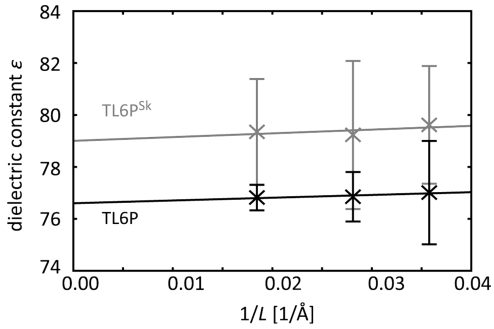


Figure 3. Size dependent dielectric constants $\epsilon(N_iV_iT_0)$ calculated for TL6P (black crosses) and TL6P^{Sk} (gray crosses) together with error bars indicating statistical uncertainties, which were calculated as described in the text. The depicted linear regressions serve to extrapolate to the respective dielectric constants ϵ_0 of the infinite systems at $L^{-1} = 0$.

extrapolations to the infinite systems at $L^{-1} = 0$. The linear regressions assign errors $\sigma(\epsilon_0)$ to the extrapolated values ϵ_0, which are smaller than 1% (cf. Table S10 in the Supporting Information). To estimate the statistical errors of the simulation results $\epsilon(N_iV_iT_0)$, the associated trajectories were partitioned into two parts. The deviations of the two values $\epsilon_k(N_iV_iT_0)$, = 1, 2, from their common average were taken as measures for the associated standard deviations and are drawn as error bars in Figure 3.

In agreement with the above expectations, which were based on the TL6P values of the average dipole moments $\langle\mu\rangle$, the extrapolated dielectric constants ϵ_0 deviate from the experimental value by only about 1%. A similarly close value had been previously obtained for TL4P, whereas the dielectric constant of TL5P showed a much larger (14%) deviation (Table 4).

In the Introduction we have pointed out that PMM models, whose electrostatic signatures are described by three or four charges at fixed positions within the molecule, cannot concomitantly describe the dielectric constant ϵ and the thermal expansion coefficient α_p at p_0 and T_0 with a comparable accuracy. Having ascertained that the TL6P models reproduce ϵ very well at these conditions, we now turn to their performances on the average density $\langle n\rangle$ and its temperature derivative α_p.

3.3.3.6. Average Density at p_0. As is documented by Table S12 in section S8.4 of the Supporting Information, the $N_ip_0T_0$ simulations predict for both TL6P models average densities $\langle n_i\rangle(T_0,p_0)$ that agree, independently of the system size i ∈ {s, m, l}, remarkably well with the experimental value[26] of 0.9965 g/cm^3. The deviations of the simulated values $\langle n_i\rangle(T_0,p_0)$ from the experimental one are smaller than the largest statistical error of the simulation data, which is 0.3%.

3.3.3.7. Thermal Expansion Coefficient. The isobaric thermal expansion coefficients were approximately calculated from the $N_sp_0T_\pm$ simulations according to[81]

$$\alpha_p \approx -\frac{\ln[\langle n\rangle(p_0,T_+)] - \ln[\langle n\rangle(p_0,T_-)]}{T_+ - T_-} \quad (11)$$

by numerical differentiation with $T_\pm = T_0 \pm 10$ K.

Surprisingly, the resulting values, which are listed in Table 4, match the experimental reference extremely well for both TL6P models showing deviations of at most 4%. Note that these descriptions of α_p are even better than the ones achieved by those previous models (TIP4P/2005[40] and BKd3[34]), which performed exceptionally well (≈±10%) on this particular observable, because α_p had been among the targets of the respective empirical optimization. Within the statistical uncertainties $\sigma(\alpha_p) = 0.6 \times 10^{-4}$ K^{-1} given in Table S10 of the Supporting Information, the two TL6P simulation results for α_p are indistinguishable from the experimental value.

The electrostatically less complex predecessor models, in contrast, show overestimates by factors of 2.1−2.6. When TL6P is compared to TL5P, it is the addition of a negative charge $q_M = -0.57$ e on the bisectrix of the HOH triangle at the quite large distance $l_{OM} = 0.38$ Å together with slightly modified positions and values of the "lone pair" charges q_L (cf. Table 2) that apparently causes the substantially improved description of the isobaric thermal expansion coefficient $\alpha_p(T_0,p_0)$.

Concerning α_p, the details of the electrostatic signatures of the TL6P models are of key importance as one can see by reconsidering Figure 2, which compares the electrostatic geometry of TL6P with that of the empirically parametrized[65] seven-point PMM model SWM6. Whereas SWM6 reproduces, just like TL6P, the dielectric constant $\epsilon(T_0,p_0)$ very well, it overestimates, just like TL5P and in contrast to TL6P, α_p by a factor of about 2.6 (we have extracted the SWM6 value $\alpha_p \approx 7.2 \times 10^{-4}$ K^{-1} from the temperature dependence of the density depicted in Figure 5 of ref 65).

4. DISCUSSION

With the mixed computational and empirical optimization of the six-point potentials TL6P and TL6P^{Sk} we finally succeeded with the construction of excellent PMM models for liquid water at T_0 and p_0. The TL6P potentials reproduce, of course, all those liquid phase properties that were targeted by the empirical optimization of the Buckingham potential employed for the modeling of the van der Waals interactions. Beyond that, they accurately predict a series of liquid phase properties, i.e., the self-diffusion constant D_0, viscosity η, isothermal compressibility κ_T, isobaric heat capacity C_p, and dielectric constant ϵ, for which also other PMM models with a related design typically yield reasonable descriptions. But finally and most remarkably they even provide accurate descriptions of the isobaric thermal expansion coefficient α_p, for which PMM models tend to fail.

4.1. Electrostatic Signatures. Moreover, the method of parameter optimization employed by us, which derives the electrostatic signatures of water models almost exclusively from DFT/PMM hybrid calculations, has eventually revealed the microscopic physical reason, why previous PMM models failed in accurately predicting $\alpha_p(T_0,p_0)$. According to the analysis of the quality, by which TL6P and its TL5P and TL4P predecessors can represent the average electrostatic signature of the DFT/PMM reference models (cf. the second line in Table 1 in section 3.1), it is the prominent feature of the TL6P potentials that they reproduce the average static part of the surface potential of a DFT water model embedded in a PMM liquid phase environment substantially better than than TL5P and TL4P and almost as good as an isolated DFT model of the same liquid phase geometry. The latter fact that the isolated DFT model excellently approximates the average static part of the surface potential of a DFT model in a PMM liquid (cf. section 3.1) furthermore explains why the TL6P quadrupole moments are much closer to those of an isolated DFT water molecule than TL5P or TL4P and as close to the experimental

dx.doi.org/10.1021/jp4125765 | J. Phys. Chem. B 2014, 118, 1589−1602

gas phase data[69] as the DFT result for an isolated molecule (cf. Table 3).

Thus, the electrostatic signatures of TL5P, TL4P, SWM6, and related models do not describe the almost constant higher electrostatic moments (quadrupole, octupole, hexadecapole, ...) of liquid phase water molecules with sufficient accuracy. In contrast, DFT and DFT/PMM descriptions and the correspondingly optimized TL6P potentials can cover these higher moments with errors of at most 3% (cf. Tables 1 and 3).

Because the electrostatic fields generated by the higher electrostatic moments are increasingly short ranged, they predominantly shape the local liquid structure in the immediate neighborhood of a given water molecule. Apparently, only models like TL6P, which reproduce the local fields generated by liquid phase water molecules very well, do not grossly overestimate the very small density reduction, which is induced by a temperature decrease near T_0 and is expressed by the small value of $\alpha_p(T_0,p_0)$.

It should be stressed that the high quality, by which DFT and DFT/PMM descriptions can cover the properties of water molecules in the liquid, lies at the heart of the apparent success of the PMM model construction reached with the TL6P potentials. Together with adopting the experimental gas phase values μ^{g}_{exp} and α^{g}_{exp} for the static dipole moment and the polarizability of the model, it was the separate DFT/PMM optimization of the width σ of the Gaussian inducible dipole at the oxygen and of the constant electrostatic signature $\Lambda_{e,6}$ that reduced the parameter space to a sufficiently low dimension such that the unusual electrostatic TL6P geometry depicted in Figure 1 could be determined.

4.2. van der Waals Potentials. In contrast to the electrostatic properties, the empirically optimized Buckingham parameters turned out to be in a range in which also previous potentials can be found. As compared to the cases for TL5P and TL4P, the dispersion attraction of the two TL6P models, for instance, turned out to be reduced by about a factor of 2, moving it close to the dispersion attraction of the well-known nonpolarizable TIP5P and TIP4P/2005 models.[40,82] On the other hand, the resulting differences between the Buckingham potentials of the two TL6P models were quite small despite the fact that different experimental targets[46,48] were used for the optimization of A_1 (cf. Figure S8 of the Supporting Information). In view of the slightly superior performance of TL6P as compared to TL6P^{Sk} concerning most observables [with the exception of $\alpha_p(T_0,p_0)$, cf. Table 4], we decided to consider the latter TL6P variant not any further.

4.3. Transferability Issues. As mentioned above at several locations, the TL6P potential turned out[42] to predict the temperature density profile $n_{exp}(T,p_0)$ of liquid water for $T \in [250, 320]$ K with an unprecedented accuracy as expressed by a root-mean square density deviation of only 0.0005 g/cm^3. In line with the small 0.1×10^{-4} K^{-1} overestimate of $\alpha_p(T_0,p_0)$ by TL6P, this deviation is a likewise small density overestimate, which increases for decreasing temperatures $T < T_0$.

For PMM water models with an inducible dipole and otherwise static partial point charges like TL6P, an overestimated density of the supercooled liquid indicates that the structure forming local electrostatic interactions, which are due to the higher electrostatic multipole moments and reduce the density,[42] do not sufficiently compete with the general attractive forces, which are caused by the dipolar and dispersive interactions. Correspondingly, one expects that such a model underestimates the stability of ice and overestimates that of the liquid entailing an underestimate of the melting temperature T^m (see further below).

Because TL6P^{Sk} performs on the temperature derivative $\alpha_p(T_0,p_0)$ of the density even better than TL6P, it might have been preferential to use this variant model potential in the computationally very costly 20 ns RE-MD simulations[42] on $n_{exp}(T,p_0)$. The answer to the question, however, to what extent the excellent performance of TL6P on $n_{exp}(T,p_0)$ can be further optimized by minor parameter variations must be left to future studies. In the case of TL6P, the observed[26] temperature T^{md}_{exp} = 277.134 K of maximum density was hit by the prediction T^{md}_{TL6P} = 277.055 ± 0.125 K within the limits of a very small statistical uncertainty.[42]

As compared to these structural properties, the TL6P predictions on the temperature profiles of energetic liquid phase properties like the heat of vaporization $\Delta H_{vap}(T,p_0)$ or the heat capacity $C_p(T,p_0)$ turned out to be a little worse.[42] Both quantities showed small but, toward lower temperatures, increasing overestimates of the corresponding experimental findings. As a possible cause the use of partial point charges instead of Gaussian charges (with their locally slightly softer electrostatic potentials) was identified (see the Supporting Information to ref 42).

4.3.1. Melting Temperature. In section S10 of the Supporting Information we have presented and applied a simulation setup for mixtures of ice Ih and liquid water. The setup was designed to measure the average potential energies $E_{pot}(t|T_z)$ of the water molecules as a function of the time t in Np_0T_z simulations covering N = 4424 water molecules at a series of temperatures T_z near the melting temperature T^m_{model} of the respective PMM or MM model. Measuring the temporal slope of $E(t|T_z)$ as a function of temperature, we determined the prediction T^m_{model} as the temperature of vanishing slope.

The simulation setup has been tested using the well-known nonpolarizable model TIP4P/2005 as a reference, which predicts[40,83] a melting temperature of about 250 ± 3 K. Our setup yielded the value $T^m_{TIP4P/2005}$ = 247.5 K, demonstrating that it gives a reasonable estimate and lower bound. In line with the fact that the TL4P predecessor model strongly overestimates[42] the liquid's density $n_{exp}(T^m_{exp},p_0)$ by 0.96% at the experimental melting temperature[73] T^m_{exp} = 273.15 K, our setup locates its melting temperature T^m_{TL4P} way below 250 K, which was the lowest temperature that could be reasonably treated by our PMM setup. This result is similar to that of the related PMM four-point model SWM4/NDP,[31] which overestimates $n_{exp}(T^m_{exp},p_0)$ by as much as 1.32% and underestimates[84] T^m_{exp} by 88 K.

In contrast, and as expected from the very small 0.02% overestimate of $n_{exp}(T^m_{exp},p_0)$ by TL6P, this PMM model predicts T^m_{TL6P} = 262.5 K, thus underestimating T^m_{exp} by at most 11 K. As shown in section S10 of the Supporting Information, the TL6P model apparently predicts T^m_{exp} more accurately than any other PMM model known to us.

4.3.2. Vapor. Applying a novel DFT/PMM technology,[14] the electrostatics of the TLvP models was optimized for liquid phase environments at T_0 and p_0. Therefore, one expects a suboptimal performance on gas phase properties. As an illustration, we compute in section S4 of the Supporting Information the second virial coefficients $B_2(T)$ of the TLvP models (v = 4, 5, 6) for $T \in [300, 450]$. The results are compared with those of other models (including a variant PMM six-point model called TL6P^g, which has been tentatively

dx.doi.org/10.1021/jp4125765 | J. Phys. Chem. B 2014, 118, 1589−1602

optimized by DFT calculations for the gas phase) and with experimental data.[72]

On average, the TL6P predictions on the absolute values $|B_2(T)|$ turned out to be by 9% smaller than the experimental values, TL4P showed underestimates of only 3%, whereas the "gas phase" model TL6P^g was very close to the experimental data closely resembling the behavior of the BK3 model.[36] As expected, TL6P is definitely not an optimal model for water vapor.

4.3.3. Isolated Dimer. Scrutinizing the performance in further environments we present and discuss in section S5 of the Supporting Information the geometries $\mathbf{G}_d$, binding energies $E_{pot,d}$, and total dipole moments μ_d of the isolated TL6P dimers. Here it is shown that the two TL6P models generally furnish reasonable descriptions, which are of quality similar to those provided by their less complex predecessors[8] but cannot compete, e.g., with SWM6, because dimer properties had been optimization targets in this case.[65]

4.3.4. DFT/PMM Settings. As opposed to descriptions of the gas phase, the TL6P water model should be excellently suited for DFT/PMM-MD simulations of solute−solvent systems, which aim, by choosing the solute as the DFT fragment, at the accurate calculation of its condensed phase infrared spectrum (see, e.g., refs 85 and 86 for less accurate DFT/MM-MD examples, which were severely hampered[87] by the low quality of the respective MM solvent model). The high quality, by which the TL6P/DFT and DFT/TL6P dimers could reproduce corresponding experimental data (cf. section S6 and Table S8 in the Supporting Information) are a first indication for the favorable properties of TL6P in DFT/PMM settings. A second indication is the near identity of the TL6P liquid phase dipole distribution with the one calculated for a liquid phase DFT/TL6P ensemble (cf. section 3.3.3 and Supporting Information S8.3).

4.4. Computational Aspects. Issues of computational manageability are discussed in some detail in section S9 of the Supporting Information. Comparing MD simulations of systems comprising either 1500 polarizable TL3P or non-polarizable TIP3P models,[88] one sees that the computer time spent by our PMM-MD program IPHIGENIE[51] increases by a factor of 3.5 upon adding an inducible Gaussian dipole, whose self-consistency iterations have to be brought to convergence. The transition to TL4P implies another factor of 1.3, which then marks the minimum effort required for a reasonably accurate description. A comparatively small factor of about 1.15 eventually characterizes the transition from TL4P to TL6P.

Interestingly, this increase of computer time is almost identical to the time that has to be spent on exchanging the three partial point charges of TL4P by Gaussian charge distributions. The thus obtained model then resembles GCPM[30] and the various[34,36,64,89] BK models. In summary, replacing TL4P by TL6P implies only a small computational overhead while yielding a significant accuracy gain. Compared to the simplistic TIP3P model, the computer time consumed by TL6P is 5.2 times larger, which is, however, compensated by an enormous gain of accuracy.

5. SUMMARY AND OUTLOOK

The mixed computational and empirical approach to the optimization of PMM models for complex liquids such as water, which has been developed in ref 8, relies on a novel DFT/PMM technology[14] and led to the suggestion of the quite reasonable PMM four- and five-point water models TL4P and TL5P, respectively, has now been successfully applied to the development of two corresponding six-point PMM models. The parametrization and a series of test simulations were executed at the single point T_0 = 300 K and p_0 = 1 bar of the thermodynamic state space. Here, the TL6P and TL6P^{Sk} potentials showed a remarkably good performance, which could be identified to follow from their closely resembling but unusual electrostatic signatures.

In particular, the TL6P potentials were shown to accurately predict a series of important liquid phase properties. These properties were, of course, not among the three experimental targets employed for the empirical optimization, which served to tune solely the three van der Waals parameters (A_1, A_2 B). Among the predicted properties were the dielectric constant $\epsilon(T_0,P_0)$, the density, and its negative temperature derivative, the isobaric thermal expansion coefficient $\alpha_p(T_0,p_0)$, of liquid water.

The latter finding had nourished the hope that one of the TL6P potentials might also correctly predict the density $n(T,p_0)$ over a sufficiently wide range of other temperatures, which extends beyond the temperature T^{md} = 277 K of maximum density to lower values. Therefore, this issue has been addressed in a thorough follow-up study,[42] which applied 20 ns RE-MD simulations spanning the temperature range [250, 320] K to the TLνP models with ν = 4, 5, 6. In line with the quite accurate value of $\alpha_p(T_0,p_0)$ also the density profile $n(T,p_0)$ and the temperature T^{md} of maximum density were predicted by TL6P with a hitherto unprecedented accuracy (as compared to TL4P, TL5P, and other PMM models).

This success of TL6P and the corresponding failure of its less complex predecessors TL4P and TL5P, which had been parametrized by the same procedure, led to the conclusion[42] that the TL6P distribution of partial point charges depicted in Figure 2 provides a minimal cartoon for those aspects of the continuous charge distributions in real water molecules, which are responsible for the density anomaly.

In line with the very slight (0.02%) overestimate of the density at the experimental melting temperature T^m_{exp}, TL6P underestimates this temperature only a little by at most 11 deg. As is shown in section S10 of the Supporting Information, no other PMM water model known to us performs with a comparable accuracy in predicting T^m_{exp}. The predecessor model TL4P, in particular, assigns the melting point of ice to temperatures way below 250 K. This finding strengthens the conclusions of ref 42 concerning the enormous jump of modeling accuracy, which can be achieved in a DFT/PMM parametrization setting through the transition from four- and five- to six-point PMM models.

The polarizable six-point model TL6P, which was successfully developed in this contribution by applying DFT/PMM hybrid techniques, does not represent a final word but should be understood, instead, as a new beginning. Here, one first should check whether and, if so, to what extent the variant model TL6P^{Sk}, which was also described above, can further reduce the already small TL6P overestimate of $n_{exp}(T,p_0)$ at temperatures $T < T_0$ and the small TL6P underestimate of the melting temperature T^m_{exp}

Next, the influence of the DFT functional, which is employed in the DFT/PMM parametrization procedure, on the resulting bulk phase properties should be studied. We employed, for reasons of reduced computational cost, the rather simple gradient-corrected Becke−Perdew functional.[58,59] Including Hartree−Fock exchange[90] into the popular Becke−Lee−

dx.doi.org/10.1021/jp4125765 | J. Phys. Chem. B 2014, 118, 1589−1602

Yang−Parr approach,[58,91] that is, using the so-called B3LYP functional promises more accurate descriptions (see, e.g., ref 87). Hence, six-point PMM models of the TL6P-type should be calculated by the B3LYP/PMM technique and evaluated. All these six-point PMM models will have the same computational complexity as TL6P.

Finally, and at the expense of an enhanced computational cost, one could try to construct PMM six-point models, in which the five static partial point charges are modeled as Gaussian distributions. This should be done if (and only if) the thus far constructed PMM six-point models still feature at low temperatures a somewhat overestimated heat of vaporization.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information provides on 27 pages in ten sections a total of nine figures (S4−S12), eight tables (S5−S12), and various pieces of text explaining and documenting several issues. First we specify the computation of the long-range electrostatics in some detail, and then we sketch the correlations between DFT/PMM and PMM dipoles, which are used for the optimization of the widths of the Gaussian inducible dipoles. Next we give the parameters of the two TL6P models with all relevant digits. After a discussion of the second virial coefficient and the PMM and DFT/PMM dimer properties, we study all those properties, which were targets of the empirical optimization. Furthermore, we provide TL6P predictions on bulk phase properties and associated statistical errors. After discussing computational issues, we finally describe MD simulations on mixtures of ice Ih and liquid water. This material is available free of charge via the Internet at http://pubs.acs.org/.

■ AUTHOR INFORMATION

Corresponding Author

*P. Tavan: e-mail, tavan@physik.uni-muenchen.de.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work has been supported by the Deutsche Forschungsgemeinschaft (SFB 749/C4).

■ REFERENCES

(1) Ball, P. *H_2O: A Biography of Water*; Phoenix, Orion Books Ltd: London, U.K., 2000.

(2) Guillot, B. A Reappraisal of What We Have Learnt During Three Decades of Computer Simulations on Water. *J. Mol. Liq.* **2002**, *101*, 219−260.

(3) Lopes, P.; Roux, B.; MacKerell, A. Molecular Modeling and Dynamics Studies with Explicit Inclusion of Electronic Polarizability: Theory and Applications. *Theor. Chem. Acc.* **2009**, *124*, 11−28.

(4) Dünweg, B.; Kremer, K. Molecular Dynamics Simulation of a Polymer Chain in Solution. *J. Chem. Phys.* **1993**, *99*, 6983−6997.

(5) Salacuse, J. J.; Denton, A. R.; Egelstaff, P. A. Finite-Size Effects in Molecular Dynamics Simulations: Static Structure Factor and Compressibility. I. Theoretical Method. *Phys. Rev. E* **1996**, *53*, 2382−2389.

(6) Rapaport, D. C. *The Art of Molecular Dynamics Simulation*; Cambridge University Press: Cambridge, U.K., 2004.

(7) Tazi, S.; Botan, A.; Salanne, M.; Marry, V.; Turq, P.; Rotenberg, B. Diffusion Coefficient and Shear Viscosity of Rigid Water Models. *J. Phys.: Condens. Matter* **2012**, *24*, 284117.

(8) Tröster, P.; Lorenzen, K.; Schwörer, M.; Tavan, P. Polarizable Water Models from Mixed Computational and Empirical Optimization. *J. Phys. Chem. B* **2013**, *117*, 9486−9500.

(9) Allen, M. P.; Tildesley, D. *Computer Simulations of Liquids*; Clarendon: Oxford, U.K., 1987.

(10) Vega, C.; Abascal, J. Simulating Water with Rigid Non-Polarizable Models: A General Perspective. *Phys. Chem. Chem. Phys.* **2011**, *13*, 19663−19688.

(11) Guidon, M.; Hutter, J.; VandeVondele, J. Robust Periodic Hartree-Fock Exchange for Large-Scale Simulations Using Gaussian Basis Sets. *J. Chem. Theory Comput.* **2009**, *5*, 3010−3021.

(12) Guidon, M.; Hutter, J.; VandeVondele, J. Auxiliary Density Matrix Methods for Hartree-Fock Exchange Calculations. *J. Chem. Theory Comput.* **2010**, *6*, 2348−2364.

(13) Tavan, P.; Carstens, H.; Mathias, G. Molecular Dynamics Simulations of Proteins and Peptides: Problems, Achievements, and Perspectives. In *Protein Folding Handbook*; Buchner, J., Kiefhaber, T., Eds.; Wiley-VCH: Weinheim, 2005; Vol. *1*, pp 1170−1195.

(14) Schwörer, M.; Breitenfeld, B.; Tröster, P.; Lorenzen, K.; Tavan, P.; Mathias, G. Coupling DFT to Polarizable Force Fields for Efficient and Accurate Hamiltonian Molecular Dynamics Simulations. *J. Chem. Phys.* **2013**, *138*, 244103.

(15) The letters "TL" in the names TL*v*P of or water models are the initials of the first two authors of ref 8 the number $v \in \{3, 4, 5, 6\}$ counts the points of force action within the respective model, and "P" stands for "point".

(16) Clough, S.; Beers, Y.; Klein, G.; Rothman, L. Dipole Moment of Water from Stark Measurements of H_2O, HDO, and D_2O. *J. Chem. Phys.* **1973**, *59*, 2254−2259.

(17) Murphy, W. F. The Rayleigh Depolarization Ratio and Rotational Raman Spectrum of Water Vapor and the Polarizability Components for the Water Molecule. *J. Chem. Phys.* **1977**, *67*, 5877−5882.

(18) Buckingham, R. A.; Corner, J. Tables of Second Virial and Low-Pressure Joule-Thomson Coefficients for Intermolecular Potentials with Exponential Repulsion. *Proc. R. Soc. London, Ser. A* **1947**, *189*, 118−129.

(19) Ichikawa, K.; Kameda, Y.; Yamaguchi, T.; Wakita, H.; Misawa, M. Neutron-Diffraction Investigation of the Intramolecular Structure of a Water Molecule in the Liquid Phase at High Temperatures. *Mol. Phys.* **1991**, *73*, 79−86.

(20) Thiessen, W. E.; Narten, A. H. Neutron Diffraction Study of Light and Heavy Water Mixtures at 25 °C. *J. Chem. Phys.* **1982**, *77*, 2656−2662.

(21) Wallqvist, A.; Berne, B. Effective Potentials for Liquid Water Using Polarizable and Nonpolarizable Models. *J. Phys. Chem.* **1993**, *97*, 13841−13851.

(22) de Pablo, J. J.; Prausnitz, J. M.; Strauch, H. J.; Cummings, P. T. Molecular Simulation of Water along the Liquid−Vapor Coexistence Curve from 25°C to the Critical Point. *J. Chem. Phys.* **1990**, *93*, 7355−7359.

(23) Kiyohara, K.; Gubbins, K. E.; Panagiotopoulos, A. Z. Phase Coexistence Properties of Polarizable Water Models. *Mol. Phys.* **1998**, *94*, 803−808.

(24) Straatsma, T. P.; McCammon, J. A. Molecular Dynamics Simulations with Interaction Potentials Including Polarization Development of a Noniterative Method and Application to Water. *Mol. Simul.* **1990**, *5*, 181−192.

(25) Yu, H.; Hansson, T.; Van Gunsteren, W. Development of a Simple, Self-Consistent Polarizable Model for Liquid Water. *J. Chem. Phys.* **2003**, *118*, 221−234.

(26) Kell, G. Precise Representation of Volume Properties of Water at One Atmosphere. *J. Chem. Eng. Data* **1967**, *12*, 66−69.

(27) Kaatze, U. Complex Permittivity of Water as a Function of Frequency and Temperature. *J. Chem. Eng. Data* **1989**, *34*, 371−374.

(28) Yu, H.; van Gunsteren, W. Charge-On-Spring Polarizable Water Models Revisited: From Water Clusters to Liquid Water to Ice. *J. Chem. Phys.* **2004**, *121*, 9549−9564.

(29) Lamoureux, G.; MacKerell, A., Jr.; Roux, B. A Simple Polarizable Model of Water Based on Classical Drude Oscillators. *J. Chem. Phys.* **2003**, *119*, 5185−5197.
(30) Paricaud, P.; Předota, M.; Chialvo, A.; Cummings, P. From Dimer to Condensed Phases at Extreme Conditions: Accurate Predictions of the Properties of Water by a Gaussian Charge Polarizable Model. *J. Chem. Phys.* **2005**, *122*, 244511.
(31) Lamoureux, G.; Harder, E.; Vorobyov, I.; Roux, B.; MacKerell, A. A Polarizable Model of Water for Molecular Dynamics Simulations of Biomolecules. *Chem. Phys. Lett.* **2006**, *418*, 245−249.
(32) Kunz, A.; van Gunsteren, W. Development of a Nonlinear Classical Polarization Model for Liquid Water and Aqueous Solutions: COS/D. *J. Phys. Chem. A* **2009**, *113*, 11570−11579.
(33) Viererblová, L.; Kolafa, J. A Classical Polarizable Model for Simulations of Water and Ice. *Phys. Chem. Chem. Phys.* **2011**, *13*, 19925−19935.
(34) Kiss, P. T.; Baranyai, A. Density Maximum and Polarizable Models of Water. *J. Chem. Phys.* **2012**, *137*, 084506.
(35) Lennard-Jones, J. E. Cohesion. *Proc. Phys. Soc.* **1931**, *43*, 461−482.
(36) Kiss, P. T.; Baranyai, A. A Systematic Development of a Polarizable Potential of Water. *J. Chem. Phys.* **2013**, *138*, 204507.
(37) Data extracted at T_0 from Figure 8 in ref 36.
(38) Effective mean-field type potentials, which try to compensate the fully (e.g., TP4P/2005[40]) or partially (e.g., iAMOEBA[39]) neglected polarizability by an empirically optimal but nevertheless nonphysical choice of the model parameters, may be of practical use in simulations but are almost certainly of limited value for predictions on conditions, which are not covered by the reference data (e.g., water in proteins, electrolytes, etc.). PMM models, in contrast, which are good enough to catch the essential physics of a water molecule, should be transferable to all kinds of environments as long as the assumption of linear response is valid. Furthermore they open the chance to identify the microscopic causes for the macroscopic properties of water.
(39) Wang, L.-P.; Head-Gordon, T.; Ponder, J. W.; Ren, P.; Chodera, J. D.; Eastman, P. K.; Martinez, T. J.; Pande, V. S. Systematic Improvement of a Classical Molecular Model of Water. *J. Phys. Chem. B* **2013**, *117*, 9956−9972.
(40) Abascal, J.; Vega, C. A General Purpose Model for the Condensed Phases of Water: TIP4P/2005. *J. Chem. Phys.* **2005**, *123*, 234505.
(41) Alejandre, J.; Chapela, G. A.; Saint-Martin, H.; Mendoza, N. A Non-Polarizable Model of Water that Yields the Dielectric Constant and the Density Anomalies of the Liquid: TIP4Q. *Phys. Chem. Chem. Phys.* **2011**, *13*, 19728−19740.
(42) Tröster, P.; Tavan, P. The Microscopic Physical Cause for the Density Maximum of Liquid Water. *J. Phys. Chem. Lett.* **2014**, *5*, 138−142.
(43) Schropp, B.; Tavan, P. Flexibility Does Not Change the Polarizability of Water Molecules in the Liquid. *J. Phys. Chem. B* **2010**, *114*, 2051−2057.
(44) Berweger, C.; van Gunsteren, W.; Müller-Plathe, F. Force Field Parametrization by Weak Coupling. Re-Engineering SPC Water. *Chem. Phys. Lett.* **1995**, *232*, 429−436.
(45) Hura, G.; Sorenson, J.; Glaeser, R.; Head-Gordon, T. A High-Quality X-Ray Scattering Experiment on Liquid Water at Ambient Conditions. *J. Chem. Phys.* **2000**, *113*, 9140−9148.
(46) Soper, A. K. The Radial Distribution Functions of Water and Ice from 220 to 673 K and at Pressures up to 400 MPa. *Chem. Phys.* **2000**, *258*, 121−137.
(47) Soper, A. K. Joint Structure Refinement of X-Ray and Neutron Diffraction Data on Disordered Materials: Application to Liquid Water. *J. Phys.: Condens. Matter* **2007**, *19*, 335206.
(48) Skinner, L. B.; Huang, C.; Schlesinger, D.; Pettersson, L. G.; Nilsson, A.; Benmore, C. J. Benchmark Oxygen-Oxygen Pair-Distribution Function of Ambient Water from X-Ray Diffraction Measurements with a Wide Q-Range. *J. Chem. Phys.* **2013**, *138*, 074506.
(49) Jancso, G.; Van Hook, W. A. Condensed Phase Isotope Effects. *Chem. Rev.* **1974**, *74*, 689−750.
(50) Wagner, W.; Pruß, A. The IAPWS Formulation 1995 for the Thermodynamic Properties of Ordinary Water Substance for General and Scientific Use. *J. Phys. Chem. Ref. Data* **2002**, *31*, 387.
(51) Lorenzen, K.; Schwörer, M.; Tröster, P.; Mates, S.; Tavan, P. Optimizing the Accuracy and Efficiency of Fast Hierarchical Multipole Expansions for MD Simulations. *J. Chem. Theory Comput.* **2012**, *8*, 3628−3636.
(52) Niedermeier, C.; Tavan, P. A Structure Adapted Multipole Method for Electrostatic Interactions in Protein Dynamics. *J. Chem. Phys.* **1994**, *101*, 734−748.
(53) Niedermeier, C.; Tavan, P. Fast Version of the Structure Adapted Multipole Method: Efficient Calculation of Electrostatic Forces in Protein Dynamics. *Mol. Simul.* **1996**, *17*, 57−66.
(54) Eichinger, M.; Grubmüller, H.; Heller, H.; Tavan, P. FAMUSAMM: An Algorithm for Rapid Evaluation of Electrostatic Interactions in Molecular Dynamics Simulations. *J. Comput. Chem.* **1997**, *18*, 1729−1749.
(55) Mathias, G.; Egwolf, B.; Nonella, M.; Tavan, P. A Fast Multipole Method Combined with a Reaction Field for Long-Range Electrostatics in Molecular Dynamics Simulations: The Effects of Truncation on the Properties of Water. *J. Chem. Phys.* **2003**, *118*, 10847−10860.
(56) Eichinger, M.; Tavan, P.; Hutter, J.; Parrinello, M. A Hybrid Method for Solutes in Complex Solvents: Density Functional Theory Combined with Empirical Force Fields. *J. Chem. Phys.* **1999**, *110*, 10452−10467.
(57) CPMD V3.9, Copyright IBM Corp 1990−2008, Copyright MPI für Festkörperforschung Stuttgart 1997-2001, see also www.cpmd.org.
(58) Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A* **1988**, *38*, 3098−3100.
(59) Perdew, J. P. Density-Functional Approximation for the Correlation Energy of the Inhomogeneous Electron Gas. *Phys. Rev. B* **1986**, *33*, 8822−8824.
(60) Troullier, N.; Martins, J. Efficient Pseudopotentials for Plane-Wave Calculations. *Phys. Rev. B* **1991**, *43*, 1993−2006.
(61) Bussi, G.; Parrinello, M. Stochastic Thermostats: Comparison of Local and Global Schemes. *Comput. Phys. Commun.* **2008**, *179*, 26−29.
(62) Lingenheil, M.; Denschlag, R.; Reichold, R.; Tavan, P. The Hot-Solvent/Cold-Solute Problem Revisited. *J. Chem. Theory Comput.* **2008**, *4*, 1293−1306.
(63) 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. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(64) Baranyai, A.; Kiss, P. A Transferable Classical Potential for the Water Molecule. *J. Chem. Phys.* **2010**, *133*, 144109.
(65) Yu, W.; Lopes, P. E. M.; Roux, B.; MacKerell, A. D.; Six-Site, J. Polarizable Model of Water Based on the Classical Drude Oscillator. *J. Chem. Phys.* **2013**, *138*, 034508.
(66) Nada, H.; van der Eerden, J. An Intermolecular Potential Model for the Simulation of Ice and Water Near the Melting Point: A Six-Site Model of H_2O. *J. Chem. Phys.* **2003**, *118*, 7401−7413.
(67) The electrostatic signature of the early nonpolarizable six-point model suggested in ref 66 strongly differs from that of TL6P, because it is an effective[38] potential. For instance, the angle φ_{LOL} between the very small lone-pair charges (−0.044 e) is 111° and the distance r_{OL} is 0.89 Å. As a result, its quadrupole moment differs on average by 18% from the experimental value. Its poor performance on liquid phase properties like the isothermal compressibility or the heat capacity, where it deviates by 20−35% from the experimental evidence, are additional reasons, why we do not discuss this attempt any further.
(68) Wang, L.-P.; Chen, J.; Van Voorhis, T. Systematic Parametrization of Polarizable Force Fields from Quantum Chemistry Data. *J. Chem. Theory Comput.* **2012**, *9*, 452−460.
(69) Verhoeven, J.; Dymanus, A. Magnetic Properties and Molecular Quadrupole Tensor of the Water Molecule by Beam-Maser Zeeman Spectroscopy. *J. Chem. Phys.* **1970**, *52*, 3222−3233.

(70) Dyke, T. R.; Mack, K. M.; Muenter, J. S. The Structure of Water Dimer from Molecular Beam Electric Resonance Spectroscopy. *J. Chem. Phys.* **1977**, *66*, 498−510.
(71) Odutola, J. A.; Dyke, T. R. Partially Deuterated Water Dimers: Microwave Spectra and Structure. *J. Chem. Phys.* **1980**, *72*, 5062−5070.
(72) Hill, P. G.; MacMillan, C. R. Virial Equations for Light and Heavy Water. *Ind. Eng. Chem. Res.* **1988**, *27*, 874−882.
(73) Feistela, R.; Wagner, W. A New Equation of State for H_2O Ice Ih. *J. Phys. Chem. Ref. Data* **2006**, *35*, 1021−1047.
(74) Krynicki, K.; Green, C.; Sawyer, D. Pressure and Temperature Dependence of Self-Diffusion in Water. *Faraday Discuss. Chem. Soc.* **1978**, *66*, 199−208.
(75) Harris, K.; Woolf, L. Temperature and Volume Dependence of the Viscosity of Water and Heavy Water at Low Temperatures. *J. Chem. Eng. Data* **2004**, *49*, 1064−1069.
(76) Weast, R. C. *Handbook of Chemistry and Physics*; CRC: Boca Raton, FL, 1983.
(77) Motakabbir, K. A.; Berkowitz, M. Isothermal Compressibility of SPC/E Water. *J. Phys. Chem.* **1990**, *94*, 8359−8362.
(78) Sprik, M. Hydrogen Bonding and the Static Dielectric Constant in Liquid Water. *J. Chem. Phys.* **1991**, *95*, 6762−6769.
(79) Neumann, M. Dipole Moment Fluctuation Formulas in Computer Simulations of Polar Systems. *Mol. Phys.* **1983**, *50*, 841−858.
(80) Smith, P. E.; van Gunsteren, W. F. Consistent Dielectric Properties of the Simple Point Charge and Extended Simple Point Charge Water Models at 277 and 300 K. *J. Chem. Phys.* **1994**, *100*, 3169−3174.
(81) Tironi, I. G.; Van Gunsteren, W. F. A Molecular Dynamics Simulation Study of Chloroform. *Mol. Phys.* **1994**, *83*, 381−403.
(82) Mahoney, M.; Jorgensen, W. A Five-Site Model for Liquid Water and the Reproduction of the Density Anomaly by Rigid, Nonpolarizable Potential Functions. *J. Chem. Phys.* **2000**, *112*, 8910−8922.
(83) Fernández, R. G.; Abascal, J. L.; Vega, C. The Melting Point of Ice I for Common Water Models Calculated from Direct Coexistence of the Solid-Liquid Interface. *J. Chem. Phys.* **2006**, *124*, 144506.
(84) Gladich, I.; Roeselová, M. Comparison of Selected Polarizable and Nonpolarizable Water Models in Molecular Dynamics Simulations of Ice Ih. *Phys. Chem. Chem. Phys.* **2012**, *14*, 11371−11385.
(85) Klähn, M.; Mathias, G.; Kötting, C.; Schlitter, J.; Nonella, M.; Gerwert, K.; Tavan, P. IR Spectra of Phosphate Ions in Aqueous Solution: Predictions of a DFT/MM Approach Compared with Observations. *J. Phys. Chem. A* **2004**, *108*, 6186−6194.
(86) Schultheis, V.; Reichold, R.; Schropp, B.; Tavan, P. A Polarizable Force Field for Computing the Infrared Spectra of the Polypeptide Backbone. *J. Phys. Chem. B* **2008**, *112*, 12217−12230.
(87) VandeVondele, J.; Tröster, P.; Tavan, P.; Mathias, G. Vibrational Spectra of Phosphate Ions in Aqueous Solution Probed by First-Principles Molecular Dynamics. *J. Phys. Chem. A* **2012**, *116*, 2466−2474.
(88) Jorgensen, W.; Chandrasekhar, J.; Madura, J.; Impey, R.; Klein, M. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Phys. Chem.* **1983**, *79*, 926−935.
(89) Baranyai, A.; Kiss, P. Polarizable Model of Water with Field-Dependent Polarization. *J. Chem. Phys.* **2011**, *135*, 234110.
(90) Becke, A. D. A New Mixing of Hartree-Fock and Local Density-Functional Theories. *J. Chem. Phys.* **1993**, *98*, 1372−1377.
(91) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37*, 785−789.

Supporting information for: Polarizable Six-Point Water Models from Computational and Empirical Eptimization

Philipp Tröster, Konstantin Lorenzen, and Paul Tavan*

Lehrstuhl für Biomolekulare Optik, Fakultät für Physik,

Ludwig-Maximilians-Universität München,

Oettingenstr. 67, D-80538 München, Germany

E-mail: tavan@physik.uni-muenchen.de

*To whom correspondence should be addressed

S1 Computation of Long-Range Electrostatics

The long-range electrostatic interactions were calculated with the fourth-order Structure-Adapted fast Multipole Method called SAMM_4, whose mathematics and favorable properties are explained in Ref. S1.

Like its predecessor method SAMM,[S2–S5] also SAMM_4 is based on a hierarchical decomposition of a simulation system into nested clusters of decreasing size, whose lowest cluster level $l = 0$ is formed by predefined chemical motifs, the so-called structural units (SUs), covering three to six atoms. In aqueous systems the SUs are the water molecules. The cluster hierarchy is formed by adaptive neural clustering algorithms,[S6,S7] which combine the water molecules, as specified by their centers of geometry, into optimally compact clusters at the cluster levels $l = 1, 2, \ldots, l_{\max}$ each containing about 4^l water molecules. The clustering is adaptively refined every 256 integration steps using the previous clustering results as starting values.

For a cubic periodic boundary system of side-length L the electrostatics is explicitly computed by SAMM_4 up to the distance $R_{\mathrm{MIC}} = L/2$ dictated by the minimum image convention (MIC).[S8] Beyond R_{MIC} a dielectric continuum ϵ_{RF} is assumed, whose contributions to the electrostatic forces and energies are calculated using the reaction field correction (RF) developed in Ref. S5. The nested cluster hierarchy is employed for the efficient check[S5] of the MIC and for the likewise efficient[S3,S4] top-down set-up of interaction lists, which steer the SAMM_4 computation of the long-range electrostatic interactions at each cluster level l.

More specifically, the small water system, which contained $N_{\mathrm{s}} = 728$ TL6P models in a periodic box of side-length $L_{\mathrm{s}} = 27.959$ Å, was hierarchically decomposed into 44 large clusters at level $l_{\max} = 2$ and 176 sub-clusters at level $l = 1$ for checking the MIC and for the set-up of the interaction lists. Because of the smallness of the system, the top-level of the electrostatics calculation was the level $l = 0$ of the SUs. The electrostatic interactions of SUs with a center-center distance larger than 10 Å and smaller than the boundary to the continuum at $R_{\mathrm{MIC}} = 13.979$ Å were thus calculated by fourth order multipole and

Taylor expansions.[S1] Less distant SU's were resolved into the constituent atoms and the electrostatic interactions were calculated by the usual Coulomb expressions. Note that the cluster structure of the medium sized system resembles that of the small system if all cluster numbers are multiplied by two.

The polarizable degrees of freedom were modeled by Gaussian dipoles of widths 0.8 Å at the oxygen atoms. For distances larger than 8 Å the potential [see e.g. Eq. (18) in Ref. S9 for an explicit expression] of such a distribution is at double-precision numerically indistinguishable from the potential of a point-dipole, which is why we took, here, the latter instead of the former.

For distances beyond 10 Å a separate SAMM$_4$/RF hierarchy was generated for the induced dipoles (using the multipole moment formulas for dipole distributions given in the Appendix to Ref. S9). The self-consistency iteration of the induced dipoles was stopped as soon as the changes were all below 5×10^{-5} D. The dipole iteration was accelerated using polynomial extrapolations[S10,S11] for the initial values combined with direct inversions in the iterative subspace[S12,S13] (DIIS). In both cases the history lengths covered four preceding steps.

The large water system with its $N_\mathrm{l} = 5300$ TL6P models in a periodic box of side-length $L_\mathrm{l} = 54.187$ Å featured $l_\mathrm{max} = 3$ cluster levels for interaction list generation and MIC checks. The top-level l_max of the clustering comprised 80 large clusters each containing about 64 molecules. The top-level of the SAMM$_4$ electrostatics computation was $l = 1$. It contained 1280 clusters each comprising about 4 molecules. At this level the electrostatics was explicitly calculated by SAMM$_4$ in the distance range between 15.9 Å and $R_\mathrm{MIC} = 27.093$ Å, for larger distances the implicit RF correction[S5] was applied. Clusters at level $l = 1$ with center to center distances smaller than 15.9 Å were decomposed into their constituent SU's, whose interactions were calculated by SAMM$_4$, if their distances were larger than 10 Å. Otherwise they were decomposed into their constituent atoms and the electrostatics was calculated at this level of atom-atom interactions. Like for the small system, the sketched SAMM$_4$/RF

scheme was separately applied to the polarizable and non-polarizable degrees of freedom.

Note here that a six-point model is computationally as expensive as a three-point model as soon as the interactions are calculated by $SAMM_4$, which is currently the case for distances larger than 10 Å (cf. also Section S9). In the forthcoming extension (Lorenzen, ongoing dissertation) of $SAMM_4$, which integrates the attractive part of the van der Waals interaction into the fast multipole scheme, it will be possible to confine the atom-atom computations to distances smaller than about 7 Å without loss of accuracy and with a considerably enhanced efficiency. Currently the van der Waals interactions are cutoff at 10 Å and a continuum correction[S8] is applied to the more long-range contributions to the energy. Here, the van der Waals cutoff cannot be set to smaller values without introducing serious artifacts and, therefore, currently also the electrostatics has to be computed as an atom-atom interaction up to this distance although the high accuracy of $SAMM_4$ would enable the transition toward multipole descriptions already at much smaller distances (Lorenzen, ongoing dissertation).

S2 Correlations of DFT/PMM and PMM Dipoles

Figure S4 shows the correlations between the absolute values $\mu_{\mathrm{DFT/PMM}}(s)$ and $\mu_\sigma(s)$ of the DFT/PMM and PMM dipole moments, respectively, which represent, up to the constant offset $\mu_{\mathrm{DFT}} = 1.79$ D, the optimal correlations between the absolute values $\mu^{\mathrm{i}}_{\mathrm{DFT/PMM}}(s)$ and

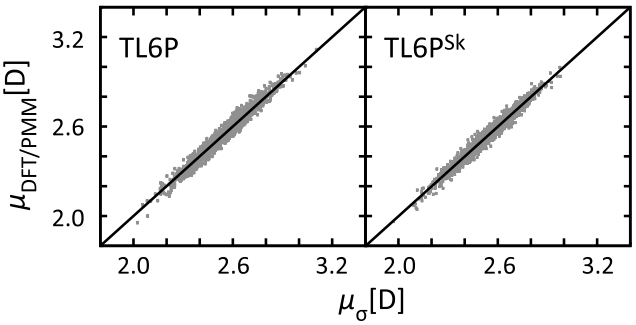


Figure S4: Correlations between the dipoles $\mu_{\mathrm{DFT/PMM}}(s)$ of the DFT fragments and the corresponding dipoles $\mu_\sigma(s)$ of the PMM test molecules for all $s \in \mathcal{S}$.

$\mu^{\mathrm{i}}_{\sigma}(s)$ of the induced dipole moments defined by Eqs. (4) and (5), respectively, and were obtained for the Gaussian widths of 0.806 Å (TL6P) and 0.802 Å (TL6P$^{\mathrm{Sk}}$).

Figure S4 demonstrates that the correlations between the DFT/PMM references and their PMM models are excellent at the optimal widths of the induced Gaussian dipoles for both models. These widths differ by only 0.5 %.

S3 Parameter List

Table S5 provides all relevant digits for the parameters λ of the two TL6P models. The parameters are discussed in Section 3.2.

The Buckingham potential specified by Eq. (1) can be fitted in the range of [2.5, 6.0] Å to a Lennard-Jones Potential $A/r^{12} - B/r^6$ with the same dispersion parameter B and a fitted repulsive parameter A. These parameters can then be converted to the parameters ϵ_{LJ} and σ_{LJ} characterizing the alternative form $4\epsilon_{\mathrm{LJ}}[(\sigma_{\mathrm{LJ}}/r)^{12} - (\sigma_{\mathrm{LJ}}/r)^6]$ of the LJ potential.

Table S5: Parameters $\boldsymbol{\lambda}$ of the TL6P Models.

λ	unit	TL6P	TL6P$^{\mathrm{Sk}}$
α	Å^3	1.470	1.470
σ	Å	0.806	0.802
q_{H}	e	0.507024	0.505580
q_{M}	e	−0.569800	−0.571100
q_{L}	e	−0.222124	−0.220530
l_{OM}	Å	0.3833	0.3860
l_{OL}	Å	0.4601	0.4619
l_{OH}	Å	0.968	0.968
φ_{HOH}	deg	105.3	105.3
φ_{LOL}	deg	175	173
A_1	10^3 kcal/mol	291	304
A_2	Å^{-1}	4.0202	4.0415
B	Å^6 kcal/mol	663	629
ϵ_{LJ}	kcal/mol	0.1378	0.1255
σ_{LJ}	Å	3.2610	3.283

These parameters are listed at the bottom of the table and should be used, if one wants to employ TL6P-like potentials in a program, which cannot deal with Buckingham potentials.

S4 Second Virial Coefficients

The TLνP models were designed for the liquid phase as is witnessed by (i) the choice of the liquid phase instead of the gas phase molecular geometry and (ii) by the optimization of the electrostatic parameters listed in Table 2 through liquid phase DFT/PMM calculations. Therefore one should not expect that these models yield highly accurate descriptions of gas phase properties such as the second virial coefficient $B_2(T)$ or the binding energy and geometry of the isolated dimer.

To give an impression how well these models generalize to the gas phase we have, nevertheless, calculated $B_2(T)$ for TL4P, TL5P, and TL6P in the temperature range from 300 K to 450 K, for which measurements are available.[S14] $B_2(T)$ was calculated by largely following the procedures described in Ref. S15 [note that Eq. (19) in Ref. S15 contains two mistakes whereas the corresponding Eq. (2) in Ref. S16 is correct]. The range of the Monte-Carlo integration was modified. Instead of stochastically integrating from 2 Å to 10 Å we moved the upper limit to 15 Å and employed 260 instead of 100 million random configurations. Beyond 15 Å we employed the analytical dipole-dipole approximation given in Ref. S17 applying the zero-field dipole moment of 1.855 D characteristic for the TLνP models with $\nu \geq 4$.

According to Figure S5 the TL4P model reproduces the experimental function $B_2(T)$ with an average deviation of 3 % better than TL6P, which deviates by 9 %. The performance of TL5P is similar to that of TL6P (data not shown). Whereas the empirical model BK3,[S15] which was optimized taking an important gas phase property (the dimer binding energy) as one of the targets performs better than the TLνP liquid phase models, another empirical model called BKd3,[S18] whose parameters were exclusively fitted to liquid phase properties, misses $B_2(T)$ toward lower temperatures.

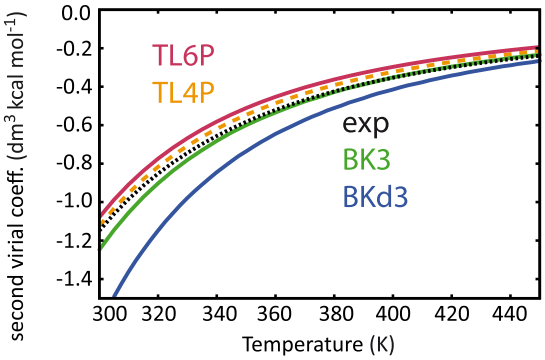


Figure S5: Second virial coefficients $B_2(T)$ of the TL6P (red) and TL4P (orange) PMM models compared with experimental data (black, dotted) and predictions[S15] of the recent empirical PMM models BK3 (green) and BKd3 (blue).

With the aim of checking whether a gas-phase six-point model, which features the experimental gas phase geometry and whose static partial charges optimally fit the surface potential of an isolated DFT water model at this geometry, reproduces $B_2(T)$ better than TL6P, we have tentatively calculated such a model (adopting the parameters modeling the van der Waals interactions, the static dipole moment, and the polarizability from TL6P instead of designing an optimization procedure for the van der Waals parameters, which is adapted to the gas phase).

Table S6: Electrostatic Signatures of TL6P$^{\mathrm{g}}$ and TL6P.

	q_{H}/e	q_{M}/e	q_{L}/e	l_{OM}/Å	l_{OL}/Å	l_{OH}/Å	φ_{HOH}/deg	φ_{LOL}/deg
TL6P$^{\mathrm{g}}$	0.524	−0.594	−0.227	0.389	0.440	0.957	104.5	178
TL6P	0.507	−0.570	−0.222	0.383	0.460	0.968	105.3	175

Table S6 compares the parameters of this gas-phase six-point model called TL6P$^{\mathrm{g}}$ with those of the liquid phase model TL6P. The changes are generally small.

Figure S6 shows that the slight changes of the electrostatic signature moves the TL6P$^{\mathrm{g}}$ prediction (magenta, dashed) on $B_2(T)$ closer to the experimental curve (black, dotted). As a result the predictions of TL6P$^{\mathrm{g}}$ and of the empirically optimized model BK3[S15] (green) are nearly identical. The improved performance TL6P$^{\mathrm{g}}$ supports the conjecture voiced above that the slightly sub-optimal performance of TL6P (red) on $B_2(T)$ is caused by its liquid

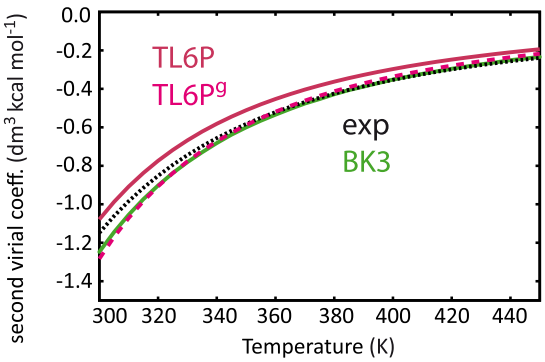


Figure S6: Second virial coefficients $B_2(T)$ of the TL6P (red) and TL6P$^{\mathrm{g}}$ (magenta, dashed) PMM models compared with experimental data (black, dotted) and predictions[S15] of the recent empirical PMM model BK3 (green).

phase optimization.

S5 Dimer Properties of the TL6P Models

The geometry $\mathbf{G}_{\mathrm{d}}$, binding energy E_{pot}, and total dipole moment μ_{d} of the water dimer are experimentally well known.[S19–S21] These data are listed in the top row of Table S7, whereas the second row provides the associated experimental uncertainties σ_{exp}. Furthermore, the table compares these data with the TL6P and TL6P$^{\mathrm{Sk}}$ predictions, which are almost identical. As references, these dimer properties are also listed for the predecessor[S22] TLνP models ($\nu = 4, 5$) and for the recent empirically parameterized PMM six-point model SWM6.[S23]

Table S7: Geometries $\mathbf{G}_{\mathrm{d}}$, Binding Energies E_{pot}, and Dipole Moments μ_{d} of the Water Dimer as Given by Experimental Data[S19–S21] and Described by Various PMM Models.

	d_{OO}/Å	β_1/deg	β_2/deg	$E_{\mathrm{pot,\,d}}/\frac{\mathrm{kcal}}{\mathrm{mol}}$	μ_{d}/D
exp[S19–S21]	2.98	51	123	−5.4	2.6
σ_{exp}	±0.03	±10	±10	±0.6	±0.05
TL6P	2.80	57	101	−5.66	1.86
TL6P$^{\mathrm{Sk}}$	2.80	58	101	−5.68	1.93
TL5P[S22]	2.71	59	85	−5.50	2.67
TL4P[S22]	2.78	59	102	−5.17	3.36
SWM6[S23]	2.79	56	123	−5.27	2.48

The experimental dimer geometry $\mathbf{G}_\mathrm{d}$ is illustrated in the left part of Figure S7. As indicated in the figure, $\mathbf{G}_\mathrm{d}$ is characterized by the oxygen-oxygen distance d_OO and by two angles (β_1, β_2).

The TL6P dimer geometry $\mathbf{G}_\mathrm{d}$, which is depicted in the right part of Figure S7 and numerically specified in Table S7, slightly differs from the experimental gas phase geometry in several respects. While the predicted angle β_1 coincides with the experimental angles well within the limits of experimental uncertainty, the value of 101° predicted for β_2 is just outside these limits and, thus is too small. The two TL6P models predict the distance $d_\mathrm{OO} = 2.80$ Å, which is by 6 % smaller than the experimental value. This value for d_OO is also predicted for the cyclic trimer, for which Pugliano et. al.[S24] determined the distance of 2.96 Å by far infrared absorption spectroscopy. Interestingly, however, the prediction for the gas phase dimer and trimer distances d_OO either coincides (TL6P$^\mathrm{Sk}$) with the targeted position r_1 of the first maximum of the liquid phase RDF $g_\mathrm{OO}(r)$ measured by Skinner et al.[S25] or is slightly larger (TL6P) than the target $r_1 = 2.76$ Å taken from Soper's[S26] data. As becomes apparent from a comparison with the TL5P and TL4P predictions,[S22] TL6P hardly performs better on the dimer geometry than its predecessors. The empirical SWM6 model, in contrast, for which the experimental angles β_1 and β_2 were optimization targets, excellently reproduces these angles while showing a likewise poor performance on d_OO.

According to Table S7 the TL6P potentials describe the binding energy E_pot of the dimer within the limits of experimental uncertainty, just like all other listed PMM models. The

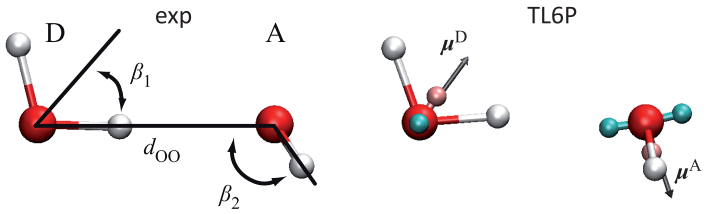


Figure S7: Geometry $\mathbf{G}_\mathrm{d}$ of the water dimer as determined by experimental data[S21] (left) and as predicted by the TL6P model (right). For TLP6 also the dipole moments $\boldsymbol{\mu}$ of the H-bond donor D and acceptor A are drawn as gray arrows.

total dipole moment μ_{d} of the dimer, however, which is very well described by SWM6, is underestimated by about 27 %.

In summary, we consider the TL6P performance on the dimer to be quite reasonable, because this very small cluster provides for each of its two components highly specific environments, which are very different from the fluctuating and more homogeneous environments found in the liquid phase at ambient temperatures and pressures. Correspondingly we doubt that dimer properties are beneficial as targets in empirical optimizations of models aimed at the liquid phase.

S6 DFT/TL6P Hybrid Dimers

In Ref. S22 the values $\tilde{\sigma}_{\mathrm{H}} = 0.281$ Å and $\tilde{\sigma}_{\mathrm{L/M}} = 0.460$ Å were chosen for the Gaussian widths $\tilde{\sigma}_i$ of those PMM charges, which are located close to a DFT atom in a DFT/PMM hybrid calculation. This choice had been derived from hybrid calculations on the water dimer, in which either the H-bond donor (DFT/TLνP) or the acceptor (TLνP/DFT) was described by DFT and its H-bonded partner by TLνP with $\nu = 4, 5$. With the aim of checking the performance of the TL6P potential in DFT/PMM hybrid settings we have calculated the DFT/TL6P and TL6P/DFT dimers applying the specified choice for the $\tilde{\sigma}_i$. The results are listed in Table S8.

A glance at the data in Table S8 immediately reveals that the two hybrid descriptions are

Table S8: DFT/PMM Geometries $\mathbf{G_d}$ and Binding Energies $E_{\mathbf{pot}}$ of TL6P Hybrid Dimers are Compared with Experimental Data[S19–S21] and Single Method Descriptions.

	d_{OO}/Å	β_1/deg	β_2/deg	$E_{\mathrm{pot,\,d}}/\frac{\mathrm{kcal}}{\mathrm{mol}}$
exp[S19–S21]	2.98	51	123	−5.4
DFT/TL6P	2.83	55	115	−5.2
TL6P/DFT	2.85	58	112	−5.3
TL6P	2.80	57	101	−5.7
DFT	2.98	56	123	−4.3

quite similar and that they interpolate between the pure TL6P and DFT dimers. Moreover, the hybrid dimers provide a better description of the experimental data than the TL6P dimer. Concerning the binding energy E_{pot} they are even better than the pure DFT description, which lacks some binding energy. A possible reason may be that the hybrid dimers include the dispersion attraction through the TL6P Buckingham potentials, which are attached also to the respective DFT fragments, whereas a pure DFT description does not account[S27] for this particular attraction.

S7 Properties Targeted by the Optimization

Table S9 documents the results of the $N_iV_iT_0$, $i \in \{\mathrm{s}, \mathrm{m}, \mathrm{l}\}$, simulations on the average potential energy $\langle E_{\mathrm{pot}} \rangle$ per molecule and on the average pressure $\langle p \rangle$. These observables were targeted by the weak-coupling[S28] optimizations for the medium-sized system using the target values $E_{\mathrm{pot}}(p_0, T_0) = -9.92$ kcal/mol and $p_0 = 1$ bar, which had been coupled to the Buckingham parameters B and A_2, respectively.

Table S9: $N_iV_iT_0$ Results for $\langle E_{\mathrm{pot}} \rangle$ and $\langle p \rangle$ at $n(T_0)$.

size i	observable	TL6P	TL6P$^{\mathrm{Sk}}$
s	$\langle E_{\mathrm{pot}} \rangle$/(kcal/mol)	-9.930 ± 0.003	-9.937 ± 0.005
m	$\langle E_{\mathrm{pot}} \rangle$/(kcal/mol)	-9.942 ± 0.004	-9.948 ± 0.003
l	$\langle E_{\mathrm{pot}} \rangle$/(kcal/mol)	-9.957 ± 0.003	-9.961 ± 0.004
s	$\langle p \rangle$/bar	15 ± 4	18 ± 6
m	$\langle p \rangle$/bar	3 ± 5	8 ± 5
l	$\langle p \rangle$/bar	-33 ± 4	-25 ± 4

The absolute values of the binding energies $\langle E_{\mathrm{pot}} \rangle$ are according to Table S9 for both TL6P models by 0.1 % - 0.4 % larger than the target value and slightly increase with the system size i. The pressures $\langle p \rangle$ correspondingly decrease with i.

Figure S8 clearly demonstrates that the positions r_1 of the first peaks of those experimental RDFs, which are marked by the vertical dotted lines and served as parametrization targets for the Buckingham parameter A_1 of TL6P and TL6P$^{\mathrm{Sk}}$, respectively, are reproduced

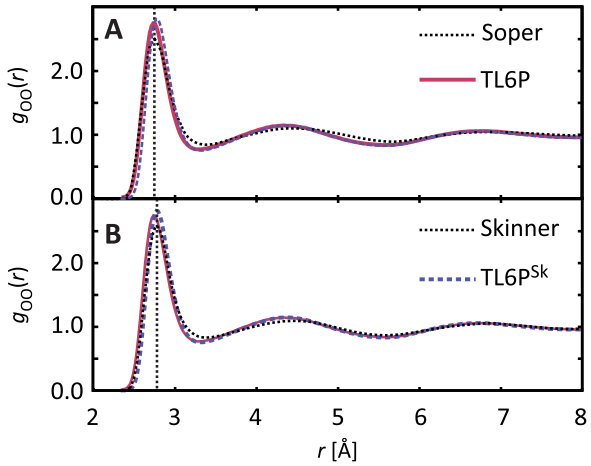


Figure S8: Radial distribution functions $g_{OO}(r)$ of TL6P (red) and TL6P^{Sk} (blue, dashed) compared with experimental references (black, dotted), for which the locations r_1 of the first maximum are marked by the vertical dotted lines; (A) RDF of Soper;[S26] (B) RDF of Skinner et al.[S25]

by the RDFs of these models, i.e. in Figure S8A by the RDF of TL6P (red) and in Figure S8B by the RDF of TL6P^{Sk} (blue dashed).

For distances larger than about 4.5 Å both TL6P RDFs are seen in Figure S8B to match the RDF of Skinner et al.[S25] extremely well, whereas at smaller distances both models predict a slightly more pronounced structure. Interestingly, differences between the RDFs of TL6P and TL6P^{Sk} become noticeable only at distances smaller than 3 Å, which is well within the range of the first peak marking the first solvation shell. Hence, differences of the TL6P and TL6P^{Sk} liquid structures at T_0 are of extremely short range.

In summary, the targeted observables agree very well with the target values.

S8 Predictions on Bulk Phase Properties

TL6P bulk phase properties were derived from the simulations sketched in Section 2 by the methods described in Sec. 3 of Ref. S22. Extending the description covered by the main text we here provide additional data and elaborate details.

S8.1 Statistical Errors of Predicted Bulk Properties

Statistical errors $\sigma(r)$ were estimated for the observables $r \in \{\kappa_T,\, C_p,\, \alpha_p\}$ listed in Table 4 of Section 3.3 by block-averaging[S29] and by the well-known laws of error propagation. As explained by Eqs. (11) - (13) in Ref. S22, these observables are calculated as finite differences of two quantities representing simple averages, whose statistical uncertainties can individually estimated by block averaging, such that the uncertainty of the composite quantity follows by error propagation. For $r \in \{\epsilon_0,\, D_0,\, \eta\}$ the statistical errors $\sigma(r)$ cannot be accessed by this approach. Here, values $\sigma(D_0)$ and $\sigma(\eta)$ were obtained from the linear regressions depicted in Figure S9 below in Section S8.2. Furthermore, also $\sigma(\epsilon_0)$ was obtained from a linear regression, namely the one depicted in Figure 3.

Table S10: Statistical Errors of TL6P Observables.

model	$\sigma(D_0)$ [a]	$\sigma(\eta)$ [b]	$\sigma(\kappa_T)$ [c]	$\sigma(C_p)$ [d]	$\sigma(\epsilon_0)$	$\sigma(\alpha_p)$ [e]
TL6P	0.05	0.19	0.3	0.4	0.1	0.6
TL6P$^{\mathrm{Sk}}$	0.02	0.11	0.3	0.4	0.5	0.6

[a] $\mathrm{nm}^2/\mathrm{ns}$. [b] mPa s. [c] $10^{-6}/\mathrm{atm}$. [d] cal/(mol K). [e] $10^{-4}\mathrm{K}^{-1}$

S8.2 Self-Diffusion Coefficients and Viscosities

The self-diffusion coefficient D measures the dynamical behavior of a water molecule in the liquid phase. It is derived from the $N_iV_iT_0$[MI] simulations by calculating the ensemble average molecular mean square displacements.[S8] The results $D(N_i, V_i, T_0)$ should depend[S30] on the box sizes V_i as measured by the inverse edge lengths $1/L_i$. This size dependence can be exploited[S30,S31] via the formula

$$D(N_i, V_i, T_0) = D_0 - \frac{k_{\mathrm{B}}T\xi}{6\pi\eta}\frac{1}{L_i}, \tag{S12}$$

where ξ has the value 2.837297 for cubic periodic boxes, for an extrapolation to the diffusion constant D_0 of the infinite system. The viscosity η is then determined by the slope of the

extrapolation line. The coefficients $D(N_i, V_i, T_0)$ obtained for the TL6P models are depicted together with extrapolating linear regressions in Figure S9. The indicated error bars are estimates σ of the statistical uncertainties associated with the coefficients $D(N_i, V_i, T_0)$, which were calculated by block averaging.[S29]

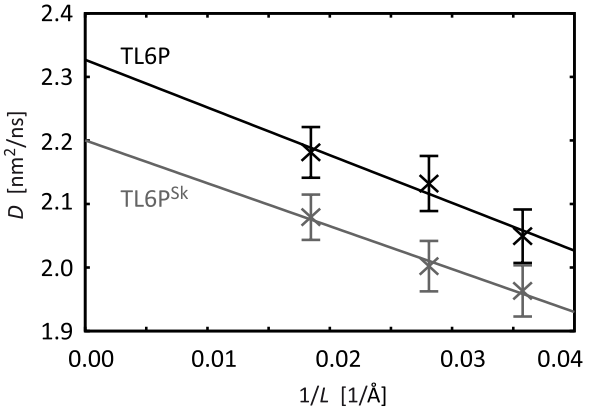


Figure S9: Linear regressions expressing Eq. (S12) for the self-diffusion coefficients D_0 and viscosities η of the TL6P models. The data points $D(N_i, V_i, T_0)$ are indicated by the crosses.

Neglecting the statistical scatter, the diffusion coefficients $D(N_i, V_i, T_0)$ show for both TL6P models the expected linear dependence on $1/L$. The regression lines yield for the infinite TL6P system the self-diffusion constant $D_0 = 2.3$ nm^2/ns and the viscosity $\eta = 0.86$ mPa s. The regression shows that only a small statistical error $\sigma(D_0)$ of 2 % is connected with the extrapolated diffusion coefficient D_0. According to Table S10 the error margin is ± 0.05 nm^2/ns for TL6P. In contrast, the statistical error $\sigma(\eta)$ associated with the viscosity η is much larger and measures 21 %. The corresponding TL6P$^{\mathrm{Sk}}$ values are $D_0 = 2.2 \pm 0.02$ ns/nm^2 and $\eta = 0.92 \pm 0.11$ mPa s and, hence, are quite similar. The data on D_0 and η have been transferred to Table 4 in Section 3.3 of the main text, where they are discussed by comparisons with experimental data and with results of the TL5P and TL4P models.

S8.3 Dipole Distributions of TL6P and TL6P$^{\mathrm{Sk}}$

For both TL6P models dipole distributions were constructed by extracting from the

$N_\mathrm{m}V_\mathrm{m}T_0$ simulations every 5 ps the absolute values μ of the total PMM dipole moments of 20 randomly selected water models. These dipole moments are normally distributed with standard deviations ρ around the average values $\langle\mu\rangle$ as listed in Table S11. The statistical uncertainties of $\langle\mu\rangle$ and ρ, which have been estimated by block-averaging,[S29] are negligibly small.

Table S11: Average Dipoles $\langle\mu\rangle$ and the Standard Deviations ρ of the TL6P and TL6P$^{\mathrm{Sk}}$ Dipole Distributions.

model	$\langle\mu\rangle$/D	ρ/D
TL6P	2.520	0.138
TL6P$^{\mathrm{Sk}}$	2.538	0.142

For both TL6P models the average dipole moments $\langle\mu\rangle$ are well within the range of 2.4-2.6 D, which has been suggested to be a prerequisite[S32,S33] for PMM models capable of reproducing the experimental dielectric constant $\epsilon(p_0, T_0)$.

Interestingly, the two TL6P values for $\langle\mu\rangle$ are almost equal to averages of the DFT/TL6P dipole moments $\mu_{\mathrm{DFT/PMM}}(s)$ discussed in connection with Figure S4. These snapshot ensembles yield for the two models the average dipole moment of 2.535 D. The average standard deviations of the DFT/TL6P data are, however, slightly larger than those of the two TL6P distributions and measure 0.155 D instead of the average TL6P value of 0.140 D. Thus, the dipole moments of the DFT fragments react a little more sensitively to changes of the surrounding liquid structure than the two TL6P models.

Table S12: Average Densities $\langle n\rangle/(\mathrm{g/cm^3})$ from the $N_i p_0 T_0$ Simulations.

size i	TL6P	TL6P$^{\mathrm{Sk}}$
s	0.997 ± 0.001	0.997 ± 0.003
m	0.997 ± 0.001	0.998 ± 0.003
l	0.998 ± 0.002	0.998 ± 0.001

S8.4 Density at T_0 and p_0

Table S12 lists the average densities observed in the $N_i p_0 T_0$ simulations on the three systems, which should be compared with experimental density[S34] $n(T_0) = 0.9965$ g/cm^3. The listed statistical errors were estimated using the block-averaging method described in Chap. 4 of Ref. S29. They are larger than the deviations of the average densities from the experimental value $n(T_0)$, i.e. the average densities cannot be distinguished from the experimental value with the given statistics.

S9 Computational Issues

While the TL6P models apparently can reproduce many properties of liquid water at p_0 and T_0 with a remarkable accuracy, one might nevertheless ask, whether this progress is bought by a possibly unmanageable computational cost.

The TL6P models are admittedly more costly than their TL4P and TL5P predecessors. Measurements with our program IPHIGENIE have shown that the TL6P computation time is, depending on the system size, by factors in the range from 1.15 to 1.4 larger than that of TL4P. Note here that the small factor applies to the large system. These surprisingly small factors are a result of the fast multipole expansions as implemented in the SAMM$_4$ algorithm[S1] for the speedy evaluation of the long-range interactions (see Section S1 for details).

Another approach, which has been chosen by several groups to parameterize complex model potentials, is the use of Gaussian instead of point charges.[S15,S18,S35–S37] The evaluation of the Gaussians, which is necessary only at small distances, implies a certain computational overhead. To estimate this overhead we have replaced the three point charges in our TL4P model by Gaussian distributions thus obtaining a model called TL4P$_\mathrm{G}$. Because IPHIGENIE can also handle such models, we have carried out $N_i V_i T_0$ simulations for the three system sizes $i \in \{\mathrm{s}, \mathrm{m}, \mathrm{l}\}$. As compared to TL4P we obtained factors in the range from 1.13 ($i = \mathrm{l}$)

to 1.38 ($i = \mathrm{s}$), which are almost as large as those measured for TL6P.

Thus, TL6P can be handled (by IPHIGENIE) almost as efficiently as $\mathrm{TL4P_G}$. For sufficiently large systems, these more complex models furthermore solely cause computational overheads in the range between 13-15 %.

If one is willing to sacrifice the accuracy of description to a large extent, then one can get, of course, considerable computational speed-ups. Exchanging the TL4P reference by TL3P, for instance, saves about 25 % of the computation time while using the non-polarizable TIP3P model can save even 80 %. Thus, the main contribution to the computational overhead of the TLνP models as compared to TIP3P are the self-consistency iterations required for the inducible dipoles.

S10 Melting Temperature of Ice

Here we describe and discuss the MD simulations on mixtures of ice Ih and water at different temperatures near the putative melting temperature T^{m}, which serve to determine this value for the PMM water models TL4P and TL6P. As a check of the applied methods we also present simulations for the non-polarizable model TIP4P/2005,[S38] whose melting temperature has been previously determined as $T^{\mathrm{m}}_{\mathrm{TIP4P/2005}} = 250 \pm 3$ K by other authors using different methods.[S38,S39]

S10.1 Methods

Within the moving-boundary reaction field approach[S5] implemented in our MD program IPHIGENIE[S1] a small block of a PMM or MM ice Ih model with the side length $L_{\mathrm{ice}} = 2.352$ nm was considered as the solute. The surrounding PMM or MM model of liquid water, which fills a periodic cubic box of side-length $L_{\mathrm{box}} = 5.176$ nm, represents the solvent consisting of 3992 molecules. For distances $> L_{\mathrm{box}}/2$ the surrounding water is treated as a dielectric continuum with the experimental[S40] dielectric constant $\epsilon_{\mathrm{RF}} = \epsilon_{\mathrm{exp}}(T, p_0)$ of liquid

water. Thus, the minimum image convention[S8] requiring $L_{\mathrm{box}} > 2L_{\mathrm{ice}}$ is fulfilled.

Periodic blocks of ice Ih containing 432 molecules were built following the algorithm of Buch.[S41] Temperature-adapted volumes $V_{\mathrm{ice}}(T_z)$ were generated for these blocks by 200 ps MD simulations in the Np_0T_z ensemble. Here, the temperature T was held at T_z with a Berendsen thermostat[S42] (coupling constant 1 ps), the pressure p was kept with an anisotropic Berendsen barostat[S42] at p_0 (coupling constant 10 ps), and the distant continuum was modeled by the dielectric constant $\epsilon_{\mathrm{RF}} = 97$ of ice.[S43] For TIP4P/2005 six target temperatures T_z were equidistantly selected from the interval [235, 260] K. Also for TL6P we chose six target temperatures T_z in 5 K steps, but this time from the interval [250, 275] K. For TL4P only the single low temperature $T_z = 250$ K had to be considered.

The initial volume $V_{\mathrm{box}}(T_z) = V_{\mathrm{ice}}(T_z) + V_{\mathrm{liq}}(T_z)$ of the respective water-ice mixture model was determined for each temperature T_z in such a way that the density $n(T_z, p_0)$ in $V_{\mathrm{liq}}(T_z)$ had either the value determined in Ref. S44 for TL6P and TL4P or, for TIP4P/2005, the experimental value.[S34] The volume $V_{\mathrm{ice}}(T_z)$ was adopted from the solid state Np_0T_z simulations described above, of course, ensuring that both the liquid and the solid subsystems had the equilibrium densities of the respective pure phases. As a result we obtained, e.g. for TL6P at $T_z = 250$ K the system depicted in Figure S10.

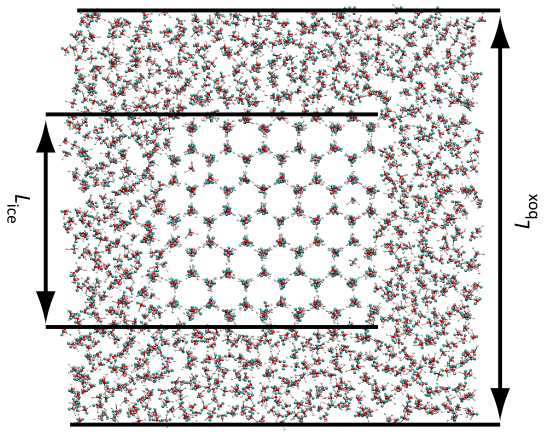


Figure S10: A section of 2 nm depth through the cubic and periodic TL6P water-ice mixture system of edge length $L_{\mathrm{box}} = 5.176$ nm at $T_z = 250$ K. An almost cubic ice block covering 432 molecules is surrounded by 3992 liquid phase molecules.

The thus generated initial simulation systems were adjusted to mixed-phase conditions in a two step procedure. First, the oxygen atoms of the water models in the ice block were softly bound to their positions by stiff harmonic potentials (force constant 10 kcal/(mol Å^2) and the system was steered toward T_z by a 2 ps MD simulation in the NVT_z ensemble using a short 0.1 fs time step and by coupling a very fast Berendsen thermostat to all atoms in the system (coupling constant 10 fs).

Next, the harmonic restraints were relaxed and the system was equilibrated for another 200 ps (time step 1 fs). This time, however, the Berendsen thermostat was solely coupled to atoms in the liquid phase (coupling time 1 ps), which was taken as the thermostat also for the solute ice block. In line with expectations,[S45] the temperature relaxation in the ice toward temperature T_z in the liquid proceeded on a time scale of about 5 ps.

Furthermore we realized a variant of the Berendsen barostat[S42] by exclusively scaling the dimensions of the cubic box and leaving the atomic coordinates invariant. Hence, in this case pressure changes are induced at the boundaries and the ice structure at the center is left untouched. A very large relaxation time of 100 ps was chosen here, such that the box volume could change only very slowly. Nevertheless, Np_0T_z simulations could be realized in this way. The final 800 ps production phases were carried out exactly in the same Np_0T_z setup for each model and at each considered temperature T_z. At every ps the average potential energy $E_{\mathrm{pot}}(t\,|\,T_z)$ per water molecule was written to file for further evaluation.

S10.2 Results

To validate the novel mixed-phase simulation setup sketched above, we first consider the six 800 ps simulations for TIP4P/2005 at the temperatures $T_z \in [235, 260]$ K. Figure S11a shows the trajectories $E_{\mathrm{pot}}(t\,|\,T_z)$ calculated at the various color-coded temperatures. Here, the trajectory of lowest energy belongs to the lowest temperature, of course. Decreasing values of $E_{\mathrm{pot}}(t)$ indicate ongoing processes of freezing whereas increasing values of $E_{\mathrm{pot}}(t)$ point toward melting. Thus, as demonstrated by Figure S11b the signs and magnitudes of

the slopes of the shown regression lines can separate these two cases.

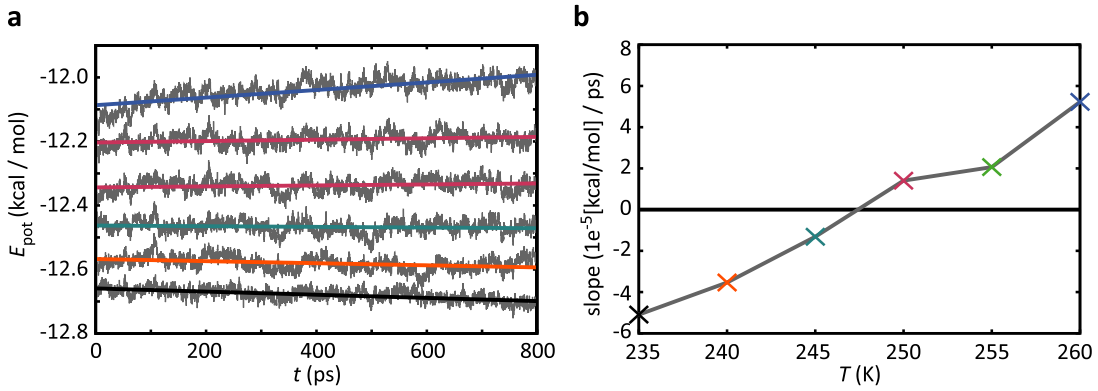


Figure S11: TIP4P/2005 Melting Temperature. a) Temporal evolution of $E_{\mathrm{pot}}(t \,|\, T_z)$ along the six 800 ps MD trajectories simulated for TIP4P/2005 at the color-coded temperatures $T_z \in [235, 260]$ K. Regression lines are fitted for each T_z to the trajectories $E_{\mathrm{pot}}(t \,|\, T_z)$. b) Slopes $\Delta E_{\mathrm{pot}}/\Delta T$ of the regression lines shown in a) as functions of the color-coded simulation temperatures T_z.

Figure S11b clearly demonstrates that the slopes of the regression lines shown in Figure S11a increase with increasing temperatures from strongly negative values at $T_z = 235$ K to likewise strongly positive values at $T_z = 260$ K. The melting temperature $T^{\mathrm{m}}_{\mathrm{TIP4P/2005}} = 247.5$ K predicted by our simulations follows from the interpolated temperature of zero slope. This value is just within the error bounds (250 ± 3 K) estimated by other authors using different methods.[S39,S46]

Estimates of the errors and uncertainties, which have to be attached to our value for $T^{\mathrm{m}}_{\mathrm{TIP4P/2005}}$, cannot yet be seriously given, because they would require much more extended simulations and the consideration of differently sized simulation systems. Nevertheless, the close similarity of our value (247.5 K) with that of Abascal and Vega[S38] (252) and Fernandez et al.[S39] (249 K) indicates that melting temperatures obtained with our simulation setup can be regarded as reasonable estimates and lower bounds.

Applying our mixed-phase simulation setup to the TL6P model yields the data shown in Figure S12. Here, the temperature of zero slope in Figure S12b yields our estimate $T^{\mathrm{m}}_{\mathrm{TL6P}} = 262.5$ K (and lower bound) for the TL6P prediction of the melting temperature.

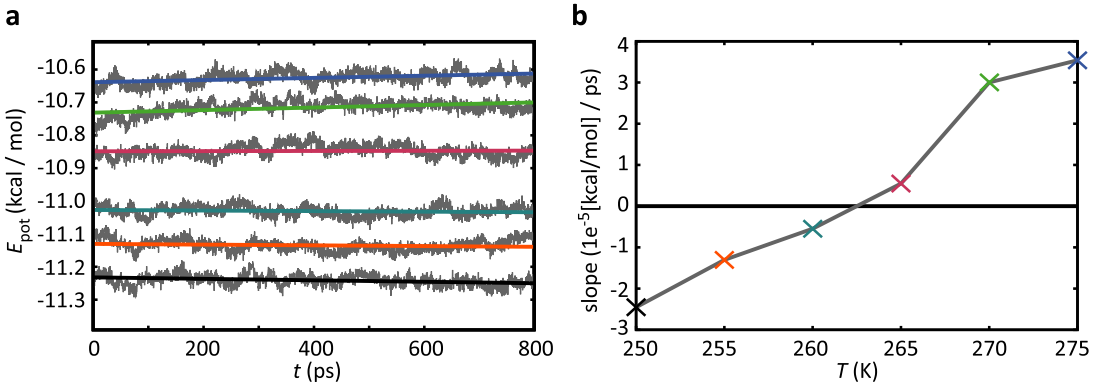


Figure S12: TL6P Melting Temperature. a) Temporal evolution of $E_{\mathrm{pot}}(t \,|\, T_z)$ along the six 800 ps MD trajectories simulated for TL6P at the color-coded temperatures $T_z \in [250, 275]$ K. Regression lines are fitted for each T_z to the trajectories $E_{\mathrm{pot}}(t \,|\, T_z)$. b) Slopes $\Delta E_{\mathrm{pot}}/\Delta T$ of the regression lines shown in a) as functions of the color-coded simulation temperatures T_z.

Hence, TL6P underestimates the observed melting temperature[S47] $T^{\mathrm{m}}_{\mathrm{exp}} = 273.15$ K of ice Ih by at most 11 K.

For TL4P, in contrast, already the lowest temperature $T_z = 250$ K yielded a strongly positive slope indicating a rapid melting at this temperature. Thus this value is solely an upper bound for the TL4P melting temperature and we expect that $T^{\mathrm{m}}_{\mathrm{TL4P}} \ll 250$ K.

S10.3 Discussion

As argued in Section 4.3 of the main text, for a PMM model like TL6P the 0.02 % overestimate of the density $n_{\mathrm{exp}}(T^{\mathrm{m}}_{\mathrm{exp}}, p_0)$ at the experimental melting temperature $T^{\mathrm{m}}_{\mathrm{exp}}$ by the simulation result $n_{\mathrm{TL6P}}(T^{\mathrm{m}}_{\mathrm{exp}}, p_0)$ renders the expectation that $T^{\mathrm{m}}_{\mathrm{exp}}$ is somewhat underestimated by $T^{\mathrm{m}}_{\mathrm{TL6P}}$. The mixed-phase simulations now have shown that this underestimate measures just 11 K.

This value is smaller than for other PMM models. In a sequence of decreasing quality the PMM five-point model POL4D,[S48] for instance, underestimates $T^{\mathrm{m}}_{\mathrm{exp}}$ by 13.5 K and overestimates $n_{\mathrm{exp}}(T^{\mathrm{m}}_{\mathrm{exp}}, p_0)$ by 0.10 %. The corresponding numbers are (23.1 K, 0.12 %) for BK3,[S15] (40.1 K, 0.18 %) for BKd3,[S18] ($\gg$ 23 K, 0.96%) for TL4P as determined above and in Ref. S44, and finally (88 K, 1.32 %) for the related PMM four-point model SWM4/NDP.[S49,S50]

Thus the suggested connection between the underestimate of $T_{\mathrm{exp}}^{\mathrm{m}}$ and the overestimate of $n_{\mathrm{exp}}(T_{\mathrm{exp}}^{\mathrm{m}}, p_0)$ is nicely supported by the available data on the properties of PMM models. Hence, a good performance of TL6P on the prediction of $T_{\mathrm{exp}}^{\mathrm{m}}$ was as much to be expected as the lousy performance of TL4P on the same observable.

We would like to note that the effective[S51] and only partially polarizable many-parameter model iAMOEBA[S52] underestimates $T_{\mathrm{exp}}^{\mathrm{m}}$ by 12 K while *underestimating* concurrently the experimental density by about 0.06 %. Here the lacking correlation between the two numbers underlines the effective character of this model. PMM models behave differently.

References

(S1) Lorenzen, K.; Schwörer, M.; Tröster, P.; Mates, S.; Tavan, P. Optimizing the Accuracy and Efficiency of Fast Hierarchical Multipole Expansions for MD Simulations. *J. Chem. Theory Comput.* **2012**, *8*, 3628–3636.

(S2) Niedermeier, C.; Tavan, P. A Structure Adapted Multipole Method for Electrostatic Interactions in Protein Dynamics. *J. Chem. Phys.* **1994**, *101*, 734–748.

(S3) Niedermeier, C.; Tavan, P. Fast Version of the Structure Adapted Multipole Method: Efficient Calculation of Electrostatic Forces in Protein Dynamics. *Mol. Simul.* **1996**, *17*, 57–66.

(S4) Eichinger, M.; Grubmüller, H.; Heller, H.; Tavan, P. FAMUSAMM: An Algorithm for Rapid Evaluation of Electrostatic Interactions in Molecular Dynamics Simulations. *J. Comput. Chem.* **1997**, *18*, 1729–1749.

(S5) Mathias, G.; Egwolf, B.; Nonella, M.; Tavan, P. A Fast Multipole Method Combined with a Reaction Field for Long-Range Electrostatics in Molecular Dynamics Simulations: The Effects of Truncation on the Properties of Water. *J. Chem. Phys.* **2003**, *118*, 10847–10860.

(S6) Martinetz, T.; Berkovich, S.; Schulten, K. 'Neural-gas" Network for Vector Quantization and its Application to Time-Series Prediction. *IEE Trans. Neur. Networks* **1993**, *4*, 558–569.

(S7) Kloppenburg, M.; Tavan, P. Deterministic Annealing for Density Estimation by Multivariate Normal Mixtures. *Phys. Rev. E* **1997**, *55*, R2089–R2092.

(S8) Allen, M. P.; Tildesley, D. *Computer Simulations of Liquids*; Clarendon: Oxford, 1987.

(S9) Schwörer, M.; Breitenfeld, B.; Tröster, P.; Lorenzen, K.; Tavan, P.; Mathias, G. Coupling DFT to Polarizable Force Fields for Efficient and Accurate Hamiltonian Molecular Dynamics Simulations. *J. Chem. Phys.* **2013**, *138*, 244103.

(S10) Bronstein, I. N.; Semendjajew, K. A.; Musiol, G.; Mühlig, H. *Taschenbuch der Mathematik*; Verlag Harri Deutsch: Frankfurt am Main, 1999.

(S11) Kvasov, B. I. *Methods of Shape Preserving Spline Approximation*; World Scientific: Singapore, 2000.

(S12) Pulay, P. Convergence Acceleration of Iterative Sequences - The Case of SCF Iteration. *Chem. Phys. Lett.* **1980**, *73*, 393–398.

(S13) Császár, C.; Pulay, P. Geometry Optimization by Direct Inversion in the Iterative Subspace. *J. Mol. Struct.* **1984**, *114*, 31–34.

(S14) Hill, P. G.; MacMillan, C. R. Virial Equations for Light and Heavy Water. *Ind. Eng. Chem. Res.* **1988**, *27*, 874–882.

(S15) Kiss, P. T.; Baranyai, A. A Systematic Development of a Polarizable Potential of Water. *J. Chem. Phys.* **2013**, *138*, 204507.

(S16) Sokolić, F.; Guissani, Y.; Baranović, G. Intermolecular Pair Potentials and the Second Virial Coefficient of Sulphur Dioxide Vapour. *Chem. Phys. Lett.* **1986**, *131*, 513–516.

(S17) Hirschfelder, J. O.; Curtiss, C. F.; Bird, R. B. *Molecular Theory of Gases and Liquids*; Wiley New York, 1954; Vol. 26.

(S18) Kiss, P. T.; Baranyai, A. Density Maximum and Polarizable Models of Water. *J. Chem. Phys.* **2012**, *137*, 084506.

(S19) Verhoeven, J.; Dymanus, A. Magnetic Properties and Molecular Quadrupole Tensor of the Water Molecule by Beam-Maser Zeeman Spectroscopy. *J. Chem. Phys.* **1970**, *52*, 3222–3233.

(S20) Dyke, T. R.; Mack, K. M.; Muenter, J. S. The Structure of Water Dimer from Molecular Beam Electric Resonance Spectroscopy. *J. Chem. Phys.* **1977**, *66*, 498–510.

(S21) Odutola, J. A.; Dyke, T. R. Partially Deuterated Water Dimers: Microwave Spectra and Structure. *J. Chem. Phys.* **1980**, *72*, 5062–5070.

(S22) Tröster, P.; Lorenzen, K.; Schwörer, M.; Tavan, P. Polarizable Water Models from Mixed Computational and Empirical Optimization. *J. Phys. Chem. B* **2013**, *117*, 9486–9500.

(S23) Yu, W.; Lopes, P. E. M.; Roux, B.; A. D. MacKerell, J. Six-Site Polarizable Model of Water Based on the Classical Drude Oscillator. *J. Chem. Phys.* **2013**, *138*, 034508.

(S24) Pugliano, N.; Saykally, R. Measurement of Quantum Tunneling Between Chiral Isomers of the Cyclic Water Trimer. *Science* **1992**, *257*, 1937–1940.

(S25) Skinner, L. B.; Huang, C.; Schlesinger, D.; Pettersson, L. G.; Nilsson, A.; Benmore, C. J. Benchmark Oxygen-Oxygen Pair-Distribution Function of Ambient Water from X-Ray Diffraction Measurements with a Wide Q-Range. *J. Chem. Phys.* **2013**, *138*, 074506.

(S26) Soper, A. K. The Radial Distribution Functions of Water and Ice from 220 to 673 K and at Pressures up to 400 MPa. *Chem. Phys.* **2000**, *258*, 121–137.

(S27) Liu, Y.; Goddard III, W. A. First-Principles-Based Dispersion Augmented Density Functional Theory: From Molecules to Crystals. *J. Phys. Chem. Lett.* **2010**, *1*, 2550–2555.

(S28) Berweger, C.; van Gunsteren, W.; Müller-Plathe, F. Force Field Parametrization by Weak Coupling. Re-Engineering SPC Water. *Chem. Phys. Lett.* **1995**, *232*, 429–436.

(S29) Rapaport, D. C. *The Art of Molecular Dynamics Simulation*; Cambridge university press: Cambridge, UK, 2004.

(S30) Dünweg, B.; Kremer, K. Molecular Dynamics Simulation of a Polymer Chain in Solution. *J. Chem. Phys.* **1993**, *99*, 6983–6997.

(S31) Yeh, I.-C.; Hummer, G. System-Size Dependence of Diffusion Coefficients and Viscosities from Molecular Dynamics Simulations with Periodic Boundary Conditions. *J. Phys. Chem. B* **2004**, *108*, 15873–15879.

(S32) Sprik, M. Hydrogen Bonding and the Static Dielectric Constant in Liquid Water. *J. Chem. Phys.* **1991**, *95*, 6762–6769.

(S33) Lamoureux, G.; MacKerell Jr, A.; Roux, B. A Simple Polarizable Model of Water Based on Classical Drude Oscillators. *J. Chem. Phys.* **2003**, *119*, 5185–5197.

(S34) Kell, G. Precise Representation of Volume Properties of Water at One Atmosphere. *J. Chem. Eng. Data* **1967**, *12*, 66–69.

(S35) Paricaud, P.; Předota, M.; Chialvo, A.; Cummings, P. From Dimer to Condensed Phases at Extreme Conditions: Accurate Predictions of the Properties of Water by a Gaussian Charge Polarizable Model. *J. Chem. Phys.* **2005**, *122*, 244511.

(S36) Baranyai, A.; Kiss, P. A Transferable Classical Potential for the Water Molecule. *J. Chem. Phys.* **2010**, *133*, 144109.

(S37) Baranyai, A.; Kiss, P. Polarizable Model of Water with Field-Dependent Polarization. *J. Chem. Phys.* **2011**, *135*, 234110.

(S38) Abascal, J.; Vega, C. A General Purpose Model for the Condensed Phases of Water: TIP4P/2005. *J. Chem. Phys.* **2005**, *123*, 234505.

(S39) Fernández, R. G.; Abascal, J. L.; Vega, C. The Melting Point of Ice I for Common Water Models Calculated from Direct Coexistence of the Solid-Liquid Interface. *J. Chem. Phys.* **2006**, *124*, 144506.

(S40) Kaatze, U. Complex Permittivity of Water as a Function of Frequency and Temperature. *J. Chem. Eng. Data* **1989**, *34*, 371–374.

(S41) Buch, V.; Sandler, P.; Sadlej, J. Simulations of H_2O Solid, Liquid, and Clusters, with an Emphasis on Ferroelectric Ordering Transition in Hexagonal Ice. *J. Phys. Chem. B* **1998**, *102*, 8641–8653.

(S42) 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. *J. Chem. Phys.* **1984**, *81*, 3684–3690.

(S43) Johari, G.; Whalley, E. The Dielectric Properties of Ice Ih in the Range 272–133 K. *J. Chem. Phys.* **1981**, *75*, 1333–1340.

(S44) Tröster, P.; Tavan, P. The Microscopic Physical Cause for the Density Maximum of Liquid Water. *J. Phys. Chem. Lett.* **2014**, *5*, 138–142.

(S45) Lingenheil, M.; Denschlag, R.; Reichold, R.; Tavan, P. The Hot-Solvent/Cold-Solute Problem Revisited. *J. Chem. Theory Comput.* **2008**, *4*, 1293–1306.

(S46) Vega, C.; Sanz, E.; Abascal, J. L. F. The Melting Temperature of the Most Common Models of Water. *J. Chem. Phys.* **2005**, *122*, 114507.

(S47) Feistela, R.; Wagner, W. A New Equation of State for H_2O Ice Ih. *J. Phys. Chem. Ref. Data* **2006**, *35*, 1021–1047.

(S48) Viererblová, L.; Kolafa, J. A Classical Polarizable Model for Simulations of Water and Ice. *Phys. Chem. Chem. Phys.* **2011**, *13*, 19925–19935.

(S49) Lamoureux, G.; Harder, E.; Vorobyov, I.; Roux, B.; MacKerell, A. A Polarizable Model of Water for Molecular Dynamics Simulations of Biomolecules. *Chem. Phys. Lett.* **2006**, *418*, 245–249.

(S50) Gladich, I.; Roeselová, M. Comparison of Selected Polarizable and Nonpolarizable Water Models in Molecular Dynamics Simulations of Ice Ih. *Phys. Chem. Chem. Phys.* **2012**, *14*, 11371–11385.

(S51) Effective mean-field type potentials, which try to compensate the fully (e.g. TP4P/2005[S38]) or partially (e.g. iAMOEBA[S52]) neglected polarizability by an empirically optimal but nevertheless non-physical choice of the model parameters, may be of practical use in simulations but are almost certainly of limited value for predictions on conditions, which are not covered by the reference data (e.g. water in proteins, electrolytes etc.). PMM models, in contrast, which are good enough to catch the essential physics of a water molecule, should be transferable to all kinds of environments as long as the assumption of linear response is valid. Furthermore they open the chance to identify the microscopic causes for the macroscopic properties of water.

(S52) Wang, L.-P.; Head-Gordon, T.; Ponder, J. W.; Ren, P.; Chodera, J. D.; Eastman, P. K.; Martinez, T. J.; Pande, V. S. Systematic Improvement of a Classical Molecular Model of Water. *J. Phys. Chem. B* **2013**, *117*, 9956–9972.

2.3 Die Mikroskopische Begründung der Dichteanomalie

Die Dichteanomalie ist die mit Abstand bekannteste der vielen Anomalien des Wassers. Der mikrokopische physikalische Grund für dieses ungewöhnliche Verhalten des Wassers war aber bislang unbekannt.

In der Publikation[3]

> „The Microscopic Physical Cause for the Density Maximum of Liquid Water“, Philipp Tröster, Konstantin Lorenzen, and Paul Tavan, *J. Phys. Chem. Lett.*, **5**, 138-142, (2014)

von mir und Paul Tavan wird der mikroskopische physikalische Grund für die Dichteanomalie identifiziert.
Durch ausgiebige 20 ns *replica exchange* Simulationen werden Temperatur-Dichte Profile für die Wassermodelle TL4P, TL5P und TL6P berechnet. Durch die gewonnenen Daten kann die große Qualität des TL6P Modells, das dieses Profil und die Temperatur maximaler Dichte sehr genau vorhersagt, weiter untermauert werden. Da lediglich das TL6P Modell experimentelle Daten reproduzieren kann, die Vorgängermodelle aber grob falsches Verhalten zeigen, kann der molekulare physikalische Grund der Dichteanomalie identifiziert werden.

[3]Reproduced with permission from the Journal of Physical Chemistry, **5**, 138-142, 2014.

THE JOURNAL OF PHYSICAL CHEMISTRY Letters

Letter

pubs.acs.org/JPCL

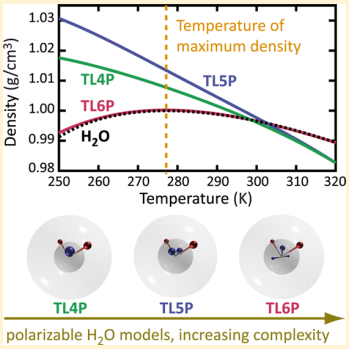
The Microscopic Physical Cause for the Density Maximum of Liquid Water

Philipp Tröster and Paul Tavan*

Lehrstuhl für Biomolekulare Optik, Fakultät für Physik, Ludwig-Maximilians-Universität München, Oettingenstr. 67, D-80538 München, Germany

Supporting Information

ABSTRACT: The existence of a density maximum at 277 K is probably the most prominent anomaly among the many very special thermodynamic properties of liquid water. While usually attributed to so-called hydrogen bonding, the microscopic physical cause of this prominent anomaly is still elusive. Here we show that the density anomaly is caused by those short-range electrostatic forces, which are generated by the quadrupole and higher moments of the charge distributions present in liquid-phase water molecules. This conclusion derives from 20 ns replica exchange molecular-dynamics simulations with closely related polarizable four-, five-, and six-point water models. As soon as the model complexity suffices to represent the higher electrostatic moments with sufficient accuracy, the density temperature profile $n(T)$ calculated for $T \in [250,320]$ K at the standard pressure 1 bar locks in to the experimental observation. The corresponding six-point model is, therefore, the most simple available cartoon for liquid-phase water molecules.



SECTION: Liquids; Chemical and Dynamical Processes in Solution

The thermodynamic behavior of liquid water differs from that of other liquids.[1,2] An anomaly of importance for life on earth is the existence of a density maximum at 277.134 K.[3] This anomaly has been attributed to a competition between attractive forces acting among the water molecules, which entail higher densities at lower temperatures, and structure-forming "hydrogen bonding" interactions, which have the opposite effect.[1,2] However, a precise physical characterization of the structure forming forces is still lacking.

Motivated by the aim to tackle this challenge and, more generally, to contribute to the physical understanding of liquid water, we and coworkers have recently developed[4,5] a series of closely related and increasingly complex polarizable four-, five-, and six-point water models, which are called[6] TL4P, TL5P, and TL6P, respectively. These polarizable molecular mechanics (PMM) potentials have been optimized at $T_0 \equiv 300$ K and at the density[3] $n_{exp}(T_0,p_0) = 0.9965$ g/cm^3 of liquid water at the standard pressure $p_0 = 1$ bar by a mixed computational and empirical approach.[4]

As empirical corner stones, the static dipole moments and polarizabilities of all of these TLνP models have the experimental gas-phase values[7,8] $\mu^g_{exp} = 1.855$ D and $\alpha^g_{exp} = 1.47$ Å^3, respectively.

As a first theoretically motivated[9] corner stone, the induced dipole moments are represented by Gaussian distributions $\boldsymbol{\mu}^i(r)$ of widths σ_ν because this choice guarantees that the required average of the electric field over the polarizable volume of a condensed phase water molecule is properly included.[4] Here σ_ν measures this volume.

As a second theoretically founded[10] corner stone, the three masses of H_2O are arranged in the liquid-phase geometry[11,12] (bond angle: 105.3 deg, bond length: 0.968 Å). This geometry is compatible with the use of the gas-phase values μ^g_{exp} and α^g_{exp} because the transfer of a water molecule from the gas into the liquid phase leaves its static dipole moment and polarizability almost invariant despite the geometry change.[10]

Furthermore, applying a novel quantum-classical hybrid approach,[13] which combines density functional theory (DFT) of a solute molecule with a PMM representation of its environment, the Gaussian widths σ_ν of the inducible dipole densities $\boldsymbol{\mu}^i(r)$ and the electrostatic signatures of the TLνP models, that is, the locations of the negative partial charges in the vicinity of the O atom, were derived by DFT/PMM calculations.

Finally, all TLνP models describe the van der Waals interactions by a three-parameter Buckingham potential[14] $E_{B,\nu}(r\,|\,A,k,B) = A\exp(-rk) - B/r^6$ centered at the oxygen, which was adjusted by weak-coupling[15] PMM molecular dynamics simulations to experimental liquid-phase quantities belonging to the thermodynamic reference conditions T_0 and $n_{exp}(T_0,p_0)$. Here the targeted experimental values were for A the position of the first peak of an oxygen−oxygen radial distribution function,[16] for k the standard pressure p_0, and for B the potential energy[17] $E_{pot}(T_0,p_0)$ per molecule.

Because of this strategy of model construction, the three selected experimental liquid-phase properties could influence the resulting parameters only at a single point $[T_0, n_{exp}(T_0,p_0)]$

Received: November 6, 2013
Accepted: December 11, 2013
Published: December 11, 2013

in the thermodynamic state space of liquid water. Therefore, TLνP simulation results obtained under other thermodynamic conditions represent predictions. Furthermore, the selected empirical properties directly solely steered the models for the van der Waals interactions, whereas the models for the zero-field electrostatic properties of liquid phase water molecules, that is, the partial charge distributions, and for the polarizable molecular volumes, that is, the Gaussian widths σ_ν, were only indirectly and weakly affected through the self-consistently constructed ensembles of PMM water models, which were employed in the associated DFT/PMM calculations.[4]

The resulting TLνP potentials mainly differ from previously developed PMM approaches of similar complexity like COS/G3[18] and SWM4-NDP[19] with four points, POL4D[20] with five points, or SWM6[21] with six points, through replacing the common charge-on-spring model for the inducible dipole by a Gaussian distribution and through the strong reliance on DFT/PMM calculations, which removes the electrostatic properties from empirical optimization leaving, here, only a few van der Waals parameters.

Figure 1 sketches the properties of the PMM TLνP models for water. The radii $r_{\mathrm{vdW},\nu}$ of the glassy van der Waals spheres

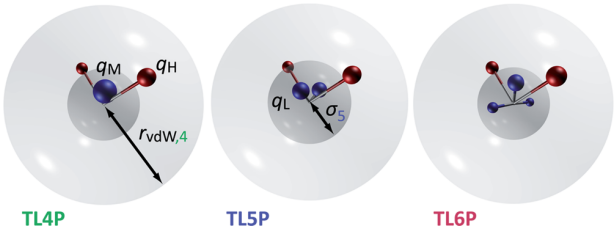


Figure 1. A TLνP model[4,5] for water has a van der Waals sphere (glassy) of radius $r_{\mathrm{vdW},\nu}$ and an inducible Gaussian dipole distribution (gray glassy) of width σ_ν centered at the O atom, two positive partial charges q_H attached to the H atoms (red spheres), and $\nu - 3$ partial point charges $q_i < 0$, $i \in \{M,L\}$ (blue spheres). For $\nu = 4$, a charge q_M occupies the so-called M-site on the bisectrix of the HOH triangle. For $\nu = 5$, two charges q_L sit at the two L sites, which are symmetrically arranged in the plane spanned by the bisectrix and by the normal of the HOH triangle. For $\nu = 6$, all three sites carry charges.

surrounding the oxygen atoms are defined by the condition $E_{B,\nu}(2r_{\mathrm{vdW},\nu} \mid A,k,B) = 0$. The sizes σ_ν of the inducible Gaussian dipole distributions are indicated by the gray glassy spheres. The static partial charges q are drawn as red ($q > 0$) and blue ($q < 0$) spheres, whose volumes scale with the absolute values $|q|$ of these charges.

All depicted partial charge distributions represent the same static dipole moment μ^{g}_{exp} but different higher multipole moments. With increasing model complexity ν, they generate electrostatic potentials on a spherical surface surrounding the oxygen at a distance of 2.75 Å, which approximate the static contribution to the potential originating from the distributed charge distributions of identically positioned DFT water models embedded in liquid-phase PMM environments at a successively better quality.[4,5] These charge distributions were calculated[4,5] by a novel DFT/PMM hybrid approach.[13] Here the quality of the surface potential approximation shows a significant jump in the transition to TL6P. A similar quality jump appears when one compares the quadrupole moments of the TLνP models with those calculated by DFT for an isolated water molecule.[5]

The less complex TL4P and TL5P models turned out to predict many properties (e.g., density, diffusion constant, viscosity, isobaric heat capacity, isothermal compressibility, dielectric constant) of liquid water at p_0 and T_0 at least as well as those previous four- and five-point PMM models, which also exhibit partial charges at fixed positions,[4] but as is the case for all such models, the isobaric thermal expansion coefficient $\alpha_p \equiv -\partial \ln[n(T,p_0)]/\partial T$ was overestimated by at least a factor of 2. TL6P, in contrast, not only predicted all enumerated properties with an improved accuracy but also even reproduced α_p exactly (within statistical limits).[5]

This success raised the question whether TL6P can reasonably predict the well-known[3] temperature dependence $n_{\mathrm{exp}}(T,p_0)$ of the density over a temperature interval covering not only the temperature $T^{\mathrm{md}}_{\mathrm{exp}} \equiv 277.134$ K of maximum density but also lower temperatures in the range of supercooling, for example, down to $T = 250$ K. Here the thermal motion of the water molecules considerably slows down; therefore, MD simulations of periodic boxes, which contain a reasonably large number N of water models and are kept at a constant temperature T, do not guarantee a statistically sufficient sampling of liquid-phase properties. To tackle this challenge, we decided to apply the replica exchange (RE) technique in the Np_0T ensemble[22] because it leads to a statistically homogeneous sampling at all rungs T_i of the chosen temperature ladder $i = 1, 2, ..., m$.

Thus, for each of the three TLνP potentials, a total of 15 cubic periodic boxes with inner radii of ~14 Å were filled with $N = 728$ water models because this system size had been shown to suffice[4,5] for accurate computations of the average density $\langle n \rangle$ at T_0 and p_0. In 20 ns RE-MD simulations, the temperatures T_i of the $m = 15$ replicas i were controlled by Bussi thermostats[23] (coupling time: 1 ps). The T_i spanned the interval [250, 320] K in steps of 5 K. The pressure p was kept at $p_0 = 1$ bar by a Berendsen barostat[24] (coupling time 10 ps, compressibility 0.46 Gbar^{-1}). Applying the so-called deterministic even−odd scheme,[25] we attempted replica exchanges every 5 ps. These RE-MD simulations were carried out with the program package IPHIGENIE,[26] taking advantage of its efficient fast multipole treatment of the electrostatics,[26] of its toroidal boundary conditions,[27] and of its moving-boundary reaction field correction for the long-range electrostatics.[28] Here temperature dependent experimental values[29] were chosen to model the dielectric constants $\varepsilon_{\mathrm{RF}}(T_i p_0)$ of the distantly surrounding dielectric continua.

The average exchange rate resulting from the chosen RE temperature ladder was ~27% for each model. Correspondingly, each of the 15 replicas completed about three so-called round trips[30] up and down the entire temperature ladder within the 20 ns of RE-MD. Thus, each replica repeatedly took advantage of the accelerated dynamics at $T_m = 320$ K and deposited the results of this enhanced phase space sampling at the lowest temperature rung $T_1 = 250$ K. As a result, the average statistical errors of the average densities $\langle n_{\mathrm{TL}\nu\mathrm{P}} \rangle (T_i p_0)$ were as small as 0.01%. Section S1 of the Supporting Information (SI) contains a table listing these densities and the remaining statistical uncertainties $\sigma_{\mathrm{TL}\nu\mathrm{P}}(T_i p_0)$, which were estimated by block-averaging.[31] Section S1 furthermore specifies for each TLνP model a fourth-order polynomial $n_{\mathrm{TL}\nu\mathrm{P}}(T,p_0)$ interpolating the simulation results.

Figure 2 compares the experimental density temperature profile $n_{\mathrm{exp}}(T,p_0)$, as represented by the interpolating function given in ref 3, with the RE-MD simulation results $\langle n_{\mathrm{TL}\nu\mathrm{P}} \rangle (T_i p_0)$

dx.doi.org/10.1021/jz4023927 | J. Phys. Chem. Lett. 2014, 5, 138−142

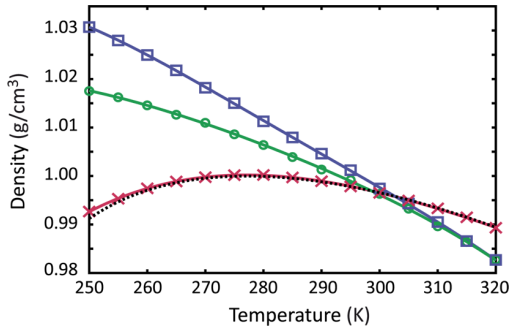


Figure 2. Comparison of the experimental temperature density profile $n_{exp}(T,p_0)$ with RE-MD results for the TLνP models. The experimental data[3] are represented by the black dotted line. The predicted average densities $\langle n_{TL\nu P}\rangle(T_i,p_0)$ are marked by symbols and are, like the graphs of the interpolating polynomials $n_{TL\nu P}(T,p_0)$, distinguished by colors. TL6P: red, TL5P: blue, TL4P: green.

and with the associated polynomials $n_{TL\nu P}(T,p_0)$. The TL6P predictions (red) are seen to closely match $n_{exp}(T,p_0)$ (black dotted). Within the depicted temperature range, the set $\langle n_{TL6P}\rangle(T_i,p_0)$ deviates from $n_{exp}(T_i,p_0)$ by a root-mean-square deviation of only 0.0005 g/cm^3.

In contrast, the less complex TL5P (blue) and TL4P (green) predecessor models match $n_{exp}(T,p_0)$ only at the temperature T_0 = 300 K used in the parametrization. At other temperatures, the density profiles $n_{TL5P}(T,p_0)$ and $n_{TL4P}(T,p_0)$ quickly deviate from $n_{exp}(T,p_0)$ because, as previously noted,[4] the slopes as expressed by the derivatives

$$\alpha_{p,TL\nu P}(T, p_0) \equiv -\partial \ln[n_{TL\nu P}(T, p_0)]/\partial T$$

are much too large at T_0 for ν = 4,5. Although TL5P is more complex than TL4P, $\alpha_{p,TL5P}(T_0,p_0)$ is larger than $\alpha_{p,TL4P}(T_0,p_0)$. Each of these models predicts a monotonous density increase for decreasing temperatures.

As we have seen in Figure 2, for TL6P, the accurate derivative[5] $\alpha_{p,TL6P}(T_0,p_0)$ actually transfers into an equally accurate description of $n_{exp}(T,p_0)$ within the whole range of temperatures covered by the RE-MD simulations. The excellence of the TL6P prediction outmatches all of those hopes and expectations, which originally inspired this study. The polynomial $n_{TL6P}(T,p_0)$ has its maximum at T^{md}_{TL6P} = 277.005 K, that is, only 0.129 K below the experimental value T^{md}_{exp}. If we estimate the statistical uncertainty of T^{md}_{TL6P} by cutting the data into halves, we find an estimated standard deviation of 0.125 K, implying that TL6P almost reproduces T^{md}_{exp} within the limits of statistical accuracy. Similarly, TL6P overestimates the experimental value (0.99997 g/cm^3) of the maximal density by only 0.02%.

As a result, TL6P is a PMM model for water, which not only accurately describes many properties of liquid water (including the dielectric constant) under the thermodynamic conditions T_0 and $n_{exp}(T_0,p_0)$ of the parametrization[5] but also likewise excellently generalizes the density $n_{exp}(T,p_0)$ to other temperatures. In the latter respect, its predecessors[4] TL4P and TL5P failed. For readers interested in further properties and possible limitations of TL6P, Sections S2−S4 of the SI present and discuss predictions on $\alpha_p(T,p_0)$, on the vaporization enthalpy $\Delta H(T,p_0)$, and on the heat-capacity $C_p(T,p_0)$ derived from the RE-MD simulations.

The thus apparent progress in the construction of PMM models for water solely required the addition of a third negative charge near the oxygen (cf. Figure 1) while leaving all other aspects and algorithmic procedures of the quantum mechanically assisted model construction[4] invariant. Because the key feature of this addition was a markedly improved description[5] of the higher moments of the charge distribution in liquid phase water molecules (as sampled by DFT/PMM calculations), we now can safely conclude that the density anomaly of water is solely caused by the associated short-range electrostatics. Concurrently, the TL6P distribution of partial point charges depicted in Figure 1 provides a minimal cartoon for those aspects of the continuous charge distributions in real water molecules, which are responsible for the density anomaly.

It should be noted that this electrostatics view is the prevalent one in the community of force-field designers and makes no reference to the so-called hydrogen-bonding interactions, which are popularly inferred for qualitative explanations.[1,2] On the basis of this view, MM force fields for water and biomolecular simulations[32,33] model the dominant structure-forming interactions through electrostatic forces. In the case of a PMM model for water, partial charges and an inducible dipole serve to approximate the electrostatic potential generated by the charge density of a liquid-phase water molecule as closely as possible. This concept was a key guideline of our DFT/PMM-based TLνP model construction, and its apparent success underlines the viability of the electrostatics view.

In conclusion, we would like to remark that minimal but nevertheless chemically accurate molecular models most likely cannot be constructed empirically. Instead, additional information from quantum-mechanical calculations of the kind employed for TL6P is deemed necessary. In support of this suggestion, Figure 3 compares the best previous PMM

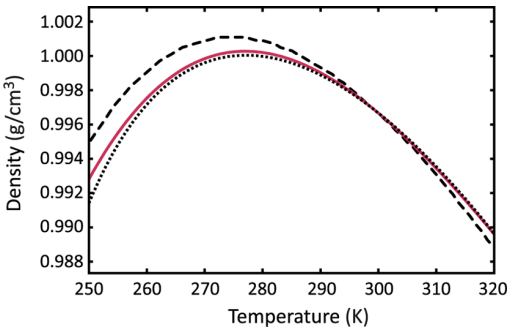


Figure 3. Comparison of the TL6P prediction $n_{TL6P}(T,p_0)$ with another PMM simulation result[34] and with $n_{exp}(T,p_0)$. Graphs of interpolating polynomials are drawn for TL6P by a red solid line and for the recent BK3 model[34] by a black dashed line (adopted from figure 5 in ref 34). Experimental ref 3 $n_{exp}(T,p_0)$: black dotted line.

description[34] (black dashed) of $n_{exp}(T,p_0)$ (black dotted), which was achieved by the empirically parametrized so-called BK3 model, with $n_{TL6P}(T,p_0)$ (red). BK3 has one parameter more than TL6P and features three Gaussian charges on springs implying a distributed polarizability to the charge points of a four-point geometry. The targets of the empirical optimization included the densities of the liquid at T = 298 K and p_0 and of hexagonal ice, implying that interpolating thermodynamic conditions were covered. Although $n_{BK3}(T,p_0)$ describes $n_{exp}(T,p_0)$ very well, it cannot compete with the DFT/PMM-based prediction represented by $n_{TL6P}(T,p_0)$. Note that in Section S5 of the SI, we briefly discuss the performance

dx.doi.org/10.1021/jz4023927 | J. Phys. Chem. Lett. 2014, 5, 138−142

of nonpolarizable and partially polarizable water models on $n(T,p_0)$, whereas Section S6 of the SI contains a corresponding discussion for PMM models with computational complexities similar to those of TL4P, TL5P, and TL6P.

In summary, we have shown that the density anomaly of liquid water is generated by the electrostatic signatures of the water molecules in the liquid, as expressed by the quadrupole and higher moments of their charge distributions. The conclusion rests on results of RE-MD simulations executed with the closely related PMM models TL4P, TL5P, and TL6P for water.[4,5] TL6P accurately predicts the observed[3] density temperature profile $n_{exp}(T,p_0)$ over the temperature range 250−320 K, whereas its less complex relatives TL5P and TL4P, like other models of the same complexity as reviewed in ref 35, grossly miss $n_{exp}(T,p_0)$. Because the only key difference between the TLνP models is the significantly better representation of the higher multipole moments by TL6P,[5] these results demonstrate that the structure forming forces generating the density maximum at 277 K are simply short-range electrostatic multipole interactions. Furthermore, the excellent performance of TL6P nourishes the hope that the underlying DFT/PMM-based methodology[4,13] for the construction of PMM force fields can help to advance the field of biomolecular simulation toward chemical accuracy.

■ ASSOCIATED CONTENT

Supporting Information

The SI provides in Section S1 two tables numerically specifying the RE-MD results depicted in Figure 2. Furthermore, it contains in Sections S2−S4 three figures and explanatory text on the temperature dependences of the thermal expansion coefficient $\alpha_p(T,p_0)$, of the vaporization heat $\Delta H(T,p_0)$, and of the isobaric heat capacity $C_p(T,p_0)$ of TL6P, which were derived from the RE-MD simulations. Section S5 discusses the practical merits and conceptual limitations of effective water models, which avoid an explicit and complete description of the polarizability. Section S6 compares the performance of the TL4P, TL5P, and TL6P models on $n(T,p_0)$ with that of similarly complex PMM models. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: tavan@physik.uni-muenchen.de.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work has been supported by the Deutsche Forschungsgemeinschaft (SFB 749/C4).

■ REFERENCES

(1) Eisenberg, D.; Kauzmann, W. *The Structure and Properties of Water*; Oxford University Press: London, 1969.
(2) Tanaka, H. Simple Physical Explanation of the Unusual Thermodynamic Behavior of Liquid Water. *Phys. Rev. Lett.* **1998**, *80*, 5750−5753.
(3) Kell, G. S. Precise Representation of Volume Properties of Water at One Atmosphere. *J. Chem. Eng. Data* **1967**, *12*, 66−69.
(4) Tröster, P.; Lorenzen, K.; Schwörer, M.; Tavan, P. Polarizable Water Models from Mixed Computational and Empirical Optimization. *J. Phys. Chem. B* **2013**, *117*, 9486−9500.
(5) Tröster, P.; Lorenzen, K.; Tavan, P. Polarizable Six-Point Water Models from Computational and Empirical Optimization. *J. Phys. Chem. B* **2013**, submitted.
(6) The letters "TL" in the name TLνP of our water models are the initials of the first two authors of ref 4. The number $\nu \in \{3,4,5,6\}$ counts the points of force action within the respective model, and "P" stands for "point".
(7) Clough, S.; Beers, Y.; Klein, G.; Rothman, L. Dipole Moment of Water from Stark Measurements of H2O, HDO, and D2O. *J. Chem. Phys.* **1973**, *59*, 2254−2259.
(8) Murphy, W. The Rayleigh Depolarization Ratio and Rotational Raman Spectrum of Water Vapor and the Polarizability Components for the Water Molecule. *J. Chem. Phys.* **1977**, *67*, 5877−5882.
(9) Schropp, B.; Tavan, P. The Polarizability of Point-Polarizable Water Models: Density Functional Theory/Molecular Mechanics Results. *J. Phys. Chem. B* **2008**, *112*, 6233−6240.
(10) Schropp, B.; Tavan, P. Flexibility Does Not Change the Polarizability of Water Molecules in the Liquid. *J. Phys. Chem. B* **2010**, *114*, 2051−2057.
(11) Ichikawa, K.; Kameda, Y.; Yamaguchi, T.; Wakita, H.; Misawa, M. Neutron-Diffraction Investigation of the Intramolecular Structure of a Water Molecule in the Liquid Phase at High Temperatures. *Mol. Phys.* **1991**, *73*, 79−86.
(12) Thiessen, W. E.; Narten, A. H. Neutron Diffraction Study of Light and Heavy Water Mixtures at 25 °C. *J. Chem. Phys.* **1982**, *77*, 2656−2662.
(13) Schwörer, M.; Breitenfeld, B.; Tröster, P.; Lorenzen, K.; Tavan, P.; Mathias, G. Coupling DFT to Polarizable Force Fields for Efficient and Accurate Hamiltonian Molecular Dynamics Simulations. *J. Chem. Phys.* **2013**, *138*, 244103.
(14) Buckingham, R. A.; Corner, J. Tables of Second Virial and Low-Pressure Joule-Thomson Coefficients for Intermolecular Potentials with Exponential Repulsion. *Proc. R. Soc. London, Ser. A* **1947**, *189*, 118−129.
(15) Berweger, C.; van Gunsteren, W.; Müller-Plathe, F. Force Field Parametrization by Weak Coupling. Re-Engineering SPC Water. *Chem. Phys. Lett.* **1995**, *232*, 429−436.
(16) Soper, A. K. The Radial Distribution Functions of Water and Ice from 220 to 673 K and at Pressures up to 400 MPa. *Chem. Phys.* **2000**, *258*, 121−137.
(17) Jancso, G.; Van Hook, W. A. Condensed Phase Isotope Effects. *Chem. Rev.* **1974**, *74*, 689−750.
(18) Yu, H.; van Gunsteren, W. Charge-On-Spring Polarizable Water Models Revisited: From Water Clusters to Liquid Water to Ice. *J. Chem. Phys.* **2004**, *121*, 9549−9564.
(19) Lamoureux, G.; Harder, E.; Vorobyov, I.; Roux, B.; MacKerell, A. A Polarizable Model of Water for Molecular Dynamics Simulations of Biomolecules. *Chem. Phys. Lett.* **2006**, *418*, 245−249.
(20) Viererblová, L.; Kolafa, J. A Classical Polarizable Model for Simulations of Water and Ice. *Phys. Chem. Chem. Phys.* **2011**, *13*, 19925−19935.
(21) Yu, W.; Lopes, P. E. M.; Roux, B.; MacKerell, A. D., Jr. Six-Site Polarizable Model of Water Based on the Classical Drude Oscillator. *J. Chem. Phys.* **2013**, *138*, 034508.
(22) Okabe, T.; Kawata, M.; Okamoto, Y.; Mikami, M. Replica-Exchange Monte Carlo Method for the Isobaric-Isothermal Ensemble. *Chem. Phys. Lett.* **2001**, *335*, 435−439.
(23) Bussi, G.; Parrinello, M. Stochastic Thermostats: Comparison of Local and Global Schemes. *Comput. Phys. Commun.* **2008**, *179*, 26−29.
(24) 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. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(25) Lingenheil, M.; Denschlag, R.; Mathias, G.; Tavan, P. Efficiency of Exchange Schemes in Replica Exchange. *Chem. Phys. Lett.* **2009**, *478*, 80−84.
(26) Lorenzen, K.; Schwörer, M.; Tröster, P.; Mates, S.; Tavan, P. Optimizing the Accuracy and Efficiency of Fast Hierarchical Multipole Expansions for MD Simulations. *J. Chem. Theory Comput.* **2012**, *8*, 3628−3636.

(27) Allen, M. P.; Tildesley, D. *Computer Simulations of Liquids*; Clarendon: Oxford, U.K., 1987.
(28) Mathias, G.; Egwolf, B.; Nonella, M.; Tavan, P. A Fast Multipole Method Combined with a Reaction Field for Long-Range Electrostatics in Molecular Dynamics Simulations: The Effects of Truncation on the Properties of Water. *J. Chem. Phys.* **2003**, *118*, 10847−10860.
(29) Kaatze, U. Complex Permittivity of Water as a Function of Frequency and Temperature. *J. Chem. Eng. Data* **1989**, *34*, 371−374.
(30) Denschlag, R.; Lingenheil, M.; Tavan, P. Optimal Temperature Ladders in Replica Exchange Simulations. *Chem. Phys. Lett.* **2009**, *473*, 193−195.
(31) Rapaport, D. C. *The Art of Molecular Dynamics Simulation*; Cambridge university press: Cambridge, U.K., 2004.
(32) Guillot, B. A Reappraisal of What We Have Learnt During Three Decades of Computer Simulations on Water. *J. Mol. Liq.* **2002**, *101*, 219−260.
(33) Ponder, J. W.; Case, D. A. Force Fields for Protein Simulation. *Adv. Protein Chem.* **2003**, *66*, 27−85.
(34) Kiss, P. T.; Baranyai, A. A Systematic Development of a Polarizable Potential of Water. *J. Chem. Phys.* **2013**, *138*, 204507.
(35) Kiss, P. T.; Baranyai, A. Density Maximum and Polarizable Models of Water. *J. Chem. Phys.* **2012**, *137*, 084506.

Supporting Information to: The Microscopic Physical Cause for the Density Maximum of Water

Philipp Tröster and Paul Tavan*

Lehrstuhl für Biomolekulare Optik, Fakultät für Physik,

Ludwig-Maximilians-Universität München,

Oettingenstr. 67, D-80538 München, Germany

E-mail: tavan@physik.uni-muenchen.de

*To whom correspondence should be addressed

S1 TLνP density profiles from RE-MD

Table S1 lists the density profiles $\langle n_{\mathrm{TL}\nu\mathrm{P}}\rangle(T_i)$ obtained by the RE-MD simulations at p_0 for $\nu \in \{6, 5, 4\}$. Also given are estimates of the statistical errors calculated by block averaging.[1] The data in the table were employed to fit for all TLνP models the coefficients of polynomials of the form

$$n(T) = a_0 + a_1 T + a_2 T^2 + a_3 T^3 + a_4 T^4. \tag{S1}$$

The resulting coefficients are displayed by Table S2.

Table S1: The average densities $\langle n_{\mathbf{TL}\nu\mathbf{P}}\rangle(T_i)$ at p_0 from RE-MD.

T_i/K	$\langle n_{\mathrm{TL6P}}\rangle(T_i)$/(g/cm^3)	$\langle n_{\mathrm{TL5P}}\rangle(T_i)$/(g/cm^3)	$\langle n_{\mathrm{TL4P}}\rangle(T_i)$/(g/cm^3)
250	0.99267 ± 0.00024	1.03071 ± 0.00025	1.01755 ± 0.00026
255	0.99531 ± 0.00022	1.02797 ± 0.00023	1.01623 ± 0.00023
260	0.99745 ± 0.00018	1.02496 ± 0.00020	1.01458 ± 0.00020
265	0.99887 ± 0.00016	1.02182 ± 0.00017	1.01264 ± 0.00017
270	0.99976 ± 0.00014	1.01820 ± 0.00016	1.01097 ± 0.00016
275	1.00018 ± 0.00013	1.01502 ± 0.00016	1.00862 ± 0.00016
280	1.00021 ± 0.00013	1.01129 ± 0.00012	1.00641 ± 0.00012
285	0.99970 ± 0.00011	1.00794 ± 0.00012	1.00391 ± 0.00010
290	0.99893 ± 0.00009	1.00462 ± 0.00010	1.00133 ± 0.00010
295	0.99784 ± 0.00009	1.00120 ± 0.00010	0.99918 ± 0.00009
300	0.99653 ± 0.00010	0.99742 ± 0.00009	0.99632 ± 0.00008
305	0.99495 ± 0.00009	0.99432 ± 0.00009	0.99326 ± 0.00008
310	0.99332 ± 0.00008	0.99050 ± 0.00009	0.98962 ± 0.00008
315	0.99148 ± 0.00007	0.98658 ± 0.00008	0.98638 ± 0.00007
320	0.98932 ± 0.00006	0.98269 ± 0.00007	0.98271 ± 0.00007

Table S2: The TLνP coefficients a_i of the polynomials Eq. (S1).

coefficient	unit	TL6P	TL5P	TL4P
a_0	g/cm^3	-2.46948	-6.27299	-2.41137
a_1	10^{-2} g/cm^3 K^{-1}	4.02195	10.4196	4.88528
a_2	10^{-4} g/cm^3 K^{-2}	-1.7264	-5.51377	-2.59393
a_3	10^{-7} g/cm^3 K^{-3}	3.26969	12.8749	6.12089
a_4	10^{-10} g/cm^3 K^{-4}	-2.33374	-11.2672	-5.4684

S2 Thermal Expansion Coefficients

Because the RE-MD simulations yielded the temperature dependent average densities $\langle n_{\mathrm{TL6P}}\rangle(T_i)$ at p_0 listed in Table S1, one can immediately calculate from these data temperature dependent thermal expansion coefficients $\alpha_p(\tilde{T}_i, p_0)$ at the temperatures $\tilde{T}_i \equiv T_i + (T_{i+1} - T_i)/2$ by the finite differences

$$\alpha_p(\tilde{T}_i, p_0) \approx -\frac{\ln[\langle n\rangle(T_{i+1}, p_0)] - \ln[\langle n\rangle(T_i, p_0)]}{T_{i+1} - T_i}. \tag{S2}$$

Then the statistical uncertainties $\sigma_\alpha(\tilde{T}_i)$ of $\alpha_p(\tilde{T}_i, p_0)$ follow from the uncertainties of $\langle n_{\mathrm{TL6P}}\rangle(T_i)$ also listed in Table S1 by error propagation. Alternatively, one can obtain a smooth estimate for $\alpha_p(T, p_0)$ by properly differentiating the negative logarithm of the polynomial $n_{\mathrm{TL6P}}(T, p_0)$ given by Eq. (S1) and Table S2.

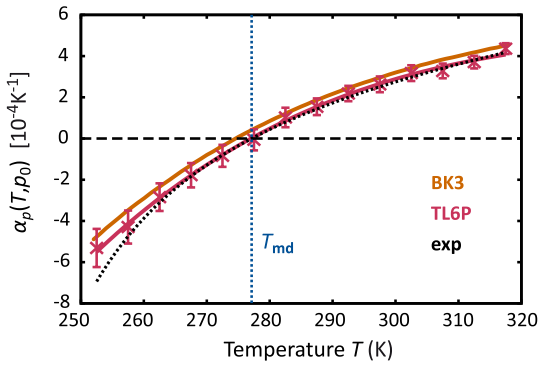


Figure S4: The temperature dependence of thermal expansion coefficient $\alpha_p(T, p_0)$ calculated for TL6P either from the interpolating polynomial Eq. (S1) specified by Table S2 (red line) or by Eq. (S2) from the data in Table S1 (red crosses) together with the associated statistical uncertainties $\sigma_\alpha(\tilde{T}_i)$ (red bars) is compared with experimental data[2] (black dotted). Data for the BK3 model (orange) were extracted from Figure 8 of Ref. 3 and are shown for comparison. The experimental temperature T_{md} of maximum density is indicated by the blue dashed line.

Figure S4 compares the TL6P predictions on $\alpha_p(T, p_0)$ with those of the BK3[3] model and with experimental data.[2] The numerical approximation Eq. (S2) to the derivative magnifies the statistical uncertainties of the density data listed in Table S1. Nevertheless, the TL6P predictions $\alpha_p(\tilde{T}_i, p_0)$ (red crosses) reproduce the experimental data (black dotted line) for temperatures above 275.5 K within the limits of the statistical uncertainties $\sigma_\alpha(\tilde{T}_i)$. In the temperature range from

267.5 K to 317.5 K the prediction by the TL6P polynomial (red line) almost perfectly matches the experimental data. BK3 (orange line), in contrast, systematically overestimates $\alpha_p(T, p_0)$ everywhere. The good match of the TL6P prediction on $\alpha_p(T, p_0)$ is, of course, an immediate consequence of the excellent match of $n(T, p_0)$ documented by Figure 3.

In view of the excellent performance of TL6P on $n(T, p_0)$ and $\alpha_p(T, p_0)$, readers, who are interested in simulation descriptions of water at all kinds of thermodynamic conditions, may ask in what respects the underlying model assumptions will lead to less accurate predictions.

Here the answer is pretty clear. TL6P, like its predecessors, employs partial point charges to model the electrostatic signature of an isolated water molecule. At temperatures below the parameterization temperature T_0, at which the structuring effect of the higher multipole moments gains weight, these point charges are likely to induce a slightly too strong binding, i.e. the heat of vaporization $\Delta H_{\mathrm{vap}}(T, p_0)$ will be increasingly overestimated at decreasing temperatures.[3–5] To illustrate this issue and possible remedies we now consider $\Delta H_{\mathrm{vap}}(T, p_0)$ and its negative derivative, the isobaric heat capacity $C_p(T, p_0)$.

S3 Heats of vaporization

Like the thermal expansion coefficients also the vaporization heats $\Delta H_{\mathrm{vap}}(T_i, p_0)$ can be extracted from the Np_0T RE-MD simulations at the temperatures T_i. These quantities are approximately given by[6]

$$\Delta H_{\mathrm{vap}}(T_i, p_0) \approx \langle E_{\mathrm{pot}} \rangle (T_i, p_0) + RT_i - p_0 \langle v \rangle (T_i, p_0) + C(T_i, p_0), \tag{S3}$$

where $\langle E_{\mathrm{tot}} \rangle (T_i, p_0)$ denotes the average potential energy per molecule, R the gas constant, and $\langle v \rangle (T_i, p_0)$ the volume per molecule at the temperature T_i. The quantities $C(T_i, p_0)$ are corrections,[6] which account for the quantum nature of the molecular vibrations and librations and for the non-ideal character of water vapor at the respective temperature. These quantities are tabulated in Ref. 6 for a series of temperatures, from which the values $C(T_i, p_0)$ at the temperatures T_i employed in our RE-MD simulation follow by interpolation.

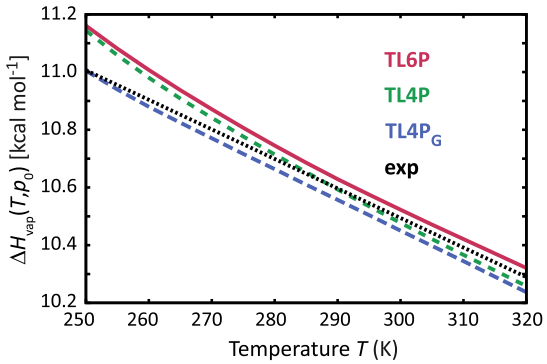


Figure S5: Predictions on the heat of vaporization $\Delta H_{\mathrm{vap}}(T, p_0)$ by TL6P (red line), by TL4P (green dashed), and by a variant TL4P$_{\mathrm{G}}$ (blue dashed), which employs Gaussian charge distributions instead of point charges (see the text for explanation) to express the electrostatic signature, are compared with experimental[7] data (black dotted line).

Figure S5 compares the predictions of TL6P (red line) on $\Delta H_{\mathrm{vap}}(T, p_0)$ with experimental data[7] (black dotted). The figure shows that TL6P overestimates the experimental values at all temperatures. Above 290 K these overestimates are small and approximately constant at about 0.03 kcal/mol. Toward lower temperatures the overestimates become larger reaching a value of 0.15 kcal/mol at 250 K. At this temperature TL6P and its much less complex predecessor TL4P (green dashed line) overestimate the enthalpy of vaporization by about equal amounts. Toward higher temperatures, however, the TL4P overestimate successively diminishes and turns into an underestimate for $T \geq 290$ K.

According to Kiss and Baranyai[3,5] large low-temperature overestimates of $\Delta H_{\mathrm{vap}}(T, p_0)$ are typical for water models employing partial point charges. Experience has shown[3,5] that these overestimates can be diminished, if one replaces the point charges by Gaussian distributions like in BK3. A corresponding water model has been first introduced by Paricaud et al.[4]

Table S3: Differing properties of TL4P$_{\mathrm{G}}$ and TL4P.

	$A/(10^3\mathrm{kcal/mol})$	$B/(\mathrm{Å}^6\mathrm{kcal/mol})$	$k/(1/\mathrm{Å})$	$\sigma_{\mathrm{G}}/\mathrm{Å}$
TL4P$_{\mathrm{G}}$	268	579	4.11	0.80
TL4P	84	992	3.55	0.84

Because we had tentatively developed also a polarizable four-point model, in which the partial

charges q_H and q_M, were replaced by Gaussian distributions of the widths $\sigma_H = 0.351$ Å and $\sigma_M = 0.560$ Å, respectively, we have checked the quoted suggestion and have carried out a short (5 ns) RE-MD simulation in the Np_0T ensemble also for this model, which we called TL4P$_G$. Note that the locations and strengths of the Gaussian charge distributions of TL4P$_G$ are identical to the locations and strengths of the TL4P point charges. The width σ_G of the Gaussian inducible dipole, however, is a little smaller in TL4P$_G$ than in TL4P (see the data in Table S3). The change of the Buckingham parameters is also documented in Table S3. It was induced by the weakened short-range electrostatic forces of TL4P$_G$, served to partially compensate the reduced short-range attraction, and was similar to the change observed earlier in the transition from TL4P/TL5P to TL6P.[8]

The blue line in Figure S5 represents the predictions of TL4P$_G$ on $\Delta H_{vap}(T, p_0)$ and shows that the TL4P low-temperature overestimate of the experimental vaporization enthalpy actually vanishes through the use of Gaussian partial charge distributions. For higher temperatures the TL4P and TL4P$_G$ predictions approach each other with TL4P$_G$ providing a slightly larger underestimate of $\Delta H_{vap}(T, p_0)$ than TL4P. If the use of Gaussian charges in a hypothetical and computationally still more complex model TL6P$_G$ would introduce, compared to TL6P, the same changes as TL4P$_G$ compared to TL4P, then one would arrive at an almost perfect prediction of $\Delta H_{vap}(T, p_0)$ with errors smaller than 0.1 %. Thus, for an improved representation of the low-temperature energetics of water the development of a corresponding model seems to be promising.

S4 Heat Capacities

Also the isobaric heat capacity $C_p(T, p_0)$ can be extracted from the Np_0T RE-MD simulations at the temperatures $\tilde{T}_i$ by the numerical derivatives

$$C_p(\tilde{T}_i, p_0) \approx \frac{\langle E_{tot}\rangle(T_{i+1}) - \langle E_{tot}\rangle(T_i)}{T_{i+1} - T_i} - p_0 \frac{\langle v\rangle(T_{i+1}) - \langle v\rangle(T_i)}{T_{i+1} - T_i} + \Delta C_{QM}(\tilde{T}_i, p_0) \qquad \text{(S4)}$$

of the liquid phase enthalpy $H(T,p_0) = \langle E_{\mathrm{tot}}\rangle(T,p_0) - p_0\langle v\rangle(T,p_0)$ per molecule, where $\langle E_{\mathrm{tot}}\rangle(T,p_0)$ and $\langle v\rangle(T)$ denote the average total energy and volume per molecule, respectively. The quantities $\Delta C_{\mathrm{QM}}(\tilde{T}_i,p_0)$ in Eq. (S4) are temperature dependent quantum corrections,[6] which account for errors originating from the classical description and from the enforced stiffness of the TL6P water models. Statistical uncertainties $\sigma_E(T_i)$ of the energies and $\sigma_v(T_i)$ of the volumes were calculated by block averaging, from which the uncertainties uncertainties $\sigma_C(T_i)$ were calculated by error propagation.

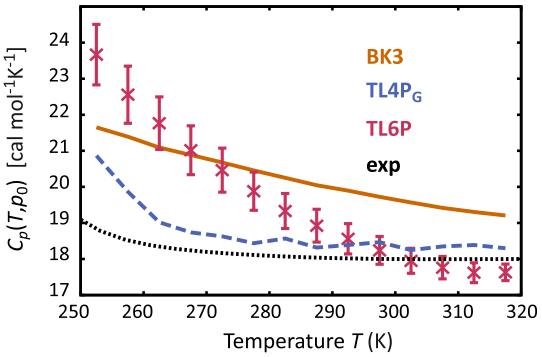


Figure S6: The TL6P predictions (red crosses) on the heat capacities $C_p(\tilde{T}_i,p_0)$ at constant pressure and their uncertainties $\sigma_C(\tilde{T}_i)$ (red bars) are compared with the predictions of BK3[3] (orange line) and TL4P$_{\mathrm{G}}$ (blue dashed line) and with experimental[7] data (black dotted line). The BK3 data are extracted from Fig. 7 in Ref. 3 and are corrected by $\Delta C_{\mathrm{QM}}(T,p_0)$.

Figure S6 compares the resulting TL6P predictions (red) on $C_p(T,p_0)$ with those of BK3[3] (orange) and of TL4P$_{\mathrm{G}}$ (blue) and with experimental data (black dotted). TL6P yields a good description of the experimental evidence on $C_p(T,p_0)$ for $T \geq 290$ K. At lower temperatures the heat capacity is increasingly overestimated and the statistical uncertainties increase.

For the Gaussian variant TL4P$_{\mathrm{G}}$ (blue dashed line) of TL4P, however, the heat capacity $C_p(T,p_0)$ is almost everywhere close to the experimental data indicating that a correspondingly improved description may be expected from the hypothetical Gaussian variant TL6P$_{\mathrm{G}}$ mentioned above. Note that also the PMM model BK3[3] (orange line) with its Gaussian partial charges performs reasonable on $C_p(T,p_0)$.

S5 Effective Models

Effective mean-field type potentials for water try to compensate the fully (e.g. TIP4P/2005[9]) or partially (e.g. iAMOEBA[10]) neglected polarizability by an empirically optimal but nevertheless non-physical choice of the model parameters. In the parameterization of several of these models, notably in the empirical optimizations of TIP4P/2005 and of iAMOEBA, the whole experimental temperature density profile[2] $n_{\mathrm{exp}}(T, p_0)$ served as a target.

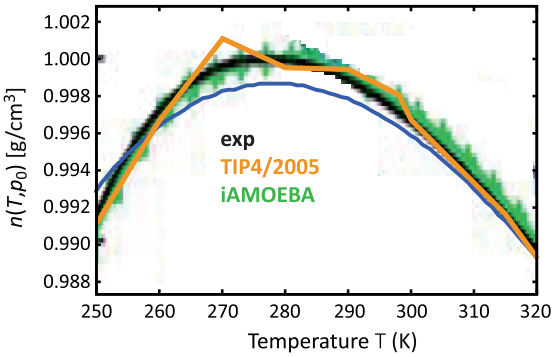


Figure S7: Comparison of computational results on $n(T, p_0)$ with experimental data for two effective models, which were empirically parameterized with the aim of reproducing the experimental curve. The close match of the experimental data (black dotted) by iAMOEBA[10] (green) and TIP5P/2005[9] (orange) solely demonstrates that one can find effective models capable of reproducing $n(T, p_0)$ (the quoted data were copied from the original papers). The noisy character of the TIP4P/2005 curve indicates insufficient statistics.

As documented by Figure S7 these models eventually did what they were taught to do, i.e. they reproduced the experimental input information. In the case of iAMOEBA the reproduction of $n(T, p_0)$ had to be expected, because it offers as many as 19 adjustable parameters, whereas in the case of TIP4P/2005 it was a remarkable finding.

On the other hand the success apparent in Figure S7 is accompanied by a non-physical behavior in other respects. For instance, the non-polarizable TIP4P/2005 model sizeably underestimates the dielectric constant by assigning a value of about 60 instead of 78, whereas the partially polarizable iAMOEBA model features a quadrupole moment, which deviates by an average of 12 % from the experimental one, and a polarizability overestimating the respective experimental value even by 24 %[10] demonstrating that iAMOEBA is an effective but not a physical model.

Effective models, like the two examples mentioned above, may be of practical use in simulations but are almost certainly of limited value for predictions on conditions, which are not covered by the reference data (e.g. water in proteins, electrolytes etc.). PMM models, in contrast, which are good enough to catch the essential physics of a water molecule in the liquid, should be transferable to all kinds of condensed phase environments as long as the assumption of linear response is valid. Furthermore they open the chance to identify the microscopic causes for the macroscopic properties of liquid water.

S6 PMM Models of comparable complexity

There have been numerous previous suggestions for PMM water models of a computational complexity similar to the one offered by TL4P, TL5P, and TL6P. In Refs. 11 and 8 we have compared a series of water properties predicted by these models for the standard liquid phase conditions of 300 K and 1 bar with the corresponding TL4P, TL5P, and TL6P predictions. In the main text of the current paper we have claimed particularly for PMM four- and five-point models that TL4P and TL5P perform at these conditions at least as good as other PMM models of comparable complexity. Furthermore, we have claimed that no other PMM model can reproduce the temperature-density profile $n_{\mathrm{exp}}(T, p_0)$ as well as TL6P. Here, we have quoted the slightly more complex BK3 model[3] as the best attempt so-far (cf. Fig.3).

Figure S8 shows for several selected PMM four-, five- and six-point models that their predictions on $n_{\mathrm{exp}}(T, p_0)$ are usually as poor as those of TL4P and TL5P and that none can compete with TL6P.

Consider first the left graph, which compares the $n(T, p_0)$ result of our PMM four-point model TL4P (green dashed) with the corresponding predictions of the similarly complex models COS/G3[12,13] and SWM4-NDP[13,14] (both green solid) as well as with the reference data from experiment[2] (black dashed) and TL6P (red dashed). Here, our TL4P prediction is seen to deviate from $n_{\mathrm{exp}}(T, p_0)$ even less than COS/G3 and SWM4-NDP. Note however, that SWM4-NDP reproduces the density at the

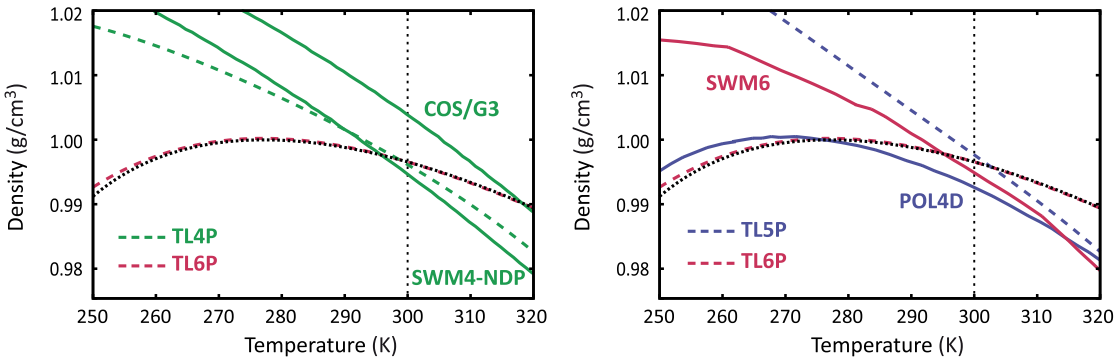


Figure S8: Comparison of computational results on $n(T, p_0)$ for various PMM models with experimental data[2] (black dotted line) and the TL6P prediction (red dashed line). Left: the PMM four-point models TL4P (green dashed), COS/G3,[12,13] and SWM4-NDP[13,14] (both green solid). Right: the PMM five-point models TL5P (blue dashed) and POL4D[15] (blue solid) as well as the PMM six-point model SWM6.[16]

liquid phase standard conditions about as well as TL4P, whereas COS/G3 sizably overestimates the density also at this point.

Turning to the right graph of Figure S8 we see that the PMM five-point model POL4D[15] (blue solid) predicts a density maximum near 269 K. Its predicted profile $n(T, p_0)$ intersects $n_{\mathrm{exp}}(T, p_0)$ near 276 K and considerably underestimates the experimental density at the liquid phase standard conditions. Disregarding the latter conditions, POL4D shows a much better overall performance on $n_{\mathrm{exp}}(T, p_0)$ than TL5P (blue dashed), which however still cannot compete with that of TL6P (red dashed) or BK3 (black dashed curve in Figure 3). The PMM six-point model SWM6[16] (red solid), which is the only model of a complexity similarly large as that of TL6P, in contrast, performs much worse on $n_{\mathrm{exp}}(T, p_0)$ than POL4D. In fact, the SWM6 prediction of $n(T, p_0)$ is very similar to that of TL4P. Correspondingly, also for SWM6 the prediction of the density at the liquid phase standard condition is likewise reasonable.

References

(1) Rapaport, D. C. *The Art of Molecular Dynamics Simulation*; Cambridge university press: Cambridge, UK, 2004.

(2) Kell, G. S. Precise Representation of Volume Properties of Water at One Atmosphere. *J. Chem. Eng. Data* **1967**, *12*, 66–69.

(3) Kiss, P. T.; Baranyai, A. A Systematic Development of a Polarizable Potential of Water. *J. Chem. Phys.* **2013**, *138*, 204507.

(4) Paricaud, P.; Předota, M.; Chialvo, A.; Cummings, P. From Dimer to Condensed Phases at Extreme Conditions: Accurate Predictions of the Properties of Water by a Gaussian Charge Polarizable Model. *J. Chem. Phys.* **2005**, *122*, 244511.

(5) Kiss, P. T.; Baranyai, A. Testing the Recent Charge-On-Spring Type Polarizable Water Models. II. Vapor-Liquid Equilibrium. *J. Chem. Phys.* **2012**, *137*, 194103.

(6) Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. Development of an Improved Four-Site Water Model for Biomolecular Simulations: TIP4P-EW. *J. Chem. Phys.* **2004**, *120*, 9665–9678.

(7) Wagner, W.; Pruß, A. The IAPWS Formulation 1995 for the Thermodynamic Properties of Ordinary Water Substance for General and Scientific Use. *J. Phys. Chem. Ref. Data* **2002**, *31*, 387–535.

(8) Tröster, P.; Lorenzen, K.; Tavan, P. Polarizable Six-Point Water Models from Computational and Empirical Optimization. *J. Phys. Chem. B, submitted* **2013**,

(9) Abascal, J. L. F.; Vega, C. A General Purpose Model for the Condensed Phases of Water: TIP4P/2005. *J. Chem. Phys.* **2005**, *123*, 234505.

(10) Wang, L.-P.; Head-Gordon, T.; Ponder, J. W.; Ren, P.; Chodera, J. D.; Eastman, P. K.; Martinez, T. J.; Pande, V. S. Systematic Improvement of a Classical Molecular Model of Water. *J. Phys. Chem. B* **2013**, *117*, 9956–9972.

(11) Tröster, P.; Lorenzen, K.; Schwörer, M.; Tavan, P. Polarizable Water Models from Mixed Computational and Empirical Optimization. *J. Phys. Chem. B* **2013**, *117*, 9486–9500.

(12) Yu, H.; van Gunsteren, W. Charge-On-Spring Polarizable Water Models Revisited: From Water Clusters to Liquid Water to Ice. *J. Chem. Phys.* **2004**, *121*, 9549–9564.

(13) Kiss, P. T.; Baranyai, A. Density Maximum and Polarizable Models of Water. *J. Chem. Phys.* **2012**, *137*, 084506.

(14) Lamoureux, G.; Harder, E.; Vorobyov, I.; Roux, B.; MacKerell, A. A Polarizable Model of Water for Molecular Dynamics Simulations of Biomolecules. *Chem. Phys. Lett.* **2006**, *418*, 245–249.

(15) Viererblová, L.; Kolafa, J. A Classical Polarizable Model for Simulations of Water and Ice. *Phys. Chem. Chem. Phys.* **2011**, *13*, 19925–19935.

(16) Yu, W.; Lopes, P. E. M.; Roux, B.; A. D. MacKerell, J. Six-Site Polarizable Model of Water Based on the Classical Drude Oscillator. *J. Chem. Phys.* **2013**, *138*, 034508.

3 *Résumé* und Ausblick

Wie insbesondere der vorangegangene Abschnitt 2.3 gezeigt hat, war mein Versuch, im Rahmen dieser Dissertation verbesserte PMM Modelle für Wasser zu konstruieren, letztlich mit der Entwicklung des 6-Punktmodells TL6P (Abschnitt 2.2) von großem Erfolg gekrönt. Damit wurde erstmals ein PMM Modellpotential für Wasser vorgelegt, das nicht nur so wichtige Eigenschaften wie die Dielektrizitätskonstante ε_0, die Diffusionskonstante, die isobare Wärmekapazität und die isotherme Kompressibilität bei den Standardbedingungen T_0 undp_0 sondern auch den isobaren thermischen Expansionskoeffizienten α_p und darüber hinaus sogar das gesamte Dichteprofil $n(T, p_0)$ für Temperaturen T aus dem Bereich [250 K, 320 K] mit bislang unübertroffener Genauigkeit vorhersagen kann.

Sogar Eigenschaften anderer Phasen konnten, wie aus der in Abschnitt 2.2 nachgedruckten *Supporting Information (SI)* zur Publikation [5] des 6-Punktmodells hervorgeht, mit recht großer Genauigkeit vorhergesagt werden. So unterschätzt TL6P die Schmelztemperatur T^{m} von Eis lediglich um höchstens 10 K (siehe Abschnitt S10 der *SI* zu [5]), was die bislang genauste Beschreibung dieser Größe durch ein PMM Modell darstellt. Obwohl TL6P dediziert für die flüssige Phase konstruiert wurde, ist die TL6P Beschreibung des zweiten Virialkoeffizienten, d.h. einer Gasphasen-Eigenschaft, für Temperaturen aus dem Bereich [300 K, 450 K] ähnlich vernünftig (siehe Abschnitt S4 der *SI* zu [5]).

Der skizzierte Erfolg bestätigt die Gültigkeit der Arbeitshypothese, welche die Leitidee bei der Konstruktion der TLν, $\nu = 3, \ldots, 6$, PMM-Modellpotentiale darstellte. Diese Hypothese bestand zum einen aus der Unterstellung, dass komplexe PMM Modellpotentiale für Moleküle oder für Fragmente von Molekülen aufgrund der Vielzahl festzulegender Parameter nicht, wie sonst üblich (vgl. z.B. [18]), durch empirische Gradientenabstiege auf hochdimensionalen Parameterräumen bestimmt werden können, da solche Suchen höchstwahrscheinlich in einem der vielen lokalen Minima steckenbleiben. Zum anderen bestand sie aus der Gewissheit, dass die elektrostatischen Eigenschaften von Molekülen, wie ihre Polarisierbarkeit und elektrostatische Signatur, QM Methoden wie der DFT zugänglich sind. Dabei verstehen wir unter der elektrostatischen Signatur eines Moleküls die mittleren nicht-polarisierten Anteile seiner Ladungsverteilung, welche die nicht-polarisierten Anteile seiner elektrostatischen Wechselwirkung mit anderen Molekülen erzeugen und im nicht-polarisierter Anteil seines Oberflächenpotentials kodiert ist.

Ausgehend von diesen Hypothesen sollten also lediglich die wenigen Parameter der Van der Waals Wechselwirkung von PMM Modellen für Wasser noch empirisch durch *weak coupling* PMM-MD Simulationen [125] bestimmt werden, während die elektrostatischen Eigenschaften vermittels der neuen DFT/PMM Technologie [9], die durch die Kombination des PMM-MD Programms IPHIGENIE [9, 10, 11] mit dem DFT Programm CPMD [12] im Verlauf meiner Dissertation verfügbar wurde, berechnet werden sollten.

Eine besondere Schwierigkeit bei der Berechnung der elektrostatischen Signatur des H_2O Moleküls ergab sich nun aus früheren Ergebnissen von Schropp und Tavan [16], die gezeigt hatten, dass das elektrische Feld $\mathbf{E}(\mathbf{r})$ auch im sehr kleinen Volumen v eines Wassermoleküls, das sich in wässriger Umgebung befindet, stark inhomogen ist.

Wäre nämlich $\mathbf{E}(\mathbf{r})$ homogen, d.h. nach $\mathbf{E}(\mathbf{r}) = \langle E(\mathbf{r})\rangle_v$ überall gleich dem Volumenmittel, dann gäbe es lediglich das induzierte Dipolmoment $\boldsymbol{\mu}^{\mathrm{i}} = \alpha\langle E(\mathbf{r})\rangle_v$ als Polarisationseffekt zu beachten und die elektrostatische Signatur eines in wässriger Umgebung gelösten Wassermoleküls wäre schon durch ein isoliertes Molekül zutreffend beschrieben. Also wäre es der üblichen Vakuum-Quantenchemie zugänglich. Inhomogene Felder $\mathbf{E}(\mathbf{r})$ können jedoch auch höhere Multipolmomente modifizieren, wobei die nicht-verschwindenden tensoriellen Gradienten [10] $\partial_{(n)}\mathbf{E}(\mathbf{r})$ n-ter Ordnung ($n = 1, 2, \ldots$) den polarisierten Anteil des $(n+1)$-ten Multipolmoments erzeugen. Falls diese Gradienten n-ter Ordnung im Ensemblemittel nicht verschwinden oder sehr klein sind, dann ist die elektrostatische Signatur eines isolierten Moleküls kein gutes Modell für seine Signatur in Lösung und kann entsprechend nicht aus Vakuum-Rechnungen abgeleitet werden. Wie groß die mittleren Gradienten $\partial_{(n)}\mathbf{E}(\mathbf{r})$ im Volumen eines Wassermolküls sind war vor meiner Arbeit nicht bekannt, weshalb das komplexe und in den Abschnitten 2.1 und 2.2 ausführlich geschilderte selbst-konsistente DFT/PMM Parametrisierungsverfahren zum Einsatz kam.

Ein für künftige Bemühungen zur Parametrisierung von PMM Modellen wichtiges und tröstliches Ergebnis enthält Tabelle 3 in [5] (nachgedruckt in Abschnitt 2.2). Hier zeigt sich nämlich, dass das Quadrupolmoment von H_2O, das von einer Vakuum-DFT Rechnung vorhergesagt wird, die elektrostatische Signatur der TL6P Flüssigphasen-Modelle relativ gut beschreibt. Ferner sind die Unterschiede des Gasphase-Modells TL6P$^{\mathrm{g}}$, das an die DFT Beschreibung eines isolierten H_2O Moleküls angepasst wurde, zu TL6P nicht groß, wie Tabelle S6 aus dem *SI* zu [5] zeigt. Man hätte also unter teilweiser Umgehung der selbst-konsistenten Parametrisierungs-Prozedur zumindest die Verteilungen der statischen Partialladungen ohne große Einbußen an Genauigkeit mit Vakuum-Quantenchemie erschließen können. Das vereinfacht die Konstruktion von PMM Kraftfeldern erheblich.

Andere Aspekte der selbst-konsistenten Prozedur scheinen dagegen unvermeidlich zu sein. So ist wohl die Bestimmung der Gaußbreiten σ aus Korrelationen (vgl. z.B. Abb. S4 aus dem *SI* zu [5]) der induzierten DFT/PMM Dipolmomente der DFT-Fragmente mit den induzierten PMM Dipolmomenten nach Glgen. (4) und (5) aus [5] unumgänglich. Auch müsste hier zusätzlich den Gaußbreiten σ_k der Partialladungen q_k von PMM Molekülen, die sich in unmittelbarer Nähe des DFT-Fragments befinden und in der IPHIGENIE/CPMD Schnittstelle [9] als Gaußladungen behandelt werden, größere Aufmerksamkeit gewidmet werden, als es in meinen Arbeiten geschehen ist, in denen diese Breiten einfach durch Dimer-Betrachtungen gesetzt wurden (cf. Section S1 of the *SI* to [5] and S6 of the *SI* to [5]). Man könnte die σ_k beispielsweise durch Berechnung von radialen Verteilungsfunktionen im DFT-PMM Setting in den gesamten Optimierungsprozess einbeziehen.

Ausblick

Im Vergleich zu früheren PMM Modellpotentialen hat TL6P eine sehr viel bessere Vorhersage für das Dichte-Temperatur-Maximum geliefert. Sie ist aber selbstverständlich auch nicht vollständig perfekt. Unterhalb von 260 K überschätzt TL6P beispielsweise die experimentelle Dichte etwas (siehe Abb. 3 in [7]). Es ist nicht ausgeschlossen, dass Details der TL6P Parametrisierung wie etwa die Wahl der radialen Verteilungsfunktion nach Soper et al. statt nach Skinner et al. wie bei der Modellvariante TL6P$^{\mathrm{Sk}}$ (siehe Abschnitt 1.3 in [5])

Da die Ableitung des von TL6P$^{\mathrm{sk}}$ vorhergesagten Temperatur-Dichte Profils bei 300 K etwas kleiner ist als der entsprechende Wert für TL6P, liegt diese Vermutung nahe.

Die TRE-Simulationen lieferten auch Einblicke in die Energetik der Wassermodelle bei verschiedenen Temperaturen, gemessen durch die Verdampfungswärme. Die Vorhersagen von TL6P für die Verdampfungswärme liegt bei allen Temperaturen oberhalb der experimentellen Messkurve. Auch bei dieser Observable wird der Verlauf für Temperaturen oberhalb von 260 K jedoch relativ gut getroffen. Bei Temperaturen unterhalb von 260 K, im Bereich in dem auch die Abweichungen in der Dichte größer werden, werden auch die Abweichungen der Verdampfungswärme von experimentellen Daten stärker. Auch hier stellt sich die Frage, ob das TL6P$^{\mathrm{sk}}$-Modell bei tiefen Temperaturen bessere Vorhersagen treffen kann.

Eine weitere Möglichkeit zur Verbesserung der Energetik von Wassermodellen besteht in der Verwendung Gaußscher Ladungsverteilungen statt Punktladungen. Natürlich hätten diese Ladungen den Nachteil eines größeren Rechenaufwandes, allerdings stellen sie ein bedeutend besseres Modell der ausgedehnten Ladungsverteilung eines Wassermoleküls dar. Durch die Entwicklung des TL4P$^{\mathrm{G}}$ Modells, einer Version des TL4P Modells mit Gaußschen Ladungsverteilungen, die in der *Supplementary Information* von Veröffentlichung [7] vorgestellt wurde, konnte der positive Einfluss Gaußscher Ladungen auf die Energetik, und damit auf Observablen wie Verdampfungswärme oder Wärmekapazität, nachgewiesen werden. Durch eine entsprechende Entwicklung eines TL6P$^{\mathrm{G}}$ Modells, eines TL6P Modells mit Gaußschen Ladungsverteilungen also, könnte dementsprechend eine noch genauere Beschreibung des Verhaltens von Wasser unter verschiedenen thermodynamischen Bedingungen gelingen.

Ein Sechspunktmodell mit Gaußschen Ladungsverteilungen zieht drei neue Parameter nach sich, die in die semiempirische Optimierungsstrategie integriert werden müssten, wollte man ein solches Modell parametrisieren. Da diese drei Parameter, die Breiten der Ladungsverteilungen auf den Wasserstoffen und den masselosen Ladungspunkten, kaum Einfluss auf das Oberflächenpotential und das Quadrupolmoment haben, ist die elektrostatische Signatur von der Einführung dieser Parameter kaum betroffen. Es müsste somit ein anderer Weg zur Optimierung der Breiten der Ladungsverteilungen gefunden werden. Bereits jetzt werden Punktladungen im Hybridszenario als Gaußladungen beschrieben, kommen sie dem DFT Fragment zu nahe. Wie in Abschnitt 2.1 erklärt wurde, wurden die Breiten dieser Verteilungen anhand des DFT/PMM, beziehungsweise des PMM/DFT Hybrid Dimers optimiert. Um eine vollständige Konsistenz zwischen PMM Modell und DFT Fragment zu erhalten, könnte man die Gaußschen Breiten der PMM Ladungsverteilungen auf ähnliche Weise optimieren. Eine weitere Möglichkeit bestünde darin, sie im Rahmen der *weak-coupling* Simulationen an eine experimentelle Zielgröße zu koppeln und somit empirisch anzupassen. Diese Vorgehensweise

widerspricht jedoch der Intention, möglichst wenig experimentelles Vorwissen in die Parameterisierung zu stecken, der die DFT/PMM Optimierungsmethode folgt. Die Tatsache, dass die TLνP Modellpotentiale vorrangig dafür entwickelt wurden in DFT/PMM Hybridsystemen als effektives Lösungsmittel eingesetzt zu werden, ist ein weiterer Grund dafür, eventuelle zusätzliche Parameter zur Optimierung der Eigenschaften solcher Hybridysteme einzusetzen.

Da das TL6P Modell mittels eines DFT/PMM Hybridverfahrens als Lösungsmittel für Biomoleküle konzipiert wurde, ist der wichtigste weitergehende Test ein Einsatz als solches. In der Arbeit von Klaehn et. al. [26], in der Infrarotspektren von Phosphatanionen berechnet wurden, wurde gezeigt, dass der Einsatz eines polarisierbaren Modellpotentials, das die Solvarstruktur um das gelöste Ion gut beschreiben kann, unerlässlich ist. Der Grund hierfür ist das starke elektrische Feld, welches ein einfach oder gar ein doppelt negativ geladenes Ion erzeugt, und welches den Dipol benachbarter Wassermoleküle stark beeinflussen muss. Dieser Einfluss kann durch das in dieser Arbeit verwendete TIP3P Modell, welches zudem noch über eine relativ flache Solvatstruktur verfügt [24], nicht berücksichtig werden. Da mittlerweile für die Spektren von gelösten Phosphatanionen, sowie über seine Solvatisierung, sehr gute DFT Referenzdaten vorliegen [27], liegt es auf der Hand, in diese Richtung weiter zu forschen und ein DFT/PMM Setting eines in TLνP Modellen gelösten Phosphatanions zu untersuchen.

Literaturverzeichnis

[1] BALL, Philip: *H_2O: A Biography of Water*. London, UK : Pheonix, Orion Books Ltd, 2000

[2] OPARIN, Aleksandr I.: *The Origin of Life*. DoverPublications, 2003

[3] EISENBERG, David S. ; KAUZMANN, Walter: *The structure and properties of water*. Clarendon Press Oxford, 2005

[4] BALL, Philip: Water: Water an enduring mystery. In: *Nature* 452, Nr. 7185, S. 291–292

[5] TRÖSTER, Philipp ; LORENZEN, Konstantin ; TAVAN, Paul: Polarizable Six-Point Water Models from Computational and Empirical Optimization. In: *J. Phys. Chem. B* 118 (2014), S. 1589–1602

[6] KELL, G.S̃.: Precise representation of volume properties of water at one atmosphere. In: *J. Chem. Eng. Data* 12 (1967), S. 66–69

[7] TRÖSTER, Philipp ; TAVAN, Paul: The Microscopic Physical Cause for the Density Maximum of Liquid Water. 5 (2014), S. 138–142

[8] TRÖSTER, Philipp ; LORENZEN, Konstantin ; SCHWÖRER, Magnus ; TAVAN, Paul: Polarizable water models from mixed computational and empirical optimization. In: *J. Phys. Chem. B* 117 (2013), S. 9486–9500

[9] SCHWÖRER, Magnus ; BREITENFELD, Benedikt ; TRÖSTER, Philipp ; LORENZEN, Konstantin ; TAVAN, Paul ; MATHIAS, Gerald: Coupling DFT to Polarizable Force Fields for Efficient and Accurate Hamiltonian Molecular Dynamics Simulations. In: *J. Chem. Phys.* 138 (2013), S. 244103

[10] LORENZEN, K. ; SCHWÖRER, M. ; TRÖSTER, P. ; MATES, S. ; TAVAN, P.: Optimizing the Accuracy and Efficiency of Fast Hierarchical Multipole Expansions for MD Simulations. In: *J. Chem. Theory Comput.* 8 (2012), S. 3628–3636

[11] BAUER, Sebastian ; TAVAN, Paul ; MATHIAS, Gerald: Electrostatics of proteins in dielectric solvent continua. II. Hamiltonian reaction field dynamics. In: *J. Chem. Phys.* 140 (2014), S. 104103

[12] CPMD V3.9, Copyright IBM Corp 1990-2008, Copyright MPI für Festkörperforschung Stuttgart 1997-2001, see also www.cpmd.org

[13] EICHINGER, M. ; GRUBMÜLLER, H. ; HELLER, H. ; TAVAN, P.: FAMUSAMM: An algorithm for rapid evaluation of electrostatic interactions in molecular dynamics simulations. In: *J. Comput. Chem.* 18 (1997), S. 1729–1749

[14] MATHIAS, Gerald ; EGWOLF, Bernhard ; NONELLA, Marco ; TAVAN, Paul: A Fast Multipole Method Combined with a Reaction Field for Long-Range Electrostatics in Molecular Dynamics Simulations: The Effects of Truncation on the Properties of Water. In: *J. Chem. Phys.* 118 (2003), S. 10847–10860

[15] EICHINGER, M ; TAVAN, P ; HUTTER, J ; PARRINELLO, M: A hybrid method for solutes in complex solvents: Density functional theory combined with empirical force fields. In: *J. Chem. Phys.* 110 (1999), S. 10452–10467

[16] SCHROPP, B. ; TAVAN, P.: The polarizability of point-polarizable water models: density functional theory/molecular mechanics results. In: *J. Phys. Chem. B* 112 (2008), S. 6233–6240

[17] SCHROPP, B. ; TAVAN, P.: Flexibility Does Not Change the Polarizability of Water Molecules in the Liquid. In: *J. Phys. Chem. B* 114 (2010), S. 2051–2057

[18] WANG, Lee-Ping ; HEAD-GORDON, Teresa ; PONDER, Jay W. ; REN, Pengyu ; CHODERA, John D. ; EASTMAN, Peter K. ; MARTINEZ, Todd J. ; PANDE, Vijay S.: Systematic Improvement of a Classical Molecular Model of Water. In: *J. Phys. Chem. B* 117 (2013), S. 9956–9972

[19] CLAPEYRON, Émile: *Mémoire sur la puissance motrice de la chaleur*. J. Gabay, 1834

[20] MASON, B.J.: The supercooling and nucleation of water. In: *Advances in Physics* 7 (1958), Nr. 26, S. 221–234

[21] KUHS, WF ; LEHMANN, MS: The structure of the ice Ih by neutron diffraction. In: *J. Phys. Chem.* 87 (1983), S. 4312–4313

[22] MASON, B.J.: The physics of water and ice. In: *Advances in Physics* 7 (1958), S. 1–3

[23] TANAKA, Hajime: Simple Physical Explanation of the Unusual Thermodynamic Behavior of Liquid Water. In: *Phys. Rev. Lett.* 80 (1998), S. 5750–5753

[24] JORGENSEN, W.L. ; CHANDRASEKHAR, J. ; MADURA, J.D. ; IMPEY, R.W. ; KLEIN, M.L.: Comparison of simple potential functions for simulating liquid water. In: *J. Phys. Chem.* 79 (1983), S. 926–935

[25] BERENDSEN, H. J. C. ; POSTMA, J. P. M. ; GUNSTEREN, W. F. ; DINOLA, A. ; HAAK, J. R.: Molecular Dynamics with Coupling to an External Bath. In: *J. Chem. Phys.* 81 (1984), S. 3684–3690

[26] KLÄHN, Marco ; MATHIAS, Gerald ; KÖTTING, Carsten ; NONELLA, Marco ; SCHLITTER, Jürgen ; GERWERT, Klaus ; TAVAN, Paul: IR spectra of phosphate ions in aqueous solution: predictions of a DFT/MM approach compared with observations. In: *J. Phys. Chem. A* 108 (2004), Nr. 29, S. 6186–6194

[27] VANDEVONDELE, Joost ; TRÖSTER, Philipp ; TAVAN, Paul ; MATHIAS, Gerald: Vibrational spectra of phosphate ions in aqueous solution probed by first-principles molecular dynamics. In: *J. Phys. Chem. A* 116 (2012), S. 2466–2474

[28] MURPHY, W.F.: The Rayleigh depolarization ratio and rotational Raman spectrum of water vapor and the polarizability components for the water molecule. In: *J. Chem. Phys.* 67 (1977), S. 5877–5882

[29] LAMOUREUX, G. ; MACKERELL JR, A.D. ; ROUX, B.: A Simple Polarizable Model of Water Based on Classical Drude Oscillators. In: *J. Chem. Phys.* 119 (2003), S. 5185–5197

[30] LAMOUREUX, G. ; HARDER, E. ; VOROBYOV, I.V. ; ROUX, B. ; MACKERELL, A.D.: A polarizable model of water for molecular dynamics simulations of biomolecules. In: *Chem. Phys. Lett.* 418 (2006), S. 245–249

[31] MECKE, R: Das Rotationsschwingungsspektrum des Wasserdampfes. I. In: *Z. Physik* 81 (1933), S. 313–331

[32] BENEDICT, W. S. ; GAILAR, N. ; PLYLER, Earle K.: Rotation-Vibration Spectra of Deuterated Water Vapor. In: *J. Chem. Phys.* 24 (1956), S. 1139–1165

[33] THIESSEN, W. E. ; NARTEN, A. H.: Neutron diffraction study of light and heavy water mixtures at 25° C. In: *J. Chem. Phys.* 77 (1982), S. 2656–2662

[34] CLOUGH, S.A. ; BEERS, Y. ; KLEIN, G.P. ; ROTHMAN, L.S.: Dipole moment of water from Stark measurements of H_20, HDO, and D_20. In: *J. Chem. Phys.* 59 (1973), S. 2254–2259

[35] DYKE, T.R. ; MUENTER, JS: Electric Dipole Moments of Low J States of H_2O and D_2O. In: *J. Chem. Phys.* 59 (1973), S. 3125–3127

[36] GREGORY, JK ; CLARY, DC ; LIU, K ; BROWN, MG ; SAYKALLY, RJ: The water dipole moment in water clusters. In: *Science* 275 (1997), S. 814–817

[37] SILVESTRELLI, P.L. ; BERNASCONI, M. ; PARRINELLO, M.: Ab initio infrared spectrum of liquid water. In: *Chem. Phys. Lett.* 277 (1997), S. 478–482

[38] SILVESTRELLI, P.L. ; PARRINELLO, M.: Structural, electronic, and bonding properties of liquid water from first principles. In: *J. Chem. Phys.* 111 (1999), S. 3572–3570

[39] SPRIK, Michiel: Hydrogen bonding and the static dielectric constant in liquid water. In: *J. Chem. Phys.* 95 (1991), S. 6762–6769

[40] VERHOEVEN, J. ; DYMANUS, A.: Magnetic Properties and Molecular Quadrupole Tensor of the Water Molecule by Beam-Maser Zeeman Spectroscopy. In: *J. Chem. Phys.* 52 (1970), S. 3222–3233

[41] NARTEN, A. H. ; LEVY, H. A.: Liquid Water: Molecular Correlation Functions from X-Ray Diffraction. In: *J. Chem. Phys.* 55 (1971), S. 2263–2269

[42] SOPER, A. K.: The Radial Distribution Functions of Water and Ice from 220 to 673 K and at Pressures up to 400 MPa. In: *Chem. Phys.* 258 (2000), S. 121–137

[43] HURA, G. ; SORENSON, J.M. ; GLAESER, R.M. ; HEAD-GORDON, T.: A high-quality X-ray scattering experiment on liquid water at ambient conditions. In: *J. Chem. Phys.* 113 (2000), S. 9140–9148

[44] FU, Ling ; BIENENSTOCK, Arthur ; BRENNAN, Sean: X-ray study of the structure of liquid water. In: *J. Chem. Phys.* 131 (2009), Nr. 23, S. 234702

[45] SOPER, AK: Joint Structure Refinement of X-Ray and Neutron Diffraction Data on Disordered Materials: Application to liquid water. In: *J. Phys.: Condens. Matter* 19 (2007), S. 335206

[46] SKINNER, Lawrie B. ; HUANG, Congcong ; SCHLESINGER, Daniel ; PETTERSSON, Lars G. ; NILSSON, Anders ; BENMORE, Chris J.: Benchmark Oxygen-Oxygen Pair-Distribution Function of Ambient Water from X-Ray Diffraction Measurements with a Wide Q-Range. In: *J. Chem. Phys.* 138 (2013), S. 074506

[47] ALLEN, M. P. ; TILDESLEY, D. J.: *Computer Simulations of Liquids*. Oxford : Clarendon, 1987

[48] VERLET, L.: Computer Experiments on Classical Fluids. 1. Thermodynamical Properties of Lennard-Jones Molecules. In: *Phys. Rev.* 159 (1967), S. 98–103

[49] MAS, Eric M. ; BUKOWSKI, Robert ; SZALEWICZ, Krzysztof: Ab Initio Three-Body Interactions for Water. I. Potential and Structure of Water Trimer. In: *J. Chem. Phys.* 118 (2003), S. 4386–4403

[50] MAS, Eric M. ; BUKOWSKI, Robert ; SZALEWICZ, Krzysztof: Ab Initio Three-Body Interactions for Water. II. Effects on Structure and Energetics of Liquid. In: *J. Chem. Phys.* 118 (2003), S. 4404–4413

[51] BEN-NAIM, A ; STILLINGER, FH: *Aspects of the Statistical-Mechanical Theory of Water*. Wiley-Interscience, New York, 1972

[52] KRÄUTLER, V. ; VAN GUNSTEREN, W.F. ; HÜNENBERGER, P.H.: A Fast SHAKE Algorithm to Solve Distance Constraint Equations for Small Molecules in Molecular Dynamics Simulations. In: *J. Chem. Phys.* 22 (2001), S. 501–508

[53] LONDON, Fritz: Zur Theorie und Systematik der Molekularkräfte. In: *Zeit. Physik* 63 (1930), S. 245–279

[54] LENNARD-JONES, J. E.: Cohesion. In: *Proc. R. Soc. Lond., A* 43 (1931), S. 461–482

[55] JONES, J.Ẽ.: On the Determination of Molecular Fields. II. From the Equation of State of a Gas. In: *Proc. R. Soc. Lond., A* 106 (1924), Nr. 738, S. 463–477

[56] BUCKINGHAM, R. A.: The Classical Equation of State of Gaseous Helium, Neon and Argon. In: *Proc. R. Soc. London, Ser. A* 168 (1938), Nr. 933, S. 264–283

[57] CHIALVO, A.A. ; CUMMINGS, P.T.: Simple transferable intermolecular potential for the molecular simulation of water over wide ranges of state conditions. In: *Fluid phase equilib.* 150 (1998), S. 73–81

[58] PARICAUD, P. ; PŘEDOTA, M. ; CHIALVO, A.A. ; CUMMINGS, P.T.: From dimer to condensed phases at extreme conditions: Accurate predictions of the properties of water by a Gaussian charge polarizable model. In: *J. Chem. Phys.* 122 (2005), S. 244511

[59] GUNSTEREN, Wilfred F. ; BERENDSEN, Herman J.: Computer Simulation of Molecular Dynamics: Methodology, Applications, and Perspectives in Chemistry. In: *Angew. Chem. Int. Ed.* 29 (1990), S. 992–1023

[60] TAVAN, Paul ; CARSTENS, Heiko ; MATHIAS, Gerald: Molecular Dynamics Simulation of Proteins and Peptides: Problems, Achievements, and Perspectives. In: BUCHNER, Johannes (Hrsg.) ; KIEFHABER, Thomas (Hrsg.): *Handbook of Protein Folding*. Wiley-VHC, Weinheim, 2004

[61] EWALD, Paul P.: Die Berechnung optischer und elektrostatischer Gitterpotentiale. In: *Ann. d. Phys* 369 (1921), S. 253–287

[62] DARDEN, Tom ; YORK, Darrin ; PEDERSEN, Lee: Particle Mesh Ewald: An N log (N) Method for Ewald Sums in Large Systems. In: *J. Chem. Phys.* 98 (1993), S. 10089–10092

[63] SMITH, Paul E. ; PETTITT, B M.: Ewald Artifacts in Liquid State Molecular Dynamics Simulations. In: *J. Chem. Phys.* 105 (1996), S. 4289–4293

[64] HÜNENBERGER, Philippe H. ; MCCAMMON, J A.: Effect of Artificial Periodicity in Simulations of Biomolecules under Ewald Boundary Conditions: a Continuum Electrostatics Study. In: *Biophys. Chem.* 78 (1999), S. 69–88

[65] WEBER, Wolfgang ; HÜNENBERGER, Philippe H. ; MCCAMMON, J A.: Molecular Dynamics Simulations of a Polyalanine Octapeptide under Ewald Boundary Conditions: Influence of Artificial Periodicity on Peptide Conformation. In: *J. Phys. Chem. B* 104 (2000), Nr. 15, S. 3668–3675

[66] KASTENHOLZ, Mika A. ; HÜNENBERGER, Philippe H.: Influence of Artificial Periodicity and Ionic Strength in Molecular Dynamics Simulations of Charged Biomolecules Employing Lattice-Sum Methods. In: *J. Phys. Chem. B* 108 (2004), Nr. 2, S. 774–788

[67] VILLARREAL, Marcos A. ; MONTICH, Guillermo G.: On the Ewald Artifacts in Computer Simulations. The Test-Case of the Octaalanine Peptide with Charged Termini. In: *Journal of Biomolecular Structure and Dynamics* 23 (2005), Nr. 2, S. 135–142

[68] NIEDERMEIER, C. ; TAVAN, P.: A Structure Adapted Multipole Method for Electrostatic Interactions in Protein Dynamics. In: *J. Chem. Phys.* 101 (1994), S. 734–748

[69] NIEDERMEIER, Christoph ; TAVAN, Paul: Fast Version of the Structure Adapted Multipole Method: Efficient Calculation of Electrostatic Forces in Protein Dynamics. In: *Mol. Simul.* 17 (1996), S. 57–66

[70] STERN, H.A. ; RITTNER, F. ; BERNE, B.J. ; FRIESNER, R.A.: Combined fluctuating charge and polarizable dipole models: Application to a five-site water potential function. In: *J. Chem. Phys.* 115 (2001), S. 2237–2251

[71] GUILLOT, B.: A Reappraisal of What We Have Learnt During Three Decades of Computer Simulations on Water. In: *J. Mol. Liquids* 101 (2002), S. 219–260

[72] MACKERELL, Alexander D. ; FEIG, Michael ; BROOKS, Charles L.: Extending the Treatment of Backbone Energetics in Protein Force Fields: Limitations of Gas-Phase Quantum Mechanics in Reproducing Protein Conformational Distributions in Molecular Dynamics Simulations. In: *J. Comput. Chem.* 25 (2004), Nr. 11, S. 1400–1415

[73] YU, H. ; GUNSTEREN, W.F. van: Accounting for Polarization in Molecular Simulation. In: *Comput. Phys. Commun.* 172 (2005), S. 69–85

[74] JORGENSEN, William L.: Special Issue on Polarization. In: *J. Chem. Theory Comput.* 3 (2007), S. 1877–1877

[75] LOPES, P.E.M. ; ROUX, B. ; MACKERELL, A.D.: Molecular Modeling and Dynamics Studies with Explicit Inclusion of Electronic Polarizability: Theory and Applications. In: *Theoret. Chem. Acc.* 124 (2009), S. 11–28

[76] LIU, Y.P. ; KIM, K. ; BERNE, BJ ; FRIESNER, R.A. ; RICK, S.W.: Constructing Ab Initio Force Fields for Molecular Dynamics Simulations. In: *J. Chem. Phys.* 108 (1998), S. 4739–4756

[77] MAPLE, Jon R. ; CAO, Yixiang ; DAMM, Wolfgang ; HALGREN, Thomas A. ; KAMINSKI, George A. ; ZHANG, Linda Y. ; FRIESNER, Richard A.: A Polarizable Force Field and Continuum Solvation Methodology for Modeling of Protein-Ligand Interactions. In: *J. Chem. Theory Comput.* 1 (2005), S. 694–715

[78] SWART, M ; VAN DUIJNEN, P T.: DRF90: A Polarizable Force Field. In: *Mol. Simul.* 32 (2006), S. 471–484

[79] MA, Buyong ; LII, Jenn-Huei ; ALLINGER, Norman L.: Molecular Polarizabilities and Induced Dipole Moments in Molecular Mechanics. In: *J. Comput. Chem.* 21 (2000), S. 813–825

[80] BÖTTCHER, Carl Johan F. ; BELLE, Oenes C. ; BORDEWIJK, Paul ; RIP, Arie: *Theory of Electric Polarization*. Bd. 1. Elsevier Amsterdam, 1973

[81] AHLSTRÖM, Peter ; WALLQVIST, Anders ; ENGSTRÖM, Sven ; JÖNSSON, Bo: A Molecular Dynamics Study of Polarizable Water. In: *Mol. Phys.* 68 (1989), S. 563–581

[82] YU, H. ; HANSSON, T. ; VAN GUNSTEREN, W.F.: Development of a simple, self-consistent polarizable model for liquid water. In: *J. Chem. Phys.* 118 (2003), S. 221–234

[83] APPLEQUIST, Jon ; CARL, James R. ; FUNG, Kwok-Kueng: Atom dipole interaction model for molecular polarizability. Application to polyatomic molecules and determination of atom polarizabilities. In: *J. Am. Chem. Soc.* 94 (1972), S. 2952–2960

[84] THOLE, B.T.: Molecular Polarizabilities Calculated with a Modified Dipole Interaction. In: *Chem. Phys.* 59 (1981), S. 341–350

[85] VAN DUIJNEN, Piet T. ; SWART, Marcel: Molecular and atomic polarizabilities: Thole's model revisited. In: *J. Phys. Chem. A* 102 (1998), S. 2399–2407

[86] ELKING, Dennis ; DARDEN, Tom ; WOODS, Robert J.: Gaussian induced dipole polarization model. In: *J. Comput. Chem.* 28 (2007), Nr. 7, S. 1261–1274

[87] DRUDE, Paul: *The Theory of Optics*. DoverPublications. com, 1933

[88] BORN, Max ; HUANG, Kun ; BORN, Max ; BORN, Max ; BRITAIN, Great: *Dynamical Theory of Crystal Lattices*. Clarendon Press Oxford, 1954

[89] STRAATSMA, T. P. ; MCCAMMON, J. A.: Molecular Dynamics Simulations with Interaction Potentials Including Polarization Development of a Noniterative Method and Application to Water. In: *Mol. Simul.* 5 (1990), S. 181–192

[90] MITCHELL, PJ ; FINCHAM, David: Shell model simulations by adiabatic dynamics. In: *J. Phys.: Condens. Matter* 5 (1993), S. 1031–1038

[91] LINDAN, PJD: Dynamics with the shell model. In: *Mol. Simul.* 14 (1995), S. 303–312

[92] LAMOUREUX, Guillaume ; ROUX, Benoît: Absolute hydration free energy scale for alkali and halide ions established from simulations with a polarizable force field. In: *J. Phys. Chem. B* 110 (2006), S. 3308–3322

[93] WHITFIELD, Troy W. ; VARMA, Sameer ; HARDER, Edward ; LAMOUREUX, Guillaume ; REMPE, Susan B. ; ROUX, Benoit: Theoretical study of aqueous solvation of K+ comparing ab initio, polarizable, and fixed-charge models. In: *J. Chem. Theory Comput.* 3 (2007), S. 2068–2082

[94] ANISIMOV, Victor M. ; VOROBYOV, Igor V. ; ROUX, Benoît ; MACKERELL, Alexander D.: Polarizable empirical force field for the primary and secondary alcohol series based on the classical Drude model. In: *J. Chem. Theory Comput.* 3

[95] LOPES, Pedro E. ; LAMOUREUX, Guillaume ; ROUX, Benoit ; MACKERELL, Alexander D.: Polarizable empirical force field for aromatic compounds based on the classical drude oscillator. In: *J. Phys. Chem. B* 111 (2007), S. 2873–2885

[96] HARDER, Edward ; ANISIMOV, Victor M. ; WHITFIELD, Troy ; MACKERELL, Alexander D. ; ROUX, Benoît: Understanding the dielectric properties of liquid amides from a polarizable force field. In: *J. Phys. Chem. B* 112 (2008), S. 3509–3521

[97] LU, Zhenyu ; ZHANG, Yingkai: Interfacing ab initio quantum mechanical method with classical Drude osillator polarizable model for molecular dynamics simulation of chemical reactions. In: *J. Chem. Theory Comput.* 4 (2008), S. 1237–1248

[98] RAPPE, Anthony K. ; GODDARD III, William A.: Charge equilibration for molecular dynamics simulations. In: *J. Phys. Chem.* 95 (1991), S. 3358–3363

[99] BARKER, JA ; WATTS, RO: Structure of water; A Monte Carlo calculation. In: *Chem. Phys. Lett.* 3 (1969), S. 144–145

[100] RAHMAN, Aneesur ; STILLINGER, Frank H.: Molecular dynamics study of liquid water. In: *J. Chem. Phys.* 55 (1971), S. 3336–3359

[101] BERENDSEN, HJC ; POSTMA, JPM ; VAN GUNSTEREN, WF ; HERMANS, J.: Interaction models for water in relation to protein hydration. In: *Intermol. Forces* 11 (1981), S. 331–342

[102] BERENDSEN, HJC ; GRIGERA, JR ; STRAATSMA, TP: The missing term in effective pair potentials. In: *J. Phys. Chem.* 91 (1987), S. 6269–6271

[103] ABASCAL, JLF ; VEGA, C.: A general purpose model for the condensed phases of water: TIP4P/2005. In: *J. Chem. Phys.* 123 (2005), S. 234505

[104] ALEJANDRE, José ; CHAPELA, Gustavo A. ; SAINT-MARTIN, Humberto ; MENDOZA, Noé: A Non-Polarizable Model of Water that Yields the Dielectric Constant and the Density Anomalies of the Liquid: TIP4Q. In: *Phys. Chem. Chem. Phys.* 13 (2011), S. 19728–19740

[105] MAHONEY, M.W. ; JORGENSEN, W.L.: A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. In: *J. Chem. Phys.* 112 (2000), S. 8910–8922

[106] RICK, Steven W.: A reoptimization of the five-site water potential (TIP5P) for use with Ewald sums. In: *J. Phys. Chem.* 120 (2004), S. 6085–6093

[107] NADA, H. ; EERDEN, J.P.J.M. van d.: An intermolecular potential model for the simulation of ice and water near the melting point: A six-site model of H_2O. In: *J. Chem. Phys.* 118 (2003), S. 7401–7413

[108] KISS, Péter T. ; BARANYAI, András: Density maximum and polarizable models of water. In: *J. Chem. Phys.* 137 (2012), Nr. 8, S. 084506

[109] BARANYAI, A. ; KISS, P.T.: Polarizable model of water with field-dependent polarization. In: *J. Chem. Phys.* 135 (2011), S. 234110

[110] KISS, Péter T ; BARANYAI, András: A systematic development of a polarizable potential of water. In: *J. Chem. Phys.* 138 (2013), S. 204507

[111] UFIMTSEV, Ivan S. ; LUEHR, Nathan ; MARTINEZ, Todd J.: Charge transfer and polarization in solvated proteins from ab initio molecular dynamics. In: *J. Phys. Chem. Lett.* 2 (2011), S. 1789–1793

[112] BARANYAI, A. ; KISS, P.T.: A transferable classical potential for the water molecule. In: *J. Chem. Phys.* 133 (2010), S. 144109

[113] HOHENBERG, Pierre ; KOHN, Walter: Inhomogeneous electron gas. In: *Phys. Rev.* 136 (1964), Nr. 3B, S. B864–B871

[114] KOHN, Walter ; SHAM, Lu J.: Self-Consistent Equations Including Exchange and Correlation Effects. In: *Phys. Rev.* 140 (1965), Nov, S. A1133–A1138

[115] DREIZLER, Reiner M. ; GROSS, Eberhard K.: *Density Functional Theory: An Approach to the Quantum Many-Body Problem*. Bd. 184. Springer-Verlag Berlin, 1990

[116] 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. ; JR. ; 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, O. ; FORESMAN, J. B. ; ORTIZ, J. V. ; CIOSLOWSKI, J. ; FOX, D. J.: Gaussian 09, Revision A.0. In: *Gaussian, Inc., Wallingford CT* (2009)

[117] VANDEVONDELE, Joost ; KRACK, Matthias ; MOHAMED, Fawzi ; PARRINELLO, Michele ; CHASSAING, Thomas ; HUTTER, Jürg: Quickstep: Fast and Accurate Density Functional Calculations Using a Mixed Gaussian and Plane Waves Approach. In: *Comput. Phys. Commun.* 167 (2005), S. 103–128

[118] STEWART, AT: Momentum Distribution of Metallic Electrons by Positron Annihilation. In: *Can. J. Phys.* 35 (1957), S. 168–183

[119] TROULLIER, N. ; MARTINS, J.L.: Efficient pseudopotentials for plane-wave calculations. In: *Phys. Rev. B* 43 (1991), S. 1993–2006

[120] BECKE, A. D.: Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. In: *Phys. Rev. A* 38 (1988), S. 3098–3100

[121] PERDEW, John P.: Density-functional approximation for the correlation energy of the inhomogeneous electron gas. In: *Phys. Rev. B* 33 (1986), S. 8822–8824

[122] STEVENS, PJ ; DEVLIN, FJ ; CHABALOWSKI, CF ; FRISCH, Michael J.: Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. In: *J. Phys. Chem.* 98 (1994), S. 11623–11627

[123] WARSHEL, A. ; KARPLUS, M.: Semiclassical trajectory approach to photoisomerization. In: *Chem. Phys. Lett.* 32 (1975), S. 11–17

[124] SENN, Hans M. ; THIEL, Walter: QM/MM methods for biological systems. In: *Atomistic approaches in modern biology*. Springer, 2007, S. 173–290

[125] BERWEGER, C.D. ; GUNSTEREN, W.F. van ; MÜLLER-PLATHE, F.: Force field parametrization by weak coupling. Re-engineering SPC water. In: *Chem. Phys. Lett.* 232 (1995), S. 429–436

Danksagung

Bedanken möchte ich mich allen voran bei meinem Doktorvater Prof. Dr. Paul Tavan für seine Geduld, seinen Einsatz und seinen guten Rat.

Meinen lieben Kollegen, Gerald, Sebi, Magnus, Christoph und Konstantin danke ich für die vielen konstruktiven, aber auch amüsanten Gespräche und ganz allgemein für die schöne Zeit, die ich am BMO erleben durfte. Es gibt viele lustige Geschichten, und ich hoffe wir werden sie uns immer mal wieder erzählen. Extra bedanken muss ich mich bei Konstantin. Ich könnte mir keinen besseren Büropartner vorstellen. Danke für soviele Gespräche, Ideen und Verrücktheiten.

Danke auch an meine Familie, die mir in dieser Zeit noch viel mehr ans Herz gewachsen ist, als es vorher schon der Fall war. Ich bin so glücklich, dass Ihr für mich da seid.

Erklärung

Mit folgender Unterschrift erkläre ich, dass ich die vorliegende Arbeit selbständig verfasst, sowie keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

München, den 24. April 2014

Philipp Tröster