An Approach to Developing a Force Field for Molecular Simulation of Martensitic Phase Transitions between Phases with Subtle Differences in Energy and Structure

ARTICLE *in* JOURNAL OF THE AMERICAN CHEMICAL SOCIETY · FEBRUARY 2004

Impact Factor: 12.11 · DOI: 10.1021/ja0356131 · Source: PubMed

3 AUTHORS, INCLUDING:

Jamshed Anwar
Lancaster University
74 PUBLICATIONS **1,598** CITATIONS
SEE PROFILE

Julian David Gale
Curtin University
270 PUBLICATIONS **14,562** CITATIONS
SEE PROFILE

J|A|C|S
ARTICLES
Published on Web 12/12/2003

An Approach to Developing a Force Field for Molecular Simulation of Martensitic Phase Transitions between Phases with Subtle Differences in Energy and Structure

Sigrid C. Tuble, Jamshed Anwar,* and Julian D. Gale†

Contribution from the Computational Pharmaceutical Sciences Laboratory, Department of Pharmacy, King's College London, Franklin-Wilkins Building, 150 Stamford Street, London SE1 9NN, U.K.

Received April 14, 2003; E-mail: jamshed.anwar@kcl.ac.uk

Abstract: D,L-Norleucine is one of only a few molecules whose crystals exhibit a martensitic or displacive-type phase transformation where the emerging phase shows a topotaxial relationship with the parent phase. The molecular mechanism for such phase transformations, particularly in molecular crystals, is not well understood. Crystalline phases that exhibit displacive phase transitions tend to be very similar in structure and energy. Consequently, the development of a force field for such phases is challenging as the phase behavior is determined by subtle differences in their lattice energies and entropies. We report an approach for developing a force field for such phases with an application to D,L-norleucine. The proposed procedure includes calculation of the phase diagram of the crystalline phases as a function of temperature to identify the best force field. D,L-Norleucine also presents an additional problem since in the solid state it exists as a zwitterion that is unstable in vacuo and therefore cannot be characterized using high-level ab initio calculations in the gas phase. However, a stable zwitterion could be obtained using Onsager's reaction-field continuum model for a solvent (SCRF) using both Hartree−Fock and density functional theory. A number of force fields and the various sets of partial charges obtained from the SCRF calculations were screened for their ability to reproduce the crystal structures of the two known phases, α and β, of D,L-norleucine. Selected parameter sets were then employed in free energy minimizations to identify the best set on the basis of a correct prediction of the $\alpha-\beta$ phase transition. The Williams' nonbonded parameters combined with partial charges from SCRF-Polarized Continuum Model calculation were found to reproduce the structures of the phases accurately and also maintained their stability in extended molecular dynamics simulations in the Parrinello−Rahman constant stress ensemble. Moreover, we were also able to successfully simulate the phase transformation of the β- to the α-phase. The identified force field should enable detailed studies of the phase transformations exhibited by crystals of D,L-norleucine and hence enhance our understanding of martensitic-type transformations in molecular crystals.

Introduction

D,L-Norleucine or 2-aminohexanoic acid ($C_6H_{13}NO_2$) is one of a very few known molecular substances that exhibit a fast, single crystal to single crystal displacive phase transformation with an orientational relationship between the parent and the emerging daughter crystal.[1] Other molecules whose crystals exhibit a similar behavior include hexamethylbenzene,[1−5] 1,2,4,5 tetrachlorobenzene,[6,7] s-triazine,[8] and aniline hydrobromide.[9,10] Such transformations are often termed martensitic, which traditionally implies a transformation velocity of the order of the speed of sound, cooperative behavior in the bulk crystal, and topotaxy between the parent and the daughter crystals. However, this definition is constantly changing as more probing investigations reveal that many of these transformations occur by nucleation and growth[11] and that the kinetics depend on the excess temperature, i.e., extent of superheating or supercooling,[12] or excess pressure.[13]

D,L-Norleucine can exist in three crystalline forms, α, β, and γ. The forms β and γ have also been referred to as the low-temperature phase (LTP) and high-temperature phase (HTP), respectively, by Mnyukh.[1] In terms of stability the β form is the stable form at 120 K,[14] the α form at ambient conditions,

† Present address: Nanochemistry Research Institute, Department of Applied Chemistry, Curtin University of Technology, P.O. Box U1987, Perth 6845, Western Australia.

(1) Mnyukh, Yu. V.; Panfilova, N. A.; Petropavlov, N. N.; Uchvatova, N. S. *J. Phys. Chem. Solids* **1975**, *36*, 127−144.
(2) Vatulev, V. N.; Prithotko, A. I. *Fiz. Tverd. tela* **1965**, *7*, 42.
(3) Vatulev, V. N.; Prithotko, A. I. *Fiz. tverd. tela* **1967**, *9*, 3475.
(4) Brockway, L.; Robertson, J. L. *J. Chem. Soc.* **1939**, 1324.
(5) Watanabe, T.; Saito, Y.; Chihara, H. *Sci. Papers Osaka Univ.* **1950**, No. 2, 9.
(6) Herbstein, F. H. *Acta Crystallogr.* **1965**, *18*, 997.
(7) Dean, C.; Pollak, M. *Acta Crystallogr.* **1958**, *11*, 710.
(8) Coppens, P.; Sabine, T. M. *Mol. Cryst.* **1968**, *3*, 507.
(9) Taguchi, I. *Bull. Chem. Soc. Jpn.* **1961**, *34*, 392.
(10) Nitta, I.; Watanabe, T.; Taguchi, I. *Bull. Chem. Soc. Jpn.* **1961**, *34*, 1405.
(11) Mnyukh, Y. V. *Mol. Cryst. Liq. Cryst.* **1979**, *52*, 163.
(12) Mnyukh, Y. V.; Kitaigorodskiy, A. I.; Asadov, Y. G. *Soviet Physics-JETP* **1965**, *21*, 12.
(13) RezaiFard, A. R.; Anwar, J.; Clark, S. M. *Mat. Sci. Forum* **1996**, *228*, 375−381.
(14) Harding, M. M.; Kariuki, B. M.; Williams, L.; Anwar, J. *Acta Crystallogr.* **1995**, *B51*, 1059−1062.

10.1021/ja0356131 CCC: $27.50

Table 1. Crystallographic Properties of the α- and β-Phase of D,L-Norleucine[a]

	α D,L-norleucine	β D,L-norleucine
crystal class	monoclinic	monoclinic
space group	$P2_1/c$	$C2/c$
molecules in unit cell	4	8
a (Å)	16.382	31.067
b (Å)	4.737	4.717
c (Å)	9.907	9.851
β (deg)	104.68	91.37
density (Mg m^{-3})	1.171	1.207

[a] The structure determination of the α-phase was at room temperature[14] while that of the β-phase was at 120 K.[15]

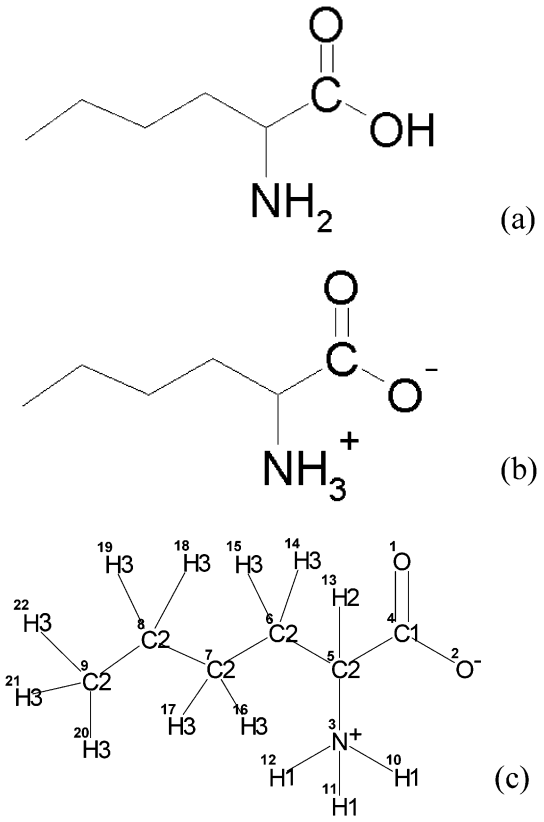

Figure 1. Molecular structure of (a) the non-charge-separated form, (b) the zwitterion, and (c) the zwitterionic form showing the force field atom types of norleucine.

and the γ form above 390 K.[1,15] The crystal structures of the α and the β forms have been determined,[14,15] and the lattice parameters are given in Table 1. Both phases crystallize in monoclinic space groups: the α form in $P2_1/c$ and the β form in $C2/c$ (which is equivalent to $I2/a$ through a transformation of the axes). In both of these phases, the D,L-norleucine molecule exists in a zwitterionic form (Figure 1b) with an almost identical geometry, while the crystal structure consists of bilayers in which the polar groups meet in the middle with the aliphatic chains on the outside (Figure 2). The individual layers are tightly linked to each other by hydrogen bonds involving the $-COO^-$ and $-NH_3^+$ in a manner typical of amino acids. The outer surfaces of the bilayers comprising the ends of the aliphatic chains form an approximate close-packed array and stack against the next bilayer with the interaction between the bilayers being of the van der Waals type only. In terms of the molecular packing, the only distinction between the two phases arises from the translation of an alternate bilayer unit by half a cell length along the *b*-axis. Thus in the α form, the packing of the bilayer units is of the form A−A−A−A, while for the β form it is A−B−A−B, with the cell parameter of the axis perpendicular to the bilayer units (the *a*-axis) for the β form being approximately twice that for the α form.

The molecular mechanism of first-order phase transformations still remains essentially a mystery with little direct experimental evidence to support any hypothesis. A key issue is the nature of the molecular processes taking place at the interface. Is the interface amorphous?[16,17] Is the transformation process essentially analogous to crystallization of the new phase, but from a solid medium?[1,11] What determines whether the parent and the daughter phase exhibit a topotaxial relationship? Our current ideas of the possible molecular mechanism come largely from diffraction studies. Such studies, however, provide absolutely no information regarding the structure of the interface. An approach that is increasingly yielding important insights into phase transformation phenomena is molecular simulation.[18−30] This methodology, in particular the molecular dynamics method employing the Parrinello−Rahman simulation cell,[18,19] has the potential to make the molecular processes occurring in phase transformations transparent.

Molecular simulations of realistic systems require that the empirically derived interaction potential characterizing the atomic interactions is sufficiently accurate to be able to reproduce the structural and thermodynamic properties of the molecule(s) of interest. While a choice of potential energy functions and associated parameters (usually termed the force field) are readily available,[31−34] one cannot be certain a priori whether a particular set will be transferable to the molecule of interest. It is therefore necessary to screen the various force fields to identify a set that reproduces the important structural and thermodynamic properties. In some instances it may be essential to optimize the parameters using the best available parameters as a starting point. We find that this approach is unlikely to be adequate for developing a force field to investigate simulations of displacive phase transitions in molecular crystals. Crystalline phases that exhibit displacive phase transitions tend, by their very nature, to be quite similar in terms of their structure and energy. Consequently, the phase transition behavior is deter-

(15) Dalhus, B.; Gorbitz, C. H. *Acta Crystallogr.* **1996**, 1761−1764.
(16) Bradley, R. S.; Hartshorne, N. H.; Thackray, M. *Nature* **1954**, *173*, 400.
(17) Bradley, R. S. *J. Phys. Chem.* **1956**, *60*, 1347.
(18) Parrinello, M.; Rahman, A. *J. Phys.* **1981**, *42 (NC6)*, 511−515.
(19) Parrinello, M.; Rahman, A. *J. Appl. Phys.* **1981**, *52*, 7182.
(20) Nosé, S.; Klein, M. *Phys. Rev. Lett.* **1983**, *50*, 1207.
(21) Refson, K. *Physica* **1985**, *131B*, 256.
(22) Car, R.; Parrinello, M. *Phys. Rev. Lett.* **1985**, *55*, 2471−2474.
(23) Sprik, M.; Cheng, A. L.; Klein, M. L. *Phys. Rev Lett.* **1992**, *69*, 11, 1660−1663.
(24) Boutin, A.; Simon, J.; Fuchs, A. *Mol. Phy.* **1994**, *81*, 1165.
(25) Hua, L.; Pawley, G. *Z. Kristallogr.* **1994**, *209*, 221.
(26) Fuchar, P.; Chiarotti, G. L.; Bernasconi, M.; Tosatti, E.; Parrinello, M. *Europhys. Lett.* **1994**, *36*, 345.
(27) Scandolo, S.; Bernasconi, M.; Chiarotti, G. L.; Focher, P.; Tosatti, E. *Phys. Rev. Lett.* **1995**, *74*, 4015.
(28) Devani, S.; Anwar, J. *J. Chem. Phys.* **1996**, *105*, 3215.
(29) De Vita, A.; Galli, G.; Canning, A.; Car, R. *Nature* **1996**, *379*, 523−526.
(30) Benoit, M.; Bernasconi, M.; Parrinello, M. *Phys. Rev. Lett.* **1996**, 2934−2936.
(31) Williams, D. E. *J. Comput. Chem.* **2001**, *22*, 11, 1154−1166.
(32) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, Jr. K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(33) Foloppe, N.; MacKerell, A. D., Jr. *J. Comput. Chem.* **2000**, *21*, 86−104.
(34) van Gunsteren, W. F.; Daura, X.; Mark, A. E. The GROMOS force field. In *Encyclopedia of Computational Chemistry*; Wiley: New York, 1998; Vol. 2, pp 1211−1216.

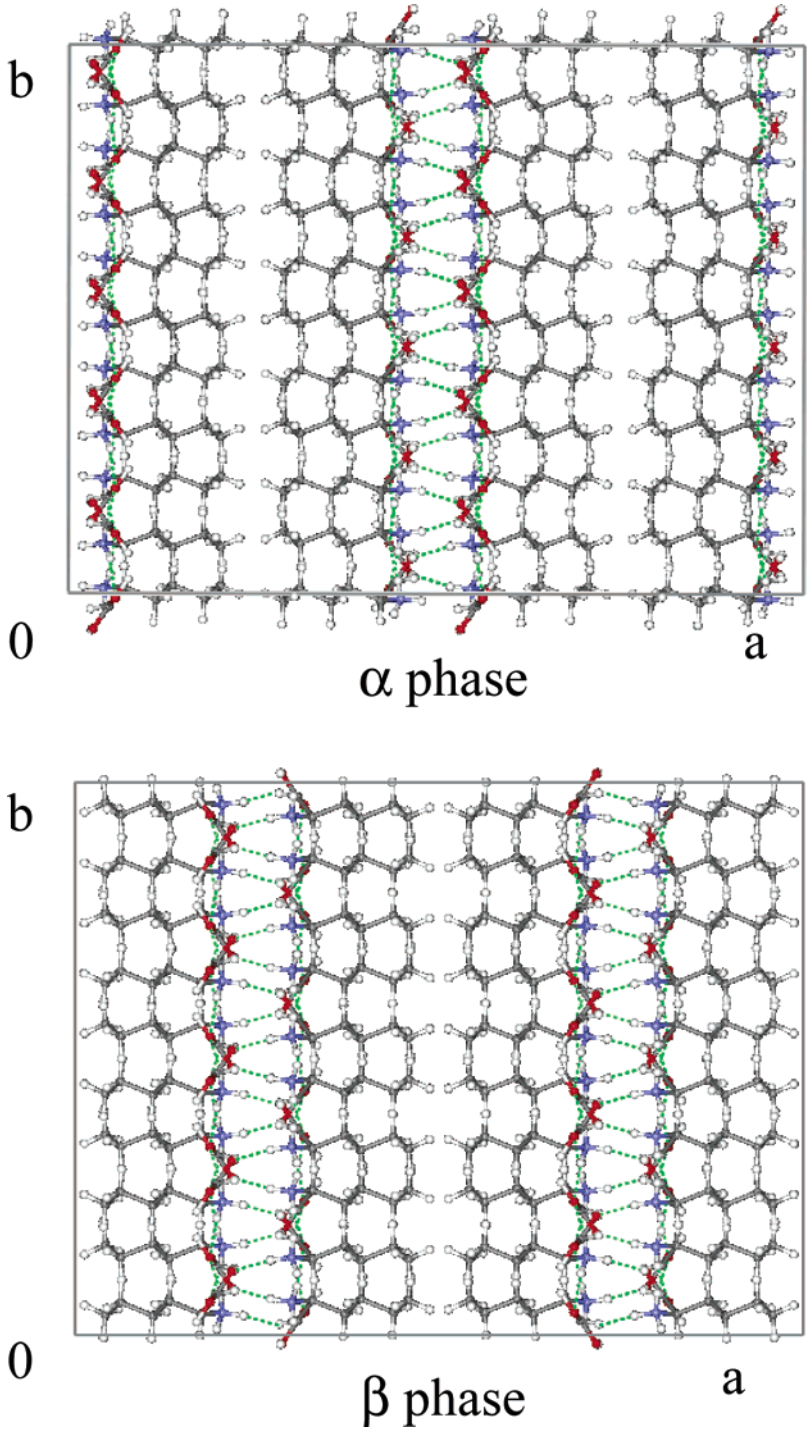

Figure 2. Views of the crystal structure of the α- and the *β*-phase of D,L-norleucine looking down the *c*-axis.

mined by a delicate balance arising from the enthalpy and entropy differences between the crystalline phases. While the crystalline structures can be distinctly characterized, the precision of the experimentally determined lattice energies tends to be too gross to discriminate between the lattice energies of the phases. Without accurate experimental values of the lattice energies, potential parameters cannot be reliably optimized to reproduce the phase equilibria of the phases of interest.

An additional problem in developing a force field for the crystalline phases of D,L-norleucine stems from the fact that this molecule exists as a zwitterion in the solid state. This can present a challenge to deriving the required partial charges that are integral to the force field. A straightforward application of the normal approach to deriving partial charges, i.e., fitting the charges to the electrostatic potential of the molecule obtained from a gas-phase calculation,[35] may not be possible since an accurate ab initio calculation may yield the non-charge-separated molecule (Figure 1a) and not the required zwitterion (Figure 1b) as has previously been discovered for glycine.[36,37]

We present here an approach for developing a force field for molecular simulation of displacive phase transitions in molecular

(35) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1984**, *5*, 129.
(36) Jensen, J. H.; Gordon, M. S. *J. Am. Chem. Soc.* **1995**, *117*, 8159.
(37) Ding, Y.; Korgh-Jespersen, K. K. *Chem. Phys. Lett.* **1992**, *199*, 261.

crystals with an application to D,L-norleucine. The approach, in addition to optimizing the crystal structures, involves the characterization of the phase equilibrium behavior using free energy calculations to identify the optimum set of potential parameters. An additional contribution is the development of a procedure for deriving partial charges for a zwitterionic form of a molecule. The identified force field for D,L-norleucine has been shown to reproduce not only the static structures but also the stability of the known crystalline forms in extended molecular dynamics simulations, and more remarkably the transformation between two of the crystalline phases.

Methodology

Derivation of Partial Charges. Partial charges are typically derived using the electrostatic potential (ESP) fitting method, where the charges are fitted to the electrostatic potential of a molecule obtained from a gas-phase calculation. The norleucine molecule exists as a zwitterion in the two known crystalline phases and is likely to exist in the neutral form in the gas phase. Even if the zwitterion were to be stable in the gas phase, one may not be able to reproduce this stability in an ab initio gas-phase calculation, given that calculations on glycine[36,37] reveal that the optimized structure can be different depending on the level of the calculation. It was observed that the zwitterion of glycine could be obtained at the Hartree−Fock (HF) level, but not using the 2nd order Møller−Plesset perturbation theory (MP2) coupled with a large basis set. In view of these considerations, the stability of the zwitterionic form of norleucine was examined using HF, MP2, and Density Functional Theory (DFT)[38] with the hybrid B3LYP method (comprising Becke's 3-parameter exchange functional[39] and the Lee−Yang−Parr correlation functional[40]). Basis sets employed included 6-31G*, 6-311G*, 6-31G**, 6-311G**, and 6-311++G**.

Should the zwitterion be unstable in a gas-phase calculation, the next most appropriate study would be to calculate the charge density for the molecule in a crystalline environment, that is, through the use of a periodic calculation. While such calculations are now routinely carried out, the derivation of ESP charges from a periodic system calculation is not straightforward due to the significant overlap of the electron density−nor is there a totally acceptable scheme for partitioning the overlapping density. The Mulliken charges can still be derived, but these can be unreliable in molecular simulations since they are rather dependent on the basis set used. An alternative is to employ the Born effective charge tensor to extract averaged charges. However, these charges are only formally appropriate for the electrical response, rather than the static potential.

The other two options are either to carry out the calculation for the molecule in a continuum environment or to perform a cluster calculation with the central molecule surrounded by other explicit molecules. The use of explicit molecules introduces additional variables, as one would need to explore the effect of the number and the positions of the added molecules, or even, ultimately, to perform an integration over the full configuration space. In view of this, we have pursued the first approach and have employed the self-consistent reaction field method (SCRF) where the environment is modeled as a continuum of uniform dielectric constant. The models employed were Onsager's reaction field model SCRF(Dipole)[41] and the Polarized Continuum Model SCRF(PCM).[42,43] For the SCRF(Dipole) model the cavity containing the molecule is spherical, whereas for the SCRF(PCM) model it comprises a union of interlocking atomic spheres that attempt to correspond more closely to the accessible surface of the molecule. These calculations were carried out using both HF and B3LYP Hamiltonians with the 6-31G** basis set. The experimentally observed molecular structure (zwitterion) in the crystalline phase was used as the initial structure. The dielectric constant of the continuum was set to 78.39, corresponding to water. Water was chosen as the effective solvent since the partial charges to be derived could then also be used in simulations of D,L-norleucine in aqueous systems. Of a particular interest to us is the simulation of nucleation of D,L-norleucine from an aqueous system.

For each of the calculations in which the zwitterion form of D,L-norleucine was found to be stable, partial charges were derived by fitting to the electrostatic potential of the optimized molecule using the Merz−Singh−Kollman (MK)[44,45] and ChelpG[46] schemes. All of the above calculations were carried out using Gaussian 98.[47]

Screening of the Force Fields. The force fields AMBER[32] and CHARMM[33] and that of Williams' [31] were tested for their ability to reproduce the α- and β-crystalline phases of D,L-norleucine. As the Williams force field only specifies nonbonded interactions, the required bond, angle, and dihedral parameters were taken from AMBER. For the initial screening, the partial charges employed were those obtained from a Hartree−Fock in vacuo calculation with the 6-311G** basis set and the ChelpG scheme. The procedure involved constant-pressure geometry optimization of the crystal structures of the two phases using the code GULP.[48] The cutoff for the nonbonded interactions was 15.0 Å, and an Ewald summation[49] was used for the Coulomb interactions. The percentage deviation between the optimized and experimental structure was then compared. The initial screening yielded Williams' potential as the most promising. This potential was then employed to screen for the best partial charges in reproducing the crystal structures of the α- and the β-crystalline phases of D,L-norleucine.

Free Energy Minimization. Having ascertained the force field results using static minimization of the internal energy, a more severe test was performed in order to examine whether the interatomic potentials were of sufficient quality to correctly predict the β−α phase transition. Free energy minimization was

(38) Parr, R. G.; Yang, W. *Density Functional Theory of Atoms and Molecules*; Oxford University Press: Oxford, 1989.
(39) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(40) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785−789.
(41) Onsager, L. *J. Am. Chem. Soc.* **1938**, *58*, 1486.
(42) Miertus, S.; Scrocco, E.; Tomasi, J. *Chem. Phys.* **1981**, *55*, 117.
(43) Miertus, S.; Tomasi, J. *Chem. Phys.* **1982**, *65*, 239.
(44) Besler, B. H.; Merz, K. M., Jr.; Kollman, P. A. *J. Comput. Chem.* **1990**, *11*, 431.
(45) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1984**, *5*, 129.
(46) Breneman, C. M.; Wiberg, K. B. *J. Comput. Chem.* **1990**, *11*, 361.
(47) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, revision A.6; Gaussian, Inc.: Pittsburgh, PA, 1998.
(48) Gale, J. D. *J. Chem. Soc., Faraday Trans.* **1997**, *93* (4), 629.
(49) Ewald, P. *Ann. Phys.* **1921**, *64*, 253.

carried out for both phases within the framework of quasiharmonic lattice dynamics over the temperature range 50−350 K in steps of 25 K. The ZSISA approach[50] was utilized, in which the unit cell is minimized with respect to the free energy while the internal coordinates are optimized with respect to the internal energy. This approach is known to be more robust over a greater range of temperatures with regard to the influence of anharmonic effects than full minimization of all degrees of freedom with respect to the free energy.[51] Since the potential model was assessed on the basis of static data and is ultimately intended for application to molecular dynamics, the contribution of the zero point energy to the free energy was specifically excluded. Again, all calculations were performed using the program GULP, which contains analytic derivatives with respect to the free energy.[51]

When calculating the free energy via the vibrational partition function, it is necessary to approximate the integration of the phonon density of states over the Brillouin zone through the use of a Monkhorst−Pack grid of sampling points.[52] For the α-phase, a 2 × 6 × 3 mesh along the reciprocal lattice vectors was found to be adequate to converge the free energy to the required degree of precision. To further maximize the precision of the free energy difference between the phases, the calculations for the *β*-phase were performed on the full centered cell rather than the primitive one. Because this cell is equivalent to the unit cell of the α-phase doubled along the *a* direction, the mesh size chosen was 1 × 6 × 3, thus guaranteeing an equivalent degree of sampling for both phases and thereby a precise free energy difference.

Molecular Dynamic Simulations. The quality of the best identified force field was then further validated by carrying out extended molecular dynamics simulations of the α- and *β*-crystalline phases of D,L-norleucine at their respective temperatures of stability, 120 K for *β*- and 298 K for the α-phase. The simulations were carried out using the code DL-POLY[53] in the constant stress ensemble of Parrinello and Rahman (PR).[18,19] The PR ensemble enables the change of both symmetry and size of the simulation cell and attempts to minimize the effect of boundary conditions on solid-state phase transformations. The simulation box contained 120 molecules and was composed of 2 × 5 × 3 unit cells for the α-phase and 1 × 5 × 3 unit cells for the *β*-phase. The system temperature was controlled using the Nosé−Hoover thermostat.[54] The relaxation constants for the thermostat and barostat were 0.1 and 1.0 ps, respectively. The cutoff for the nonbonded interactions was 10.0 Å, and an Ewald summation precision of 10^{-6} was employed for the Coulombic interactions (Ewald convergence parameter = 0.35, and maximum *k*-vectors of 8, 8, and 8 in the three directions). Appropriate long-range corrections were added to the potential because of the truncation of the nonbonded interactions. All bonds were constrained using the Shake algorithm,[55] which allowed a time step of 0.002 ps. Each simulation was carried out for up to 600 ps.

The possibility of the selected force field being able to dynamically reproduce a solid-state phase transformation between the α- and *β*-phases was also explored. A number of simulations of the low-temperature *β* phase were carried out in the temperature range 320−340 K. Within this temperature range the *β*-phase is known (from experiment) to be unstable and undergoes a transformation to the α-phase.

Table 2. Properties of the Various Optimized Zwitterion Structures of the Norleucine Molecule. RMSD is the % Root Mean Squared Deviation between the Calculated and the Experimental Molecular Structure[14,15]

type of calculation	level of theory	basis set	RMSD/%	dipole moment
in vacuo	HF	6-31G*	0.363	9.96
in vacuo	HF	6-311G*	0.357	10.21
SCRF(Dipole)	HF	6-31G**	0.340	12.55
SCRF(PCM)	HF	6-31G**	0.227	13.87
SCRF(PCM)	B3LYP	6-31G**	0.385	12.29

Results and Discussion

We consider first the stability of the zwitterionic form of the norleucine molecule as a function of the type of calculation (vacuum or reaction field solvent), the type and level of theory (HF, MP2, or DFT-B3LYP), and the basis set employed. For the gas-phase calculations, HF yields the zwitterion with the relatively restricted basis sets 6-31G* and 6-311G*, but the non-charge-separated form of the molecule with the more flexible basis sets 6-311G** and 6-311++G**. The non-charge-separated form is also obtained at the MP2 level using 6-31G**, with calculations at the B3LYP level yielding the same result. These observations can primarily be ascribed to an influence akin to basis set superposition error, i.e., the incompleteness of the atomic basis sets results in an artificial barrier to proton transfer. The higher level and DFT calculations indicate that the zwitterionic form of D,L-norleucine does not represent a stationary point on the gas-phase potential energy surface, supporting the existence of the non-charge-separated form of norleucine in vacuo. A similar observation has been noted by others for glycine.[36,37]

With the solvent reaction-field calculations, the zwitterion is not stabilized in all cases as might be expected. While the HF calculations yield the zwitterion for both types of solvent reaction-field methods, SCRF(Dipole) and SCRF(PCM), the B3LYP calculations yield the zwitterion form only for the more refined SCRF(PCM) method where the molecule cavity in the continuum complements the topology of the molecule more closely. Details of the calculations that yielded a stable zwitterion, including the root-mean-squared deviation between the optimized and the experimental structure, are given in Table 2. The root-mean-square deviations are similar with the exception of the HF calculation using SCRF(PCM) method which gave a significantly lower deviation. The partial charges obtained using the ChelpG scheme from the in vacuo HF/6-311G** optimization were employed for the initial screening of the force fields. At this stage the solvent reaction-field calculations had not been completed. The minimum energy lattice parameters, along with the % deviations from the experimental values, for the three force fields investigated (AMBER, CHARMM, and

(50) Allan, N. L.; Barron, T. H. K.; Bruno, J. A. O. *J. Chem. Phys.* **1996**, *105*, 8300.
(51) Gale, J. D. *J. Phys. Chem. B* **1998**, *102*, 5423.
(52) Monkhorst, H. J.; Pack, J. D. *Phys. Rev. B* **1976**, *13*, 5188.
(53) DL_POLY is a general purpose serial/parallel molecular dynamics simulation package developed at Daresbury Laboratory by W. Smith and T. R. Forester under the auspices of the Engineering and Physical Sciences Research Council (EPSRC) for the EPSRC's Collaborative Computational Project for the Computer Simulation of Condensed Phases (CCP5). The package is the property of the Central Laboratory of the Research Councils.
(54) Hoover, W. G. *Phys. Rev. A* **1986**, *34*, 2499−2500.
(55) Ryckaert, J.; Ciccotti, G.; Berendsen, H. *J. Comput. Phys.* **1977**, *23*, 327.

Table 3. Experimental and Calculated Lattice Parameters for the α- and β-Phases of D,L-Norleucine Using Various Force Fields[a]

	α D,L-norleucine					β D,L-norlecuine				
force field	*a*/Å	*b*/Å	*c*/Å	β/deg	*U*/kJ mol^{-1}	*a*/Å	*b*/Å	*c*/Å	β/deg	*U*/kJ mol^{-1}
experiment	16.382	4.737	9.907	104.681	−114.5 ± 0.4	31.067	4.717	9.851	91.370	-
AMBER	17.113	4.964	10.378	107.001	−174.70	32.651	4.978	10.369	90.300	−176.55
	(4.5)	(4.8)	(4.8)	(2.2)		(5.1)	(5.5)	(5.2)	(−1.2)	
CHARMM	15.704	3.719	9.975	107.559	−155.62	27.103	3.503	10.188	91.475	−167.09
	(−4.1)	(−21.5)	(0.7)	(2.7)		(−12.8)	(−25.7)	(3.4)	(0.12)	
Williams	16.241	4.813	10.267	107.316	−198.99	30.865	4.846	10.247	90.379	−200.61
	(−0.8)	(1.6)	(3.6)	(2.5)		(−0.6)	(2.7)	(4.0)	(−1.1)	

[a] The experimental lattice parameters of the α-phase correspond to a structure determination at room temperature[14] while that of the β-phase correspond to a structure determination at 120 K.[15] The percentage deviation between the experimental and calculated values is given in the parentheses. The lattice energy *U* is defined as the difference in the energy of the crystalline phase per molecule and that of the isolated optimized zwitterion.

Table 4. Williams' Force Field Parameters for the Atom Types in D,L-Norleucine[a]

atom type	*A*/kJ mol^{-1}	*B*/kJ mol^{-1}	*C*/Å^{-1}
O	1260.73	241042.0	3.96
N	5629.82	405341.0	3.48
C1	1701.73	270363.0	3.60
C2	1435.09	103235.0	3.60
H1	0.00	764.9	3.56
H2	278.37	12680.0	3.56
H3	278.37	12680.0	3.56

[a] The force field takes the form $U_{ij} = B_{ij}\exp(-C_{ij}r_{ij}) - A_{ij}r_{ij}^{-6} + q_iq_j/4\pi\epsilon_0 r_{ij}$. Only the parameters for the homo-interactions are given. The parameters for the hetero-interactions are obtained using the combination rules $A_{ij} = (A_iA_j)^{1/2}$, $B_{ij} = (B_iB_j)^{1/2}$, and $C_{ij} = (C_i + C_j)/2$.

Williams) are given in Table 3. Overall, considering both phases of D,L-norleucine, the Williams force field comes out best, albeit exhibiting a 4% deviation in the *c*-axis parameter for the β-phase. Williams' force field for the repulsive part of the nonbonded interaction is of an exponential form, which on theoretical grounds represents a better description for repulsion. The repulsive part of this potential is also a little softer than the Lennard-Jones potential employed in AMBER and CHARMM and is therefore expected to be less restrictive to phase transformations in simulations of the solid state. Consequently, Williams' potential was selected for subsequent studies, from which the parameters employed are given in Table 4.

The variation in the partial charges for each of the ab initio calculations for which the zwitterion was found to be stable using ChelpG and MK schemes is shown in Figure 3. The most significant variation in the charges was for the atom sites around the ammonium and the carboxylate groups, and thus the charge variation is only shown for these sites. The in vacuo calculations yield charges that are similar to the SCRF(Dipole) calculation, but which differ significantly from the SCRF(PCM) calculations. In general terms, the SCRF(PCM) charges tend to be lower and the molecule less polarized. There is also a difference between the charges calculated using the ChelpG and MK schemes, with the ChelpG charges tending to be lower.

The results of the constant pressure lattice optimizations using Williams' potentials and the various sets of partial charges for each of the ab initio calculations are given in Table 5. In general terms, almost all the sets of partial charges, with the exception of SCRF(PCM) B3LYP/ChelpG, accurately reproduce the *a*- and the *b*-axes of both phases. The significant departure tends to be for the *c*-axis for both phases. Such a result would be expected, since the interactions in this direction are a subtle combination of van der Waals' forces and higher order electrostatic moments, while those in the *ab* plane are dominated by hydrogen bonds and stronger Coulomb terms. Table 5 also presents the respective lattice energies as well as the lattice energy differences between the two phases for each set of partial charges. The lattice energies were calculated as the difference between the potential energy of the crystalline phase per molecule and the potential energy of the single molecule optimized in a vacuum. It should be emphasized that the molecular state for the potential model here is the zwitterion (since charge-transfer cannot occur), whereas in reality the single molecule in vacuo is likely to be in the non-charge-separated form. Looking closely at the deviations between the calculated and the experimental lattice parameters, it is clear that the sets of charges based on the in vacuo calculations seem to fare the worst, giving the highest deviations. Of the remaining partial charge sets (which are all SCRF-derived), for two of the sets, SCRF(PCM) HF 6-31G**/MK and SCRF(PCM) B3LYP 6-31G**/MK, the lattice energy difference between the two phases is contrary to that expected with the high temperature α-phase having the lower energy. This leaves four distinct sets of charges for which the optimized lattice parameters are acceptable and the stability order (in terms of lattice energies) for the two phases at 0 K being correct.

From the short list, the two most promising sets of partial charges are those derived from the SCRF(PCM) HF 6-31G**/ChelpG calculation and the SCRF(PCM) B3LYP 6-31G**/ChelpG, henceforth referred to as SCRF-HF and SCRF-B3LYP, respectively. The former charge set (SCRF-HF) gives the lowest lattice deviations while the latter set (SCRF-B3LYP) gives the *lowest* difference in lattice energy (as well as the correct order i.e., $U_\beta < U_\alpha$). The significance of the latter set is that it is the one that is most likely to yield the $\beta \rightarrow \alpha$ phase transition at reasonable temperatures. For the transformation to be effected, the enthalpy difference ($U_\alpha - U_\beta$) between the two phases needs to be minimal to enable the relatively small $T\Delta S$ term to dominate. The $T\Delta S$ term is likely to be small in magnitude as the two crystalline phases have very similar structures.

For both identified sets of charges, SCRF-HF and SCRF-B3LYP, the α- and β-phases of D,L-norleucine were subjected to free energy minimization as a function of temperature in order to ascertain whether the α−β phase transition is predicted, and if so, at what temperature. The calculated free energy curves for the two phases (not shown) were essentially parallel for both sets of charges, with the free energies ranging from about −214 kJ mol^{-1} at 0 K to −250 kJ mol^{-1} at 300 K with a constant difference ($G_\alpha - G_\beta$) of about 1.7 kJ mol^{-1} for the charge set SCRF-HF and −178 kJ mol^{-1} at 0 K to −205 kJ mol^{-1} at 250 K with a constant difference ($G_\alpha - G_\beta$) of about 0.9 kJ mol^{-1}

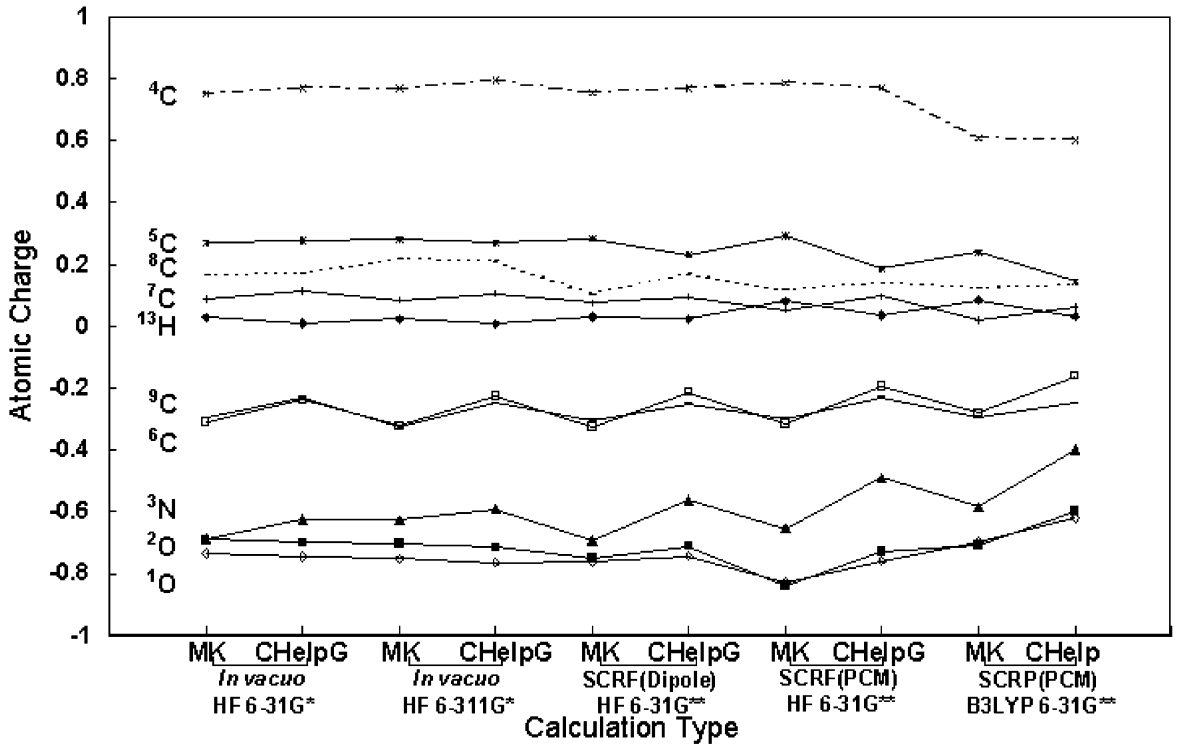

Figure 3. Variation in the ESP partial charges of the individual atoms obtained from the various ab initio calculations that yielded the zwitterionic form of norleucine. The abbreviations MK and ChelpG represent the Merz−Singh−Kollman[44,45] and ChelpG[46] electrostatic potential fitting schemes for deriving the partial charges. The prefixed superscript numbers before the atomic symbols refer to the labelling in Figure 1c.

Table 5. Experimental and Calculated Lattice Parameters of α and β D,L-Norleucine Using Williams' Nonbonded Force Field Parameters and Various Sets of ESP-Derived Partial Charges Obtained from Ab Initio Calculations[a]

	α D,L-norleucine					β D,L-norleucine					$U_\alpha - U_\beta$/
	a/Å	*b*/Å	*c*/Å	β/deg	U_α/kJ mol^{-1}	*a*/Å	*b*/Å	*c*/Å	β/deg	U_β/kJ mol^{-1}	kJ mol^{-1}
experiment	16.382	4.737	9.907	104.681	−114.5 ± 0.4	31.067	4.717	9.851	91.370	−	−
in vacuo HF 6-31G*											
MK	16.186	4.849	10.302	107.610	−192.99	30.697	4.867	10.323	90.372	−194.74	1.75
	(−1.2)	(2.4)	(3.9)	(2.8)		(−1.2)	(3.2)	(4.8)	(−1.01)		
ChelpG	16.234	4.832	10.281	107.437	−191.00	30.836	4.859	10.275	90.409	−192.55	1.55
	(−0.9)	(2.0)	(3.8)	(2.6)		(−0.7)	(3.0)	(4.3)	(−1.1)		
in vacuo HF 6-311G*											
MK	16.214	4.836	10.276	107.710	−198.07	30.372	4.855	10.294	90.146	−199.84	1.77
	(−1.02)	(2.1)	(3.7)	(2.9)		(−1.1)	(2.9)	(4.5)	(−1.3)		
ChelpG	16.241	4.813	10.267	107.316	−198.99	30.865	4.846	10.247	90.379	−200.61	1.62
	(−0.8)	(1.6)	(3.6)	(2.5)		(−0.6)	(2.7)	(4.0)	(−1.1)		
SCRF(Dipole) HF 6-31G**											
MK	16.157	4.804	10.270	107.672	−223.52	30.745	4.817	10.289	90.665	−224.47	0.95
	(−1.4)	(1.4)	(3.7)	(2.9)		(−1.0)	(2.1)	(4.4)	(−0.8)		
ChelpG	16.253	4.826	10.210	107.188	−197.90	30.923	4.854	10.204	90.502	−199.67	1.77
	(−0.8)	(1.9)	(3.1)	(2.4)		(−0.5)	(2.9)	(3.6)	(−0.9)		
SCRF(PCM) HF 6-31G**											
MK	16.149	4.733	10.162	107.104	−284.23	30.953	4.718	10.212	91.145	−284.02	−0.21
	(−1.4)	(−0.1)	(2.6)	(2.6)		(−0.4)	(0.0)	(3.1)	(−0.3)		
ChelpG	16.263	4.810	10.134	106.859	−210.13	31.006	4.834	10.123	90.686	−211.84	1.71
	(−0.7)	(1.5)	(2.3)	(2.1)		(−0.2)	(2.5)	(2.8)	(−0.7)		
SCRF(PCM) B3LYP 6-31G**											
MK	16.202	4.820	10.069	107.273	−233.76	31.029	4.812	10.141	91.111	−233.14	−0.62
	(−1.1)	(1.8)	(1.6)	(2.5)		(−0.10)	(2.0)	(2.9)	(−0.3)		
ChelpG	16.337	4.907	10.065	106.979	−164.11	31.149	4.916	10.108	90.670	−164.85	0.74
	(−0.3)	(3.6)	(1.6)	(2.2)		(0.3)	(4.2)	(2.6)	(−0.8)		

[a] The percentage deviation between the experimental and the calculated value is given in parentheses. MK and ChelpG are the Merz−Singh−Kollman[44,45] and the ChelpG[46] electrostatic potential fitting schemes for deriving the partial charges. The SCRF descriptor specifies the self-consistent reaction field method where the molecule is in a continuum solvent environment. The lattice energies U_α and U_β are defined as the difference in the energy of the crystalline phase per molecule and that of the isolated, optimized zwitterion. As the lattice energies are effectively 0 K calculations and the β-phase is the low-temperature phase, the energy difference ($U_\alpha - U_\beta$) should be positive.

for SCRF-B3LYP. For both sets of charges, the low-temperature β-phase remained stable throughout the temperature ranges examined. As specified above, the free energy difference between the two phases was lower (i.e., the curves were closer) for the charge set SCRF-B3LYP. Furthermore, for this set, the data are limited at the higher temperatures to a maximum of

Table 6. Final ESP (CHelpG Scheme) Partial Charges for the Norleucine Molecule Obtained from a B3LYP/6-31G** Calculation Using the Self-Consistent Reaction Field SCRF(PCM) Model

atom number	atom type	charge/e	atom number	atom type	charge/e
1	O	−0.6194	12	H1	0.2940
2	O	−0.5984	13	H2	0.0316
3	N	−0.3995	14	H3	0.0621
4	C1	0.6014	15	H3	0.0488
5	C2	0.1438	16	H3	−0.0496
6	C2	−0.1625	17	H3	0.0009
7	C2	0.0607	18	H3	−0.0197
8	C2	0.1312	19	H3	−0.0013
9	C2	−0.2497	20	H3	0.0568
10	H1	0.3059	21	H3	0.0548
11	H1	0.2411	22	H3	0.0670

Table 7. Calculated Lattice Parameters for the α- and β-Phases of D,L-Norleucine from Molecular Dynamics (MD) Simulations Using the Parrinello−Rahman Constant Stress Ensemble[a]

lattice parameter	α D,L-norleucine			β D,L-norleucine		
	experiment	calculated	deviation/%	experiment	calculated	deviation/%
a/Å	32.76	33.12	1.1	31.06	31.33	0.9
b/Å	23.68	24.77	4.6	23.58	24.30	3.1
c/Å	29.21	31.08	6.4	29.55	30.39	2.8
α/deg	90.00	90.04	0.0	90.00	90.01	0.0
β/deg	104.68	100.88	−3.6	91.37	91.93	0.6
γ/deg	90.00	90.02	0.0	90.00	89.98	0.0

[a] Also given are the experimental values and the % deviation between the experimental and the calculated values. The calculated parameters were obtained by averaging over the last half of the simulation trajectories (300−600 ps). The MD simulations were carried out at 298 K/0.001 kbar for the α-phase and at 120 K/0.001 kbar for the β-phase, which correspond to their respective conditions of stability.

250 K, above which the β-phase begins to show instability as demonstrated by the appearance of imaginary phonon frequencies. This tendency is also exhibited by the α-phase but at a correspondingly higher temperature of approximately 50 K above that of the β-phase. In contrast, for the charge set SCRF-HF, the free energies for both phases could be obtained up to 300 K without any instability being exhibited. Above 300 K this charge set also revealed instabilities in the α- and β-phases, again with α-phase instability setting in at a temperature of about 50 K higher than that of the β-phase. In summary, both charge sets reveal a similar behavior, but with the difference that the SCRF-B3LYP charge set (which gave the lowest lattice energy difference between the two phases) suggests the desirable outcome of the α−β phase transition being at a lower temperature, more in line with the experimental observation of the α-phase being the stable phase at ambient conditions.

Clearly, the phase instabilities observed in the free energy calculations at the higher temperatures are an indication that in this region anharmonic effects become significant, and hence the quasiharmonic free energies are expected to be unreliable. Since it is not presently possible to extend these static calculations beyond the quasiharmonic approximation, we cannot draw any absolute conclusions about the relative stability of the two polymorphs near the expected experimental transition temperature, except via molecular dynamics (where anharmonicity is allowed for). However, the free energy calculations do indicate that the β-phase becomes unstable at temperatures approaching ambient ones, while the α-phase resists this tendency for approximately a further 50 K. It is possible that this anharmonic behavior in the case of the α-phase may be associated with the transition to a further higher temperature

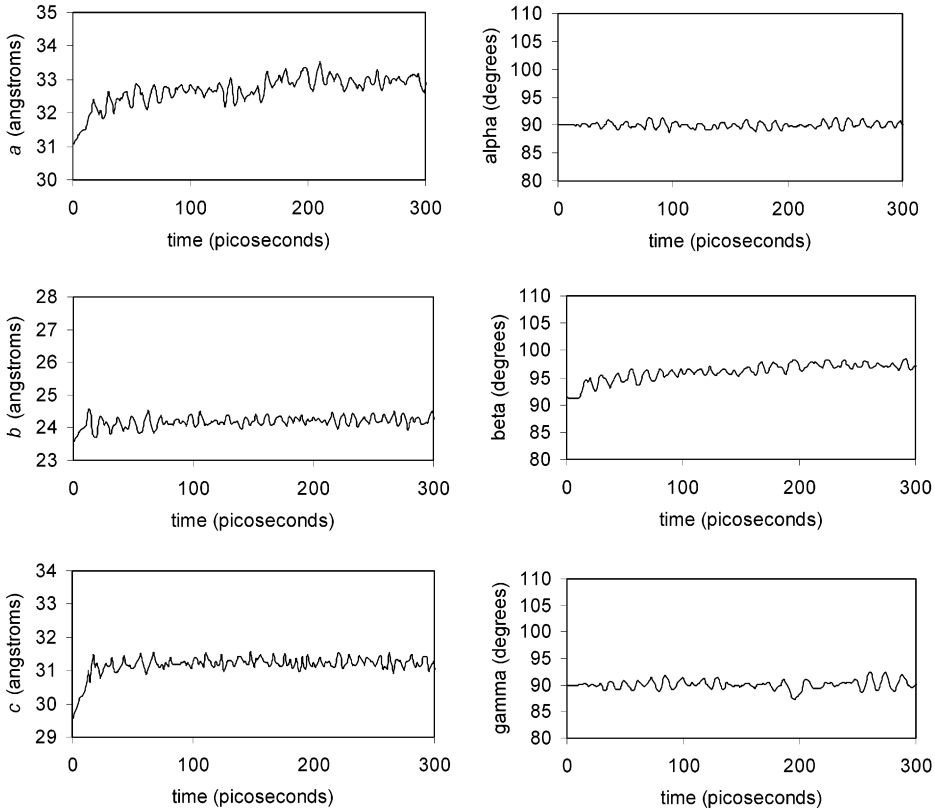

Figure 4. Simulation cell constants as a function of time for a molecular dynamics simulation of the β-phase of D,L-norleucine in the Parrinello−Rahman constant stress ensemble at 340 K/0.001 kbar. The β-phase structure transforms to α-phase over the interval 150−170 ps.

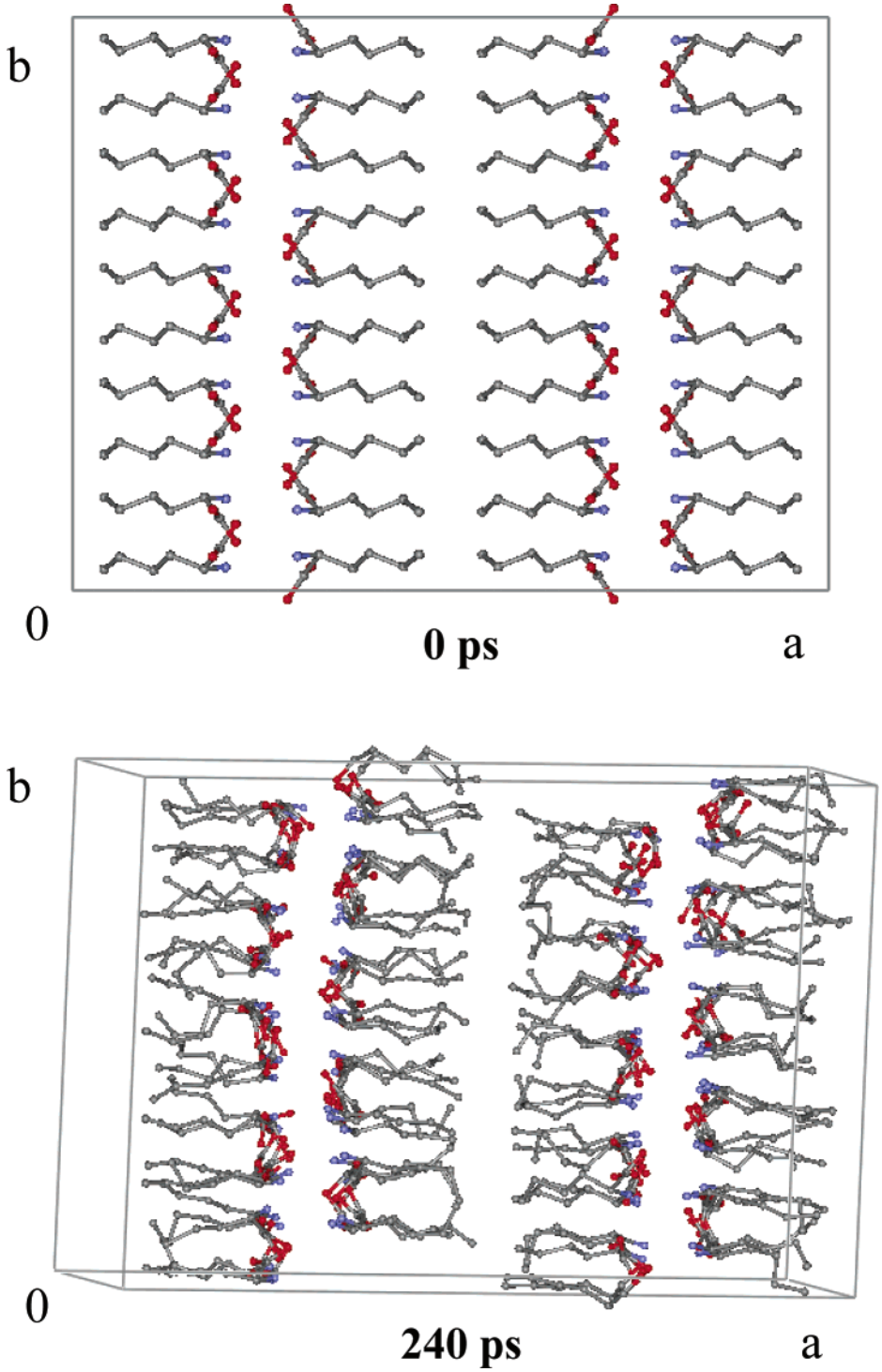

Figure 5. Snapshots of the initial β-phase and of the newly formed α-phase of D,L-norleucine (after 240 ps) from a molecular dynamics simulation of the β-phase of D,L-norleucine in the Parrinello−Rahman constant stress ensemble at 340 K/0.001 kbar. The view is looking down the *c*-axis and hydrogen atoms have been excluded to enhance structural clarity.

phase, perhaps indicating the formation of the γ-phase (about which we know very little) that is observed experimentally above about 390 K.

Examination of the structure variation with temperature for both charge sets reveals that all the cell lengths increase as the temperature rises, as would be expected for most materials. The most significant positive thermal expansion occurs for the *a*-cell parameter since this controls the interlayer spacing that is determined by weaker forces. It is this expansion, with the associated increased freedom of movement for the hydrocarbon branches, which ultimately leads to the breakdown of the quasiharmonic approximation. Again this is consistent with the view that the instability is related to the phase transition, given that it involves displacement of adjacent layers with respect to each other. Considering now the monoclinic cell angle, the value for the α-phase decreases to a value of 105.1/104.9° at 300 K for the two sets of charges SCRF-HF and SCRF-B3LYP, while that of the β-phase steadily increases with temperature to a value

of 92.3/92.7° at the temperature of 250 K. Hence the monoclinic distortions of the two phases are tending toward each other, but are unlikely (given the relative thermal expansion coefficients) to converge to a common value before the phase transition occurs. However, the highly nonlinear nature makes it impossible to rule it out such convergence. On the basis of the free energy data, it appears that the best set of charges is that derived from the solvent reaction field SCRF(PCM) B3LYP 6-31G**/ChelpG calculation and which yields the lowest lattice energy difference between the phases. These charges are given in Table 6.

The identified force field was further validated by carrying out extended molecular dynamics simulations of the α- and β-phases at their respective temperatures of stability using the constant stress ensemble of Parrinello and Rahman. Both phases were found to be stable over the 600 ps simulations. The lattice parameters averaged over the last half of the trajectory (given in Table 7) show a maximum discrepancy of 6.4%. In general, a small positive deviation is expected as the force field and charges were screened in the first instance using potential energy minimization, that is, at 0 K. Examination of the mean atomic positions revealed that the molecular structures in the two phases remained essentially identical to the experimental structures. Clearly the force field reproduces the structural aspects of the crystal phases reasonably accurately.

The lattice energies of the α- and β-phases, defined as the difference in the energy of the optimized crystalline phase per molecule and that of a single optimized molecule (zwitterion) in a vacuum, using the final definitive force field are $U_\alpha = -164.11$ kJ mol^{-1} and $U_\beta = -164.85$ kJ mol^{-1}. The experimental lattice energy for the α-phase is -114.5 ± 0.4 kJ mol^{-1}.[56] The experimental and calculated values differ markedly, but this should not be surprising as the calculated values assume that the isolated molecule exists as a zwitterion, which is not the case. The *difference* between the lattice energies of the two phases yields an energy change for the β- to α-phase transition of $\Delta U_{\beta\rightarrow\alpha} = 0.74$ kJ mol^{-1}. This value characterizes the shifting of the bilayers along the van der Waals surface at the terminus of the aliphatic chains. The experimental enthalpy for the β- to α-phase transition is not known, but a value for the α- to γ-phase transition has been reported as 4.92 kJ mol^{-1}.[57] The calculated phase transition energy for the β- to α-transition is therefore of a reasonable physical magnitude, considering the calculated energies are from a 0 K calculation and we neglect the $p\Delta V$ term. The implication is that the α- to γ-phase transition, as for the β to α one, is probably also likely to involve shifting of the bilayers at the surface defined by the terminus of the aliphatic chains. This is valuable information since the crystal structure of γ-phase is not known and continues to be elusive.

Probably the most rigorous test for any force field is whether it can describe the phase equilibrium behavior of the material. Preliminary molecular dynamics simulations using the identified force field of the β-phase of D,L-norleucine over the temperature range 320−340 K (where it should become unstable in favor of the α-phase) show the expected phase transformation to the α-phase. The changes in the lattice parameters as a function of the simulation time during the transformation at 340 K are given in Figure 4. The β-phase remains stable for about the first 150 ps, after which the structure transforms to the α-phase over an interval of about 20−40 ps. Snapshots of the initial β-phase and the emergent α-phase after 240ps are given in Figure 5. The α-phase is essentially identical in structure to the experimental structure. The simulation cell is too small to show any nucleation event and the transformation, a shift of half a cell length along the *b*-axis of one of the bilayers with respect to the other, is observed to occur relatively sharply and in a concerted mode. The transformation begins by de-stabilization of the van der Waals surface between the bilayers, with some of the molecules shifting along the *b*-axis and losing their characteristic β-phase bilayer−bilayer registration. The packing of these molecules at the interface oscillates between that of the α- and β-phase structures until eventually one of the bilayers shifts by a half a cell length (relative to the other) to yield the α-phase. These oscillations may be artifacts of the small system size, resulting from the large fluctuations (relative to the cell dimensions) of the simulation cell associated with the barostat. Clearly larger scale simulations are required to make definitive inferences in respect of the molecular mechanism. Furthermore, there are a number of other important technical variables that include the thermostat and barostat relaxation constants, bond constraints as opposed to harmonic bonds, and inclusion of vacancies in the system, whose effects still need to be investigated.

In conclusion, a force field has been defined for D,L-norleucine that reproduces the crystal structures of the two known phases in static calculations and their stability in extended molecular dynamics simulations. More remarkably, the force field enables simulation of the β- to α-phase transformation, with the simulated α-phase being essentially identical to the experimental structure. The defined force field should enable the elucidation of the molecular mechanisms underpinning the phase transformations in crystals of DL-norleucine and hence enhance our general understanding of the extraordinary class of phase transitions that exhibit topotaxy.

JA0356131

(56) Chickos, J. S.; Acree, W. E. JR. *J. Phys. Chem. Ref. Data* **2002**, *31* (2), 537−698.
(57) Matsumoto, M.; Kunihisa, K. S. *Chem. Lett.* **1984**, *8*, 1279−1282.