Models for transition metal oxides and for protein design

by

Fei Zhou

B.S. Physics, University of Science and Technology of China (2001)

Submitted to the Department of Physics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Physics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

. September 2006

© Fei Zhou, MMVI. All rights reserved.

Author . . . .
Department of Physics
August 8, 2006

Certified by . . . .
Gerbrand Ceder
R. P. Simmons Professor of Materials Science
Thesis Supervisor

Certified by . . . .
John D. Joannopoulos
Francis Wright Davis Professor of Physics
Thesis Supervisor

Accepted by . . . .
Thomas J. Greytak
Professor of Physics
Associate Department Head for Education
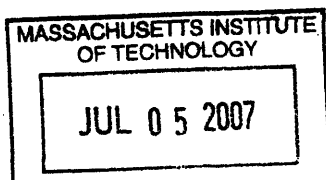
Models for transition metal oxides and for protein design

by

Fei Zhou

Submitted to the Department of Physics
on August 8, 2006, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Physics

Abstract

A large number of properties of solid state materials can now be predicted with standard first-principles methods such as the Local Density (LDA) or Generalized Gradient Approximation (GGA). However, known problems exist when using these methods for predicting the electronic structure and total energy of transition metal oxides. We demonstrate that the LDA+U method, in which localized d electrons are treated with a Hubbard-like model, gives considerably better predictions in band gap, magnetization, lattice constants, redox potential, mixing energy and other properties of lithium transition-metal oxides. The systematic improvement of both electronic structure and total energy properties suggests that the more accurate treatment of electron correlation on the transition metal sites will greatly enhance the predictive character of *ab-initio* methods for these materials. In a even more dramatic example, the more accurate electronic and total energy descriptions with LDA+U have enabled us to study the finite temperature phase diagram of $Li_x FePO_4$. It is found that an unexpected solid solution phase are mainly stabilized by the entropy contribution of localized d electrons. This brings the less visited problem of the effects of electronic correlation on thermodynamics into our attention.

The second part of this thesis is focused on application of methods used in solid state materials to proteins. Protein systems are well known for having extremely large conformation space, which makes their energy evaluation expensive even with empirical energy models. We are implementing a coarse graining scheme, called the cluster expansion, to help alleviate this problem. Traditionally, the cluster expansion is a powerful tool to model the energetics of solid-state materials with configurational disorder. We have developed a generalized cluster expansion framework suitable for protein studies, in which properties such as the energy can be unambiguously expanded in the amino acid sequence space. The cluster expansion coarse-grains over non-sequence degrees of freedom (e.g. sidechain configurations) and thereby simplifies the problem of designing proteins, or predicting the compatibility of a sequence with a given structure, by many orders of magnitude. With the help of this method, we have been able to perform a full-sequence design for a small peptide within a relatively short period of time.

Thesis Supervisor: Gerbrand Ceder
Title: R. P. Simmons Professor of Materials Science

Thesis Supervisor: John D. Joannopoulos
Title: Francis Wright Davis Professor of Physics

Acknowledgments

Many thanks to ...

Contents

List of Figures

12

14

15

List of Tables

Chapter 1

Introduction

1.1 Motivations

Theoretical and computational studies of real-world materials have come a long way in the accuracy that can be achieved, in the complexity of the targeted materials, and in the scale of the problem. With the advancement in modern condensed matter theory and the advent of powerful, inexpensive computer resources, these research efforts have attained such credibility in certain cases that they sometimes achieve the level of *"in silico"* or numeric experiments. Nowadays the major challenges for such research can be put into two categories. First, there are cases when the physics or mechanism controlling the relevant property in a material is so intriguing that it is still at the frontier of basic physics research. A convenient example is high-temperature superconductivity — after decades of intensive research, the nature of high T_C superconductivity is still illusive. Secondly, some materials properties are in principle adequately understood, but the scale of computation needed to resolve their properties is well beyond what current computer technology can offer. The protein folding problem is probably the best known example in this category.

This thesis deals with two distinct problems. The first one, on the accurate prediction of both ground-state and finite-temperature properties of certain transition metal oxides, is somewhat closer to the first category. A brief introduction is given later in this chapter, and more detailed results and discussion can be found in chapter

2 and 3. The second part of the thesis, chapter 4 and 5, is focused on reducing the complexity of the energetics in a protein system with fixed backbone structure — a problem similar to the second category.

1.2   Correlation in transition metal oxides

First-principles calculations employing density functional theory (DFT) have proven to be a powerful method in understanding the electronic, structural and thermodynamic properties of a large class of materials [135, 79, 89]. The density functional is not known exactly, and is usually modeled within the Local Density Approximation (LDA) or Generalized Gradient Approximation (GGA). For many systems LDA or GGA gives remarkably good agreement with experiments, which has made these techniques valuable tools to predict the behavior of materials.

A fundamental problem with LDA is that its exchange-correlation energy functional $E_{xc}^{\mathrm{LDA}}[n]$ is defined as a local function of electron density $n(r)$. As a result its variational derivative $\delta E_{xc}^{\mathrm{LDA}}[n]/\delta n(r)$, the exchange-correlation potential, is a continuous function of density $n(r)$. More sophisticated approximations like GGA [20, 141, 137] incorporate the gradient corrections $\nabla n(r)$ into the exchange-correlation functional $E_{xc}^{\mathrm{GGA}}[n, \nabla n]$, but unfortunately have a continuous potential as well. This is in drastic contradiction to the exact density functional, whose derivative (potential) displays a discontinuity at integral number of electrons [140]. As shown in Fig. 1-1 (taken from [39]), the correct total energy E versus number of electrons N curve (red lines labeled "exact") should consist of series of straight-line segments with derivative discontinuities at integral N. However, the LDA total energy (black curve labeled "LDA") has smooth derivative. Any attempts to improve LDA as an approximation to exact density functional theory should be able to (at least approximately) reproduce this discontinuity.

The lack of discontinuity in LDA (or GGA) is related to the fact that the energy functional is defined in such a way as to treat electron interactions based on a homogeneous electron gas. When strong localization of the electrons occurs, the errors
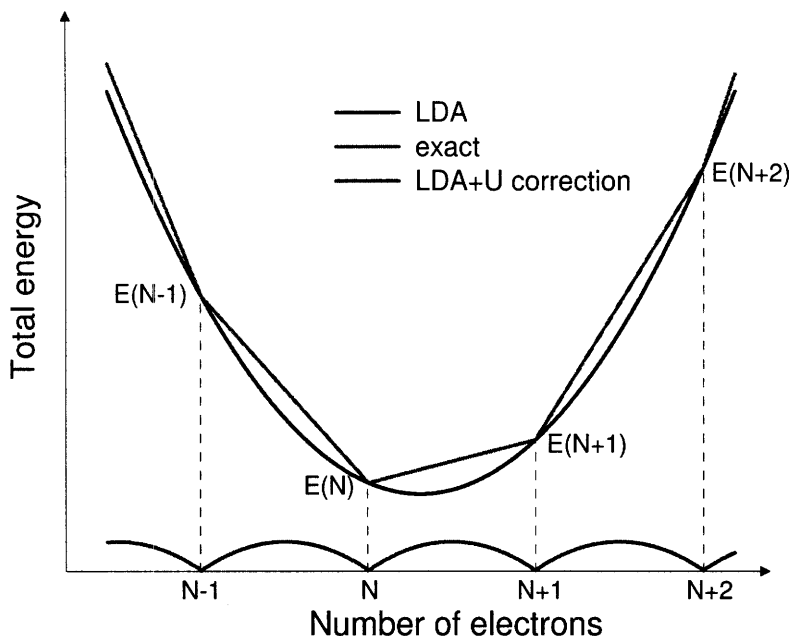
Figure 1-1: Schematic curve of total energy versus number of electrons.

in LDA electron interactions are particularly pronounced. A well known example is Mott-Hubbard insulators, e.g. transition metal oxides, in which LDA largely misses the Coulomb repulsion between electrons localized at the same site [144, 86]. The resulting failure to predict many transition metal oxides as insulators, as well as the underestimation of magnetization on TM ions, has been well documented (for reviews see [10]).

Among the many attempts to improve the LDA scheme, the most commonly adopted are the self-interaction-corrected density functional theory (SIC-DFT) [142, 176], the LDA+U method [13, 11], the LDA+DMFT [163] and the GW approximation [77]. The SIC and the LDA+U methods in principle do not add to the computational costs of normal LDA or GGA calculations, but treat electron correlations in a static fashion, which is believed to be the main source of inaccuracy in these methods. The LDA+DMFT and the GW methods generally have better performance in weakly or moderately correlated systems, though at a substantial computational cost. In this thesis we focus on the LDA+U method, because of the materials we are studying (mostly insulating transition metal oxides with strong electron correlations), because of the accuracy goal of about 10 meV relevant to these materials' practical applications

(higher accuracy is of course welcome, but not essential), and because of its advantage in computational efficiency. It will be shown that the LDA+U methods provides a reliable tool to capture the relevant physics–correlation of localized d-electron on transition metal ions–and make accurate predictions.

1.3 The LDA+U method

The DFT+U method, developed in the 1990's [13, 12, 106], is now a well-established model to deal with strong electron correlations. The method combines the high efficiency of LDA/GGA, and an explicit treatment of correlation with a Hubbard-like model for a subset of states in the system. As a leading correction to the LDA, the LDA+U method has been very successful in the transition metal oxides (see [10] and references therein). For clarity in the future discussions, we will refer to the name GGA+U when the GGA functional is used instead of LDA, or the name DFT+U when we are talking about the method itself without explicit reference to LDA or GGA.

The essence of the method can be summarized by the expression for the total energy

$$E_{\text{LDA}+U}[\rho, \hat{n}] = E_{\text{LDA}}[\rho] + E_{\text{Hub}}[\hat{n}] - E_{\text{dc}}[\hat{n}] \equiv E_{\text{LDA}}[\rho] + E_U[\hat{n}] \qquad (1.1)$$

where ρ denotes the charge density and $\hat{n}$ is the iron on-site $3d$ occupation matrix. The Hatree-Fock like interaction E_{Hub} from the Hubbard model replaces the double counting (dc) term E_{dc} representing the LDA on-site interaction. The U correction term $E_U \equiv E_{\text{Hub}} - E_{\text{dc}}$ is defined by Eq. 1.1.

However E_{dc} is not uniquely defined, and here we consider three common approaches (see Ref. [143] and references therein), expressed in the rotationally invariant form by Liechtenstein *et al* [106]. The "around mean field" dc scheme [12, 143]

is (simplified with the spherical average, as in dc3 below) given by:

$$E_{\text{dc}}^{\text{AMF}}(\hat{n}) = \frac{U-J}{2}\text{Tr}\hat{n} \cdot \hat{n} = \frac{U'}{2}(\text{Tr}\hat{n})^2, \tag{1.2}$$

$$E_U^{\text{AMF}}(\hat{n}) = -\frac{U-J}{2}\text{Tr}\,(\hat{n} \cdot \hat{n}) = -\frac{U'}{2}\text{Tr}\,(\hat{n} \cdot \hat{n}) \tag{1.3}$$

$$,V_U^{\text{AMF}}(\hat{n}) = -U'(\hat{n} - \langle n_\sigma \rangle I), \tag{1.4}$$

where we have defined $U' = U - J$, and I is the identify matrix. $\langle n_\sigma \rangle$ is the average occupancy of the σ spin orbitals. The exchange parameter J is of the order of 1 eV for most late transition metal oxides [13]. The other schemes are the dc functional defined in [106] (dc2) and its spherically averaged version [61] (dc3), i.e. the so-called "fully-localized limit" (FLL). The latter reads

$$E_{\text{dc}}^{\text{FLL}}(\hat{n}) = \frac{U-J}{2}\text{Tr}\hat{n}(\text{Tr}\hat{n} - 1) = \frac{U'}{2}\text{Tr}\hat{n}(\text{Tr}\hat{n} - 1), \tag{1.5}$$

$$E_U^{\text{FLL}}(\hat{n}) = \frac{U-J}{2}\text{Tr}\,(\hat{n}(1 - \hat{n})) = \frac{U'}{2}\text{Tr}\,(\hat{n}(1 - \hat{n})) \tag{1.6}$$

$$,V_U^{\text{FLL}}(\hat{n}) = -U'(\hat{n} - \frac{1}{2}I). \tag{1.7}$$

Because dc2 and dc3 give practically the same results for most systems, the simpler dc3 is usually favored.

The AMF (dc1) and the FLL (dc3) schemes are meant to work for different materials. In weakly or moderately correlated systems, such as transition metal or its alloy, there is evidence that the AMF version works better [143, 167], while for strongly correlated materials where electrons are more localized, the FLL version is preferred. In this thesis we use the FLL LDA+U because of the materials we are dealing with.

1.4   Calculation of the U parameter from linear-response theory

A main disadvantage of the LDA+U method is that the $U' = U - J$ (hereby referred to as just U) parameter is external, rendering the results somewhat arbitrary. Here we

determine the U parameter using the method presented in Ref. [39] which we briefly outline below. This method is based on calculating the response in the occupation of TM states to a small perturbation of their local potential. Such determination of the U parameter makes our approach a real *ab-initio* technique.

As evident from Eq. 1.6, the effective U can be evaluated through the second derivative of the total energy of the system with respect to the atomic occupations of the "Hubbard" ions. The meaning of the second derivative of the total energy with respect to the electronic occupation of the outermost orbitals of an (atomic) system has been clearly pointed out in Ref. [140, 139] and used in [39] where a connection with the physical meaning of the Hubbard U entering the "+U" correction has been established. As shown in Fig 1-1 this quadratic behavior of the total energy is basically a spurious effect introduced by LDA because of incomplete cancellation of the self-interaction of the electrons (or, equivalently, poor representation of their correlation). It is shown in [39] that, in this context, the correction to LDA introduced by the LDA+U is exactly the one needed to recover, in the atomic limit, the physical linear behavior of the total energy with respect to the occupation of the atomic orbital that gets filled. This linearity and the discontinuity of dE/dn are essential to correctly account for important physical quantities such as the difference between ionization potential and electron affinity, which basically corresponds to U and to the band gap of the system (Janak's theorem). In crystals the situation is complicated by partial occupations of atomic orbitals, but the removal of the curvature of the total energy is still fundamental in correctly describing properties involving electron transfer processes such as redox potentials.

In the current approach, the second derivative of the LDA total energy, which defines U, is evaluated indirectly using a linear response method inspired by Ref [145]. We start from an LDA/GGA ($U = 0$) calculation for a transition metal compounds in a supercell with metal sites $i, j, \cdots$ as the reference point. Then a small perturbation

$$dV = \alpha P_d^i, \quad P_d^i = \sum_{m=-2}^{2} |m^i\rangle\langle m^i|$$

in the local d-orbital potential is exerted on metal site i, where P_d^i represents the projector on the d states manifold of ion i, and α is the amplitude of the potential shift applied to the d levels. This induces a change in the occupation number of ion i as well as other ions. Thus we can calculate directly the response matrices,

$$\chi_{ji} = \frac{dn_d^j}{d\alpha_i}, \quad \chi_{0ji} = \frac{dn_{0d}^j}{d\alpha_i}, \tag{1.8}$$

which measure the variation of the d-manifold charge density n_d^j, on ion j, produced by a potential shift at ion i. The subscript "0" denotes the bare response, calculated without self-consistency (the Kohn-Sham potential apart from dV is frozen at the value obtained in LDA/GGA before the perturbation), and corresponds to the response from an independant electron system, while χ_{ji} is the screened response (charge density and potential relaxed to reach self-consistency). The effective interaction parameter U is then obtained as

$$U = (\chi_0^{-1} - \chi^{-1})_{ii}. \tag{1.9}$$

This is a well-known result in linear response theory, in which the effective electron-electron interaction kernel is given as a difference among the interacting density response and the non-interacting one [2, 200]. Since DFT is used, a finite contribution from the exchange-correlation potential is also included in the effective U. As we use the integrated quantity n_d^i to probe the responses, the calculated effective interaction is averaged over the ion in the same spirit as DFT+U. The matrix in Eq. 1.9, whose diagonal term defines the on-site Hubbard U, also contains non-diagonal terms corresponding to inter-site effective interactions in LDA/GGA. These are not used in the DFT+U model.

This method to compute U contains full account of the screening to the external perturbation operated by the electron-electron interactions. In order to obtain physically correct results, the linear response calculations for χ and χ_0 are performed in larger and larger supercells until full convergence of U is achieved. This procedure is required to guarantee that the applied perturbation is isolated and the system con-

tains all the necessary degrees of freedom for screening. To reduce the computational cost of the linear response calculations a mathematical procedure has also been used in Ref. [39] which enforces the charge neutrality of the applied perturbation in the response matrices. This constraint enhances the local character of the response of the system and speeds up the convergence with cell size.

A very important advantage of this linear response approach is that the external perturbation is applied to the same system we want to treat with LDA+U and one obtains full consistency of its physical behavior with the interaction we compute. Furthermore, the linear response approach allows us to choose different definitions of atomic orbital occupations, as long as we adopt the same definition when calculating U. Thus the above procedure is a fully consistent ab-initio scheme, making possible LDA+U calculations without the need of any a-priori assumption or semi-empirical evaluation of the interaction parameter.

Chapter 2

Accurate *ab-initio* calculation of the ground-state properties of transition metal compounds

2.1 Introduction

In the previous chapter we gave an introduction to the general problem of accurately predicting materials properties in the technologically important transition metal oxides. More specifically, we mainly fucus on the lithium-intercalated transition metal oxides that are used in the cathode of rechargeable lithium-ion batteries.

First principles computations have shown to be relevant for predicting many of the properties of lithium metal oxides [18, 30, 32, 21, 119, 190, 205, 55, 103, 147, 99, 177, 84, 153, 88, 90, 122, 59]. The focus has been on bulk properties such as lithium insertion voltage [18, 17, 31], phase stability [189, 151, 33] and lithium diffusivity [192, 191, 193, 194, 124], in which electrostatic interactions dominate. These studies are based on the LDA or GGA method, and are generally in good agreement with experiment. This shows that even though LDA and GGA do not adequately address the electron correlation effects in transition metal oxides, they give a reasonable description of the leading energy effects, i.e. electrostatics, especially when they deal

with lithium ions. However, in the DFT scheme the electronic distribution and the total energy are closely tied. Without proper treatment of electron correlations, certain errors in both total energy and electronic properties prediction naturally occur, the most well known of which are the underestimation of redox potential and band gap in these materials.

We demonstrate that the DFT+U method, which takes into account on-site d-electron correlations explicitly, is a reliable tool to model a good number of transition metal oxides. In this chapter we focus on zero-temperature properties of these materials, such as band gap, magnetic moment, lithium-insertion voltage and formation energies. For insulating, strongly correlated Li_xMO_y the DFT+U predictions are very accurate. For metallic materials, the method at least represents qualitative corrections over the LDA or GGA where explicit on-site correlation is absent. We will deal with finite temperature thermodynamical properties in the next chapter.

2.2 Materials

Transition metal oxides (TMO) are arguably the most important class of materials in modern solid state physics. They exhibit a rich collection of interesting and intriguing properties, which can be tailored for a wide variety of applications including low-loss power delivery, quantum computing using cooper pairs, ultra high-density magnetic data storage and spintronic [150, 109, 65].

Considerable interest exists in certain lithium-intercalated TMOs (denoted by Li_xMO_y) due to their application as rechargeable battery cathodes [125, 164]. The lithium content with respect to the metal may vary in a large range without substantial changes to the crystal structure of the material, making possible its application as an electrode in a rechargeable battery. Fig. 2-1 (taken from [18]) shows the basic operation mechanism of such a lithium-ion battery. The process can be represented by the following reversible redox reaction:

$$xLi + MO_y \leftrightarrow Li_xMO_y.$$

28
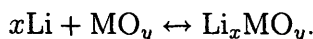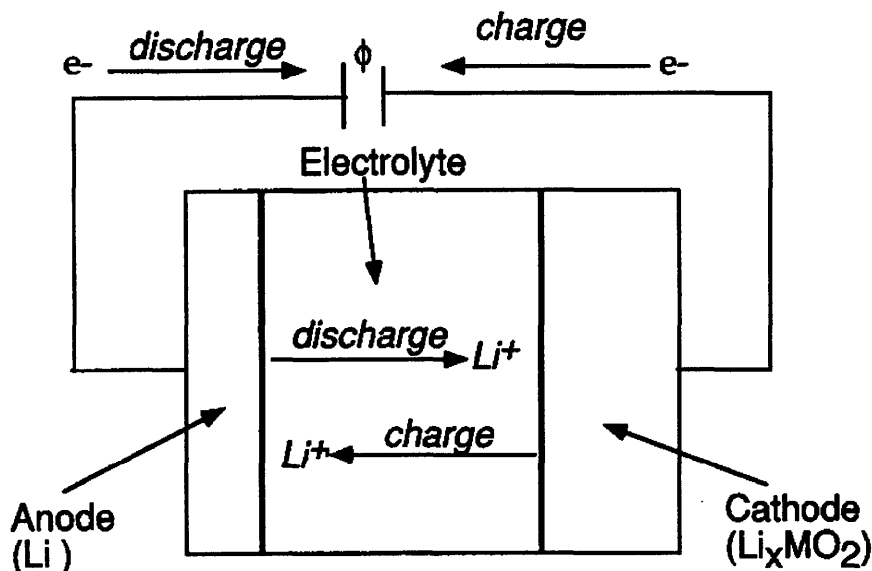
Figure 2-1: Schematic picture of a rechargeable electrochemical Li cell. When the battery discharges Li is intercalated into the cathode compound. Upon charging the Li ions are removed from the cathode.

It is believed that Li is fully ionized in most lithium-metal oxides and donates its electron to the host oxide, which is redistributed mainly to TM ions as well as to the oxygen atoms, due to TM-O rehybridization [18, 131, 111]. The electrons that enter the transition metal d-orbitals are under the influence of strong on-site correlations, dramatically different from their original s-character in lithium metal before insertion. Therefore we expect electron correlation effects to play an important role in determining the ground-state properties of Li_xMO_y.

As a representative set of Li-insertion compounds, we have selected several materials representing different environments for Li and TM ions, which are well characterized experimentally.

2.2.1 Layered lithium metal oxide Li_xMO_2

The layered $LiMO_2$ are traditional cathode materials that have been thoroughly studied experimentally (see ref. [181] and references therein) and theoretically [18, 205, 191, 152]. The layered structure can be envisioned as two interpenetrating fcc lattices, one consisting of oxygen, and the other consisting of alternating (111) planes of Li
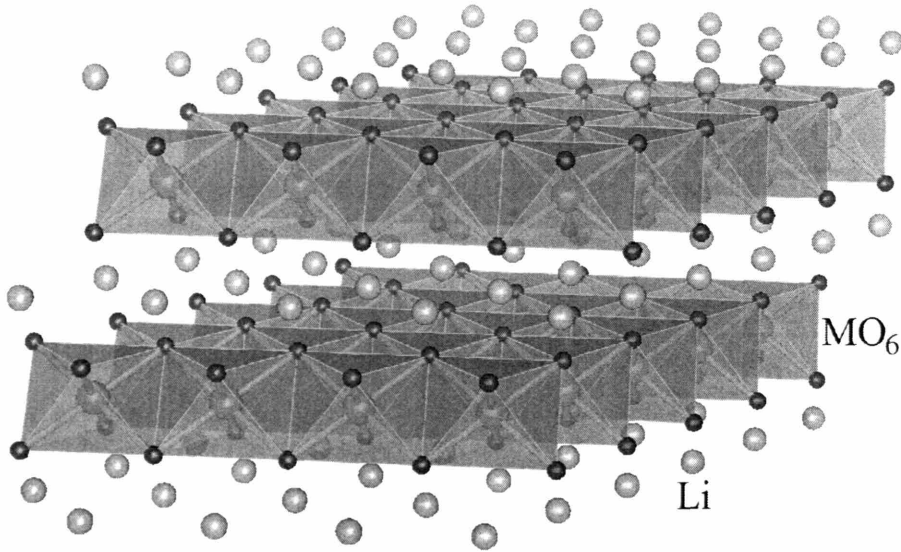
Figure 2-2: The layered structure with MO_6 octahedra and lithium atoms.

and TM ions (see Fig. 2-2). In the $R\bar{3}m$ space group the Li and the metal ions remain fixed in the ideal rock salt positions, but the whole (111) oxygen planes may relax in the [111] direction giving rise to different slab spaces that affect lithium mobility [87].

2.2.2   Spinel lithium metal oxide $Li_xM_2O_4$

The spinel-like structure $Li_xM_2O_4$ is so named because at $x = 1$ it has the same structure as the spinel mineral $MgAl_2O_4$ (see Fig. 2-3). We shall refer to it as spinel even when $x = 2$. It can be envisioned as an fcc oxygen sublattice, with TM in half of the octahedral oxygen interstices, and lithium either in part of the tetrahedral sites at $x = 1$ or in the octahedral sites not occupied by the TM ions at $x = 2$ [181].

The layered and spinel lithium metal oxides can be thought of as a close-packed oxygen host lattice with lithium and metal ions in the interstitial tetrahedral and/or octahedral sites.

2.2.3   Olivine-type Li_xMPO_4

$LiFePO_4$ and $Li_3V_2(PO_4)_3$ have attracted considerable interest as materials with poly-anion groups since their electrochemical activities were first investigated by Padhi *et*

30
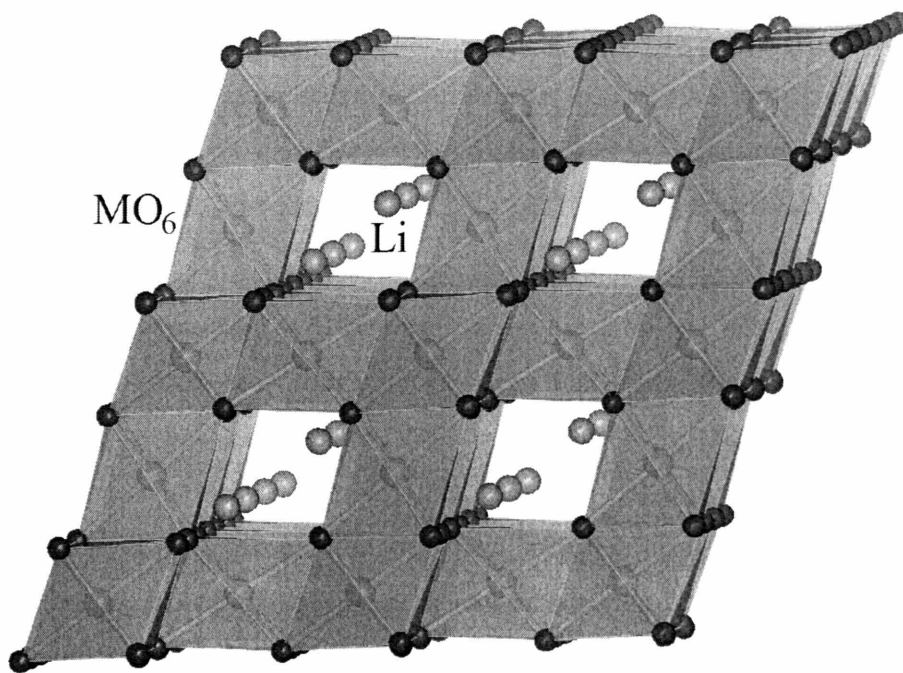
Figure 2-3: The spinel-like structure when half-lithiated ($x = 1$, Li atoms taking tetrahedral positions) with MO_6 octahedra and lithium atoms.
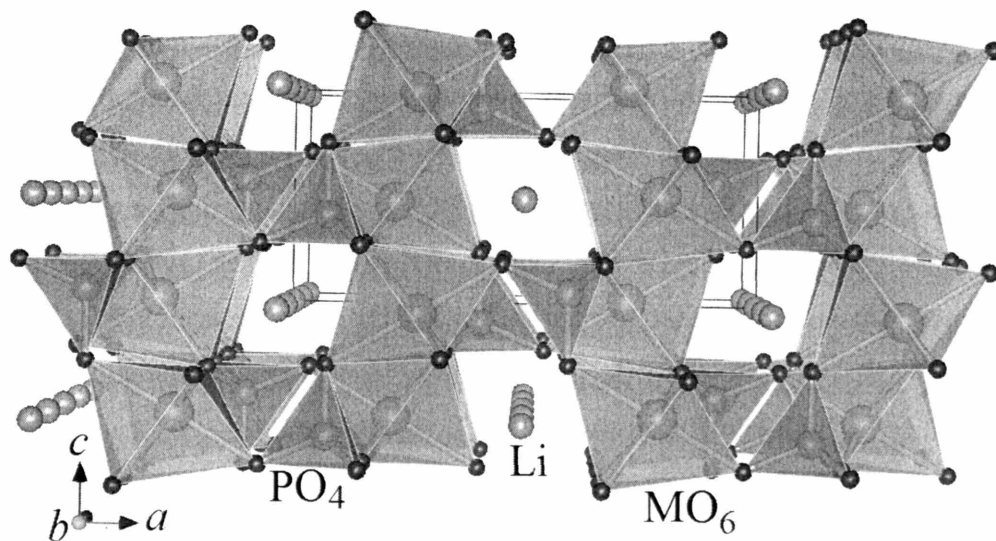
Figure 2-4: Polyhedra representation of the structure of $LiMPO_4$, with arrays of Li ions (green) along b axis, nearly co-planar MO_6 (brown) octahedra and PO_4 (purple) tetrahedra.

al in 1997 [132, 133, 122, 82, 158], as they may be considerably more stable than close-packed oxides at the end of charging. Especially, the superb thermal safety, non-toxicity and low cost of $LiFePO_4$, a naturally occurring mineral, make it the most likely candidate for rechargeable Li-batteries electrodes in large applications such as electric and hybrid vehicles [132, 134, 9, 8, 208, 207, 83, 149, 128, 37, 35, 38, 209, 166, 78, 53, 50, 34, 171, 210].

Olivine-type $LiFePO_4$ and the de-lithiated structure $FePO_4$ have an orthorhombic unit cell with four formula units and space group Pnma (see Fig. 2-4). The lattice constants are A=10.334Å, B=6.008Å, C=4.693Å for $LiMPO_4$ and A=9.821Å, B=5.792Å, C=4.788Å for $FePO_4$ [132]. Inside the material tetrahedral polyanion clusters $(PO_4)^{3-}$ are networked to form distorted oxygen octahedra occupied by lithium and iron ions. The PO_4 groups hybridize less with the TM than an oxygen anion does in simple close-packed oxides, and this leads to more localized $3d$ states on the TM than in an oxide. In a $FePO_4/LiFePO_4$ redox couple *v.s.* Li/Li^+ battery a flat charge/discharge profile over a large compositional range at 3.5 volt [132] indicates that the two-phase redox reaction proceeds via a first-order transition between $FePO_4$ and $LiFePO_4$. The olivine-type structure contains a distorted hexagonal close-packing of oxygen anions, with three types of cations occupying the interstitial sites: 1) corner-sharing FeO_6 octahedra which are nearly coplanar to form a distorted 2-d square lattice perpendicular to the **a** axis, 2) edge-sharing LiO_6 octahedra aligned in parallel chains along the **b** axis, and 3) tetrahedral PO_4 groups connecting neighboring planes or arrays. Electrochemical experiments and X-ray diffraction measurements have confirmed that no intermediate compound Li_xFePO_4 exists between $FePO_4$ and $LiFePO_4$ [132, 134], so that its phase diagram consists of a wide two-phase region with limited solubility on both the $FePO_4$ and $LiFePO_4$ sides. The magnetic structure of $LiFePO_4$ and $FePO_4$ was determined from neutron diffraction data [161, 156]. Below the Nèel temperature $T_N = 50K$ [162, 161] and 125K [156], respectively, the iron spins align in an anti-ferromagnetic (AFM) array collinear with the **b** axis, induced by Fe-O-Fe super-exchange interactions between neighboring iron atoms.

The excellent properties of $LiFePO_4$ has spawned interests in the family $LiMPO_4$

	Mn^{2+}	Mn^{3+}	Mn^{4+}	Fe^{2+}	Fe^{3+}	Co^{2+}	Co^{3+}	Co^{4+}	Ni^{2+}	Ni^{3+}	Ni^{4+}
Olivine	3.92	5.09		3.71	4.90	5.05	6.34		5.26	6.93	
Layered							4.91	5.37		6.70	6.04
Spinel		4.64	5.04				5.62	6.17			
MO [145]	3.6			4.6		5.0			5.1		

Table 2.1: Calculated U in eV.

of olivine structures where the Fe ions are replaced by other metal M=Mn, Co and Ni [132, 128, 78, 50, 209, 105, 7].

2.3 Self-consistent calculation of effective U

In the previous chapter we have outlined the linear-response approach to calculate the effective interaction parameter U. We present here the results of such calculations for the above mentioned lithium metal oxides in Table 2.1 for Mn, Fe, Co and Ni in different structures. For each structure, the effective U (to be precise $U - J$) is calculated for the low and high valence states in a fully lithiated and a de-lithiated structure, respectively. In all cases, except Ni^{3+}/Ni^{4+} in the layered structure, a higher valence state leads to a higher U. For the three cases ($Mn^{3+}/Co^{3+}/Ni^{3+}$) for which we have the parameter in a close-packed (layered or spinel) oxides and in an olivine phosphate structure, U is higher for the olivine structure. This may be related to the fact that the TM-octahedra in the olivine are only corner sharing in two directions but separated from each other by phosphate groups in the third direction, leading to very narrow bandwidth and well localized TM-d states (see the next section for more details). For comparison we also list the U values calculated in [145] for TM monoxides MO (M = Mn, Fe, Co and Ni) in non spin-polarized state. Good agreement with $LiMPO_4$ is found except for Fe^{2+}. We note that in [39] the U value of 4.3 eV for FeO was obtained with the same linear response approach, in good agreement with [145]. So the difference between our results and those in [145] could be mainly due to different crystal environment.

The effective U values in Table 2.1 will be used in the rest of this thesis. In each Li_xMO_y, we take the average of the U values in the high-valence ($U[x = 0]$) and in

the low-valence ($U[x = 1]$) structures. It will be shown that the DFT+U predictions are not very sensitive to the U value in the range between $U[x = 0]$ and $U[x = 1]$ and the average suffices. In certain cases we predict properties as a function of the U parameter to observe its influence.

2.4 Electronic structure and band gap of LiMPO$_4$ and MPO$_4$

2.4.1 Introduction

It is well known that LDA underestimates the band gap of semiconductors by about 50% and incorrectly predicts many insulators as metallic. The origin of this so-called band-gap problem has been attributed to the missing discontinuity in the exchange-correlation potential [140, 138] and the self-interaction error within the LDA [173, 136]. When the DFT+U method was first proposed in the early 1990's, the main purpose was to solve the band gap problem, which it did in many materials [10]. Therefore we first present the comparison of the electronic structure calculations with LDA/GGA and DFT+U in phospho-olivines (Li)MPO$_4$. The general agreement of the DFT+U results with experiment validates further investigations on total energy properties.

There is also significant practical interest in the electronic properties of Li$_x$MPO$_4$ (M=Fe, Mn, Ni, Co) because of the importance of electronic conductivity in batteries for high rate applications. A major challenge in using the olivine phosphates is their low electrical conductivity. For example, the most studied Li$_x$FePO$_4$ in its pure form has very poor conductivity, greatly inhibiting high-rate applications [37, 78]. Similar problems are believed to inhibit Li exchange from Li$_x$MnPO$_4$ [207, 209]. Efforts to increase conductivity of electrodes made from these materials have focused on particle size reduction [208], intimate carbon coating [83], cation doping [37] and creation of conduction network [78]. Significant disagreement exists on the origin of the low electronic conductivity. *Ab initio* studies with LDA/GGA have found a small

gap, or no gap at the Fermi level, which seems to be in contradiction to experiment [207, 209, 177, 166, 206]. For example, Xu, *et al.* [206] found that $LiFePO_4$ is a semi-metal, which seems surprising, given the experimentally observed lack of electrical conductivity. Our results with the more appropriate DFT+U method can help clarify the conductivity problem in $LiFePO_4$.

2.4.2 Computational Methods

All calculations shown in this chapter are performed within either the Generalized Gradient Approximation (GGA) [199] or the GGA+U [12, 106], with the projector-augmented wave (PAW) method [23, 94] as implemented in the Vienna Ab-initio Simulation Package (VASP) [92]. An energy cut-off of 500 eV and appropriate k-point mesh were chosen so that the total ground state energy is converged to within 3 meV per formula unit.

In order to understand the possible impact of magnetic structure, we have performed calculations with both ferromagnetic (FM) and antiferromagnetic (AFM) orderings of the magnetic transition metal ions. The AFM ordering for all calculations was taken from the magnetic states of $LiFePO_4$ and $LiMnPO_4$, which have been determined experimentally to be AFM within the approximately simple square lattices formed by the planes of transition metal cations in the olivine structure [113, 156]. For completeness we note that the low-temperature magnetic state of $FePO_4$ is non-collinear and slightly different from $LiFePO_4$ [156], and that at higher temperatures all these systems will have magnetic disorder. All the atoms and cell parameters of each structure are fully relaxed, and the lattice parameters for the Li_xMPO_4 compounds. The rotationally invariant [106] form of GGA+U is used with a spherically averaged double counting term [61]. Within this approach the on-site coulomb term U, and the exchange term J, can always be grouped together into a single effective parameter $(U - J)$ [61], and this effective parameter will simply be referred to as U in this chapter. According to Table 2.1 we use $U = 4.3, 4.5, 6$ and 6 eV for Fe, Mn, Ni and Co, respectively.

2.4.3   Results with GGA and GGA+U

Fig. 2-6 shows the FM total density of states (DOS) for Li_xMPO_4 (x = 0, 1 and M = Fe, Mn), with $U = 0$ (normal GGA) and $U = 4.3$ and 4.5 eV, for Fe and Mn, respectively, and Fig. 2-5 shows the AFM DOS. Fig. 2-7 shows the FM DOS for M= Ni and Co, respectively. Table 2.2 compares the band gaps in the different approximations, magnetic orderings, and materials, along with values from previous work. * The calculated band gaps show some sensitivity to the choice of magnetic ordering, particularly for the $MnPO_4$ material, perhaps due to coupling of the magnetic ordering and the Jahn-Teller distortion (for more information on Jahn-Teller and magnetic coupling in Mn^{3+} see [110]. However, it is clear that the qualitative impact of changing from LDA/GGA to DFT+U methods (here DFT+U is used to refer to both LDA+U and GGA+U methods) does not depend on the magnetic ordering. The previous works quoted in the last column of Table 2.2 are all obtained without U corrections, and therefore should be compared to the GGA data from this study. The results show that our pure GGA results are consistent with the ranges found in the previous literature. In general, more accurate treatment of the on-site correlations through GGA+U clearly yields larger band gaps, in some cases quite dramatically. $NiPO_4$ is the only exception where a band gap does not appear even with the introduction of U corrections, due to very strong Ni-O rehybridization. A general trend is that the band gaps in MPO_4 are smaller than in $LiMPO_4$. Comparison of the DOS shows that a simple band-filling picture applies as a first approximation. *Therefore generally the band gap is more Mott-Hubbard like in $LiMPO_4$ and closer to the charge-transfer type in MPO_4, resulting in its decrease.*

The most definitive evidence that DFT+U is more accurate than LDA/GGA for band gaps in olivine phosphate materials comes from $LiFePO_4$. The calculated band gap from GGA is somewhere in the range 0-0.3 eV, with some disagreement between

---

*All results from other work are ferromagnetic (FM) unless specified as AFM, and GGA results are quoted when available. These results are obtained with a range of different methods and/or approximations, and therefore the original references must be consulted for precise comparisons with each other or the present work. However, the previous results are useful for establishing the range of values obtained using traditional LDA and GGA methods.
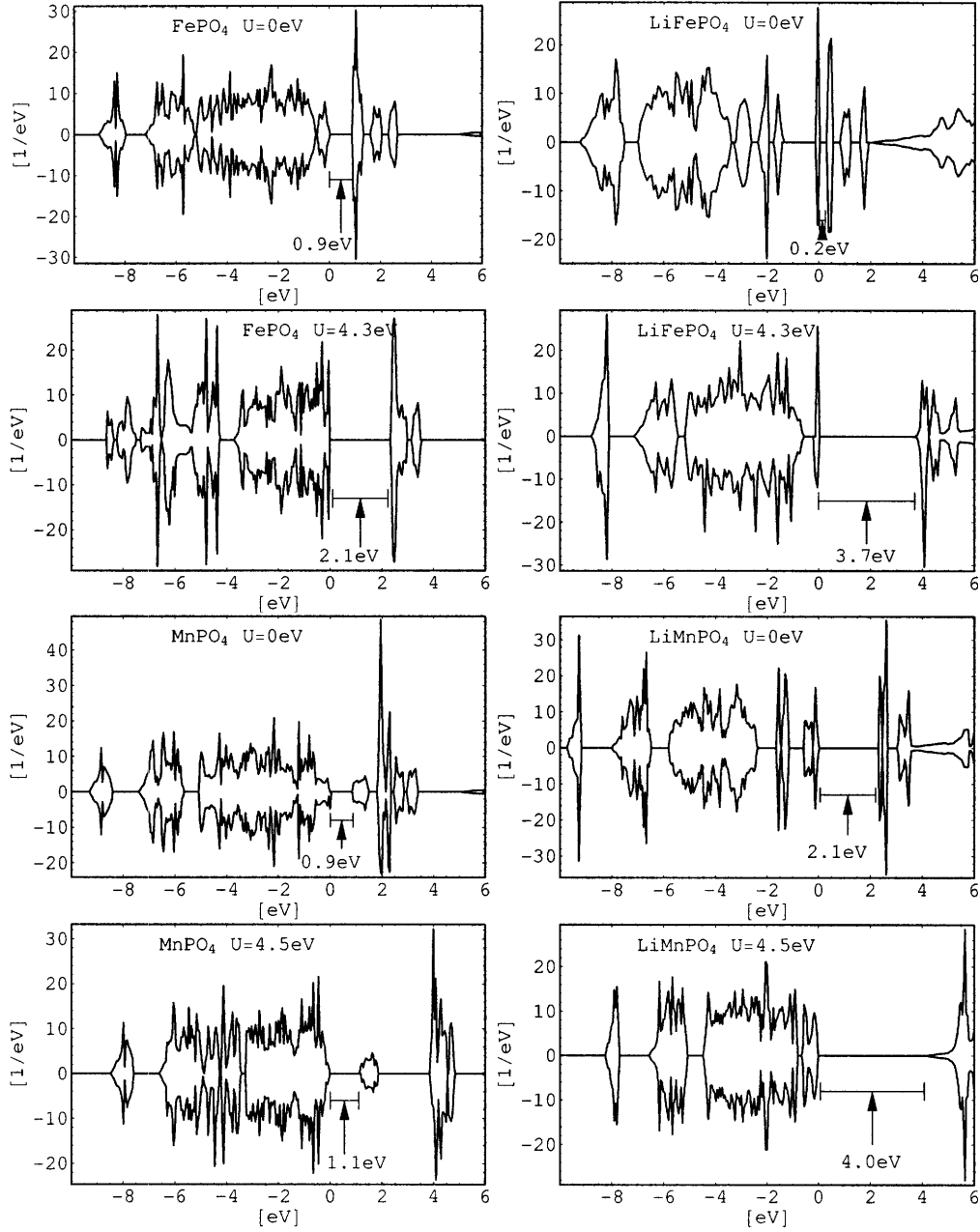
Figure 2-5: The anti-ferromagnetic total density of states for Li_xMPO_4 (x = 0, 1 and M = Fe, Mn), with U = 0 (normal GGA) and U = 4.3 and 4.5, for Fe and Mn, respectively.
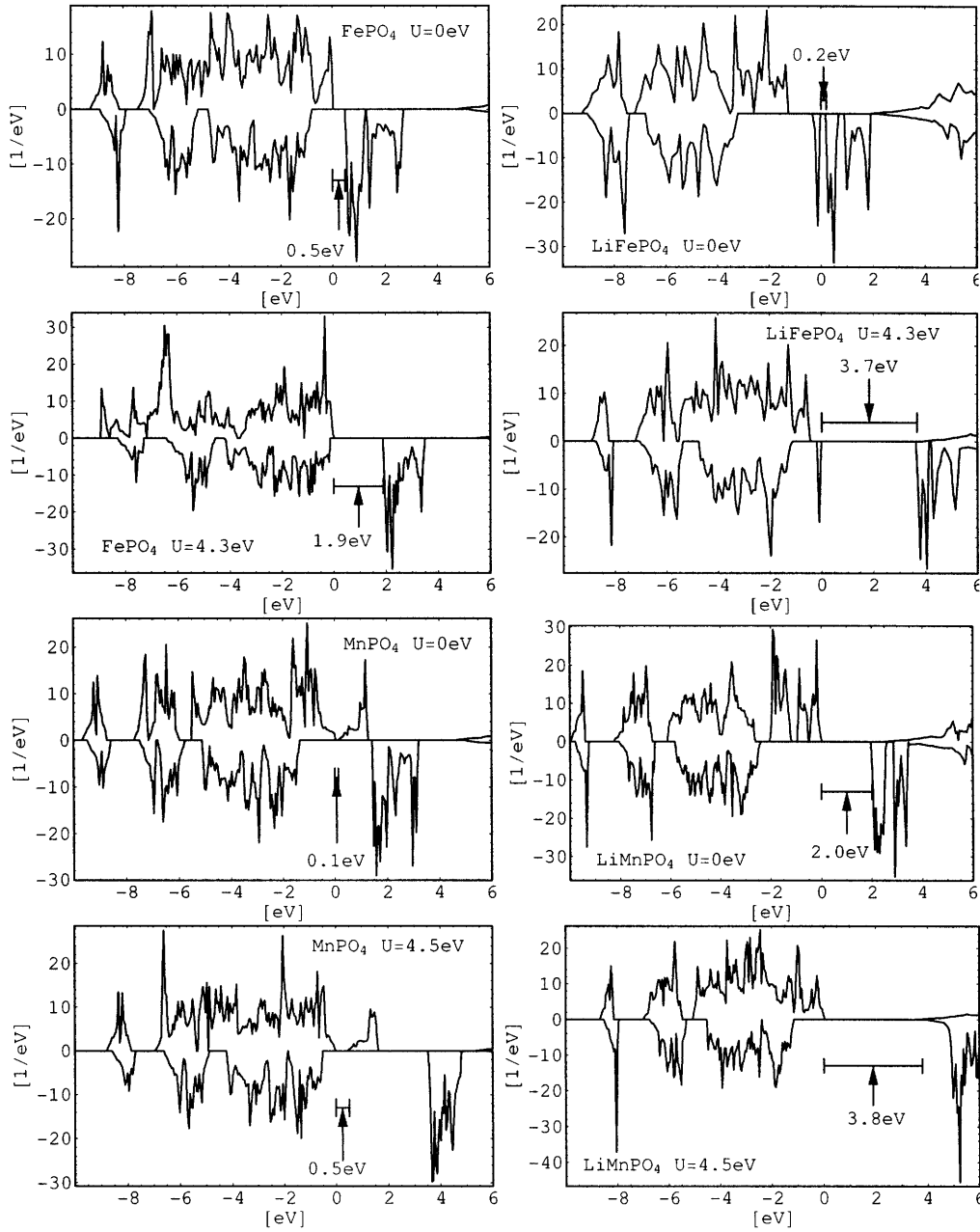
Figure 2-6: The ferromagnetic total density of states for Li_xMPO_4 (x = 0, 1 and M = Fe, Mn), with U = 0 (normal GGA) and U = 4.3 and 4.5, for Fe and Mn, respectively. The positive (negative) axis is the majority (minority) spin direction. Note that for $MnPO_4$ the FM gaps are considerably different for AFM gaps in Table 1.

38
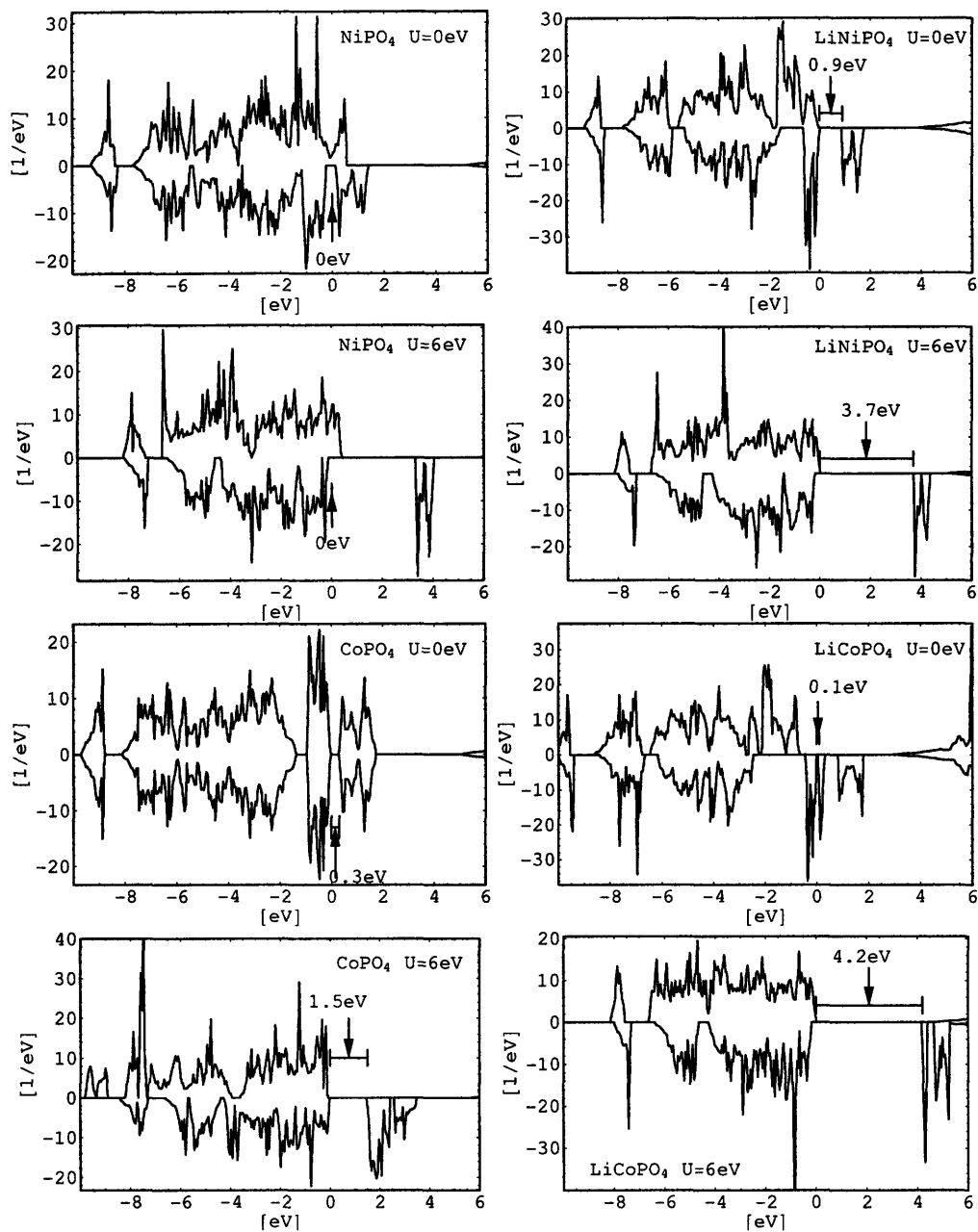
Figure 2-7: The ferromagnetic total density of states for Li_xMPO_4 (x = 0, 1), with U = 0 (GGA) and U = 6, for Ni and Co, respectively.

System	GGA Gap (FM)	GGA+U (FM)	GGA (AFM)	GGA+U (AFM)	Other work (FM)
FePO4	0.5	1.9	0.9	2.1	0.4 [207], 0.3 (0.7 AFM) [177]
LiFePO4	0.2	3.7	0.2	3.7	0.3 [207], 0.0 (0.0 AFM) [177], 0.2 [166], 0.0 [206]
MnPO4	0.1	0.5	0.9	1.1	—
LiMnPO4	2.0	3.8	2.1	4.0	2 [207], 1.7 [177]
CoPO4	0.3	1.5			
LiCoPO4	0.1	4.2			~ 0.5 [177]
NiPO4	0.0	0.0			
LiNiPO4	0.9	3.7			~ 0.5 [177]

Table 2.2: Calculated band gaps in eV for $LiMPO_4$ and MPO_4, from this work and previous references.

different authors (see Table 2.2). However, the GGA+U result is about 3.7 eV, which is close to the value of 3.8–4.0 eV from diffuse reflectance spectroscopy [219] and 4.0 eV from soft-x-ray-absorption and -emission spectroscopies [16].

2.4.4 Discussion and conclusion

Having established that the GGA+U method produces a more accurate DOS for this class of compounds, it is important to consider the implications of the DOS for the electronic conductivity. A large gap will lead to a very small number of intrinsically generated electrons or holes. For example, based on a 3.9 eV gap, all the charge carriers in $LiFePO_4$ up to 300 K should be introduced extrinsically, either by impurities or by Li deficiency, ruling out a conduction mechanism by delocalized electrons in the conduction band. Recently, our collaborators have studied the polaron mechanism of electronic conduction in $(Li)FePO_4$ [112].

In summary, our computational studies of the Li_xMPO_4 (M=Fe, Mn, Ni, Co) olivine materials reveal significant band gaps, largely induced by strong electron correlation at the transition metals. The calculated band gap agrees well with two inde-

pendent experimental reports. The generally more accurate description of electronic structure support the further study of energetic effects in lithium metal oxides.

2.5 Redox potentials

2.5.1 Introduction

Redox processes are relevant to many technological applications, including corrosion, fuel cells and rechargeable Li batteries, and the ability to study these processes from first principles is therefore crucial. In terms of lithium ion batteries, one critical property is the redox potential at which Li can be removed and inserted. While a high potential increases the energy density of the material, if this potential is too high, Li can not be practically removed, and side reactions such as electrolyte breakdown can occur in the cell. A low potential can lead to moisture sensitivity of the electrode material.

First principles calculations have led to considerable success in predicting the trends of Li insertion voltages [18] and even new phases [188], but it has been noted that LDA or GGA can give relatively large errors for the average Li insertion potential [18, 123]. For example, Table 2.3 compares the experimental voltage for different structures with the one calculated in the GGA approximation using computational settings discussed in section 2.5.3. The Li insertion potential is consistently under-

	$LiNiO_2/NiO_2$	$LiMn_2O_4/Mn_2O_4$	$LiFePO_4/FePO_4$
GGA	3.19	3.18	2.97
exp.	3.85[54]	4.15 [127]	3.5 [208]

Table 2.3: Calculated and experimental redox couple voltage in Volt.

predicted by as much as 0.5 to 1.0V. Similar results have been obtained with LDA [18].

To understand the origin of the underestimation of redox potential in LDA/GGA, we note that the key to a redox reaction is the transfer of electrons from one species to another. When the redox electron is transferred between very distinct environments

(e.g. delocalized s state to localized d state), the LDA and GGA may encounter problems. In this section, we argue that this inaccuracy is related to the lack of cancellation of electron self-interaction errors in LDA/GGA and can be improved by using the DFT+U method with a self-consistent evaluation of the U parameter. We show that, using this approach, the experimental lithium intercalation voltages of a number of transition metal compounds can be reproduced accurately.

2.5.2 Relation between redox potential and total energies

When Li is inserted into a TM-oxide, its charge is compensated by an electron absorbed from the external circuit. The insertion reaction is symbolized by the following equation:

$$\Delta x Li + Li_x MO_y \Longleftrightarrow Li_{x+\Delta x} MO_y, \tag{2.1}$$

where MO_y is the TM compound host material. Using thermodynamical arguments, it is possible to relate the voltage V of the cell to the lithium chemical potential (μ_{Li}) on both sides of Eq. 2.1 in the cathode [26]:

$$V(x) = -\frac{\mu_{Li(x)}^{cathode} - \mu_{Li}^{anode}}{F}. \tag{2.2}$$

F is the Faraday constant, and μ_{Li}^{anode} is the chemical potential in the anode, or more generally, the chemical potential of the Li source.

The average voltage $\langle V \rangle$ for Li insertion between two composition limits, $Li_{x_1} MO_y$ and $Li_{x_2} MO_y$, can be found by integrating Eqn. 2.2 (usually between $x = 0$ and 1), and is determined by the free energy of the compounds at the composition limits [18]. Neglecting the entropic and $P\Delta V$ contributions [18], $\langle V \rangle$ can simply be determined by computing the total energy of $Li_{x_2} MO_y$, $Li_{x_1} MO_y$ and Li:

$$\langle V \rangle = \frac{-[E(Li_{x_2} MO_y) - E(Li_{x_1} MO_y) - (x_2 - x_1)E(Li)]}{(x_2 - x_1)F}. \tag{2.3}$$

Typically $x_1 = 0$ and $x_2 = 1$ are taken as composition limits, as in these cases no Li-vacancy disorder occurs.

Experimentally, the voltage vs. lithium composition curve $V(x)$ can be conveniently measured for both the charging and the discharging processes. The corresponding curves differ in general because of the overcharge potential present in the circuit. We obtain the experimental average open circuit voltage values by numerically averaging the charge and discharge curves published in Refs. [54, 127, 208, 105, 7, 36] over the appropriate composition range.

2.5.3 Details of the calculations

We use the same computational settings as in section 2.4. Jahn-Teller distortions are allowed where the transition metal ions are Jahn-Teller active (Mn^{3+} and Ni^{3+} in our case) by explicitly breaking the symmetry of the unit cell. Our relaxed cells of layered $LiNiO_2$ and spinel $Li_2Mn_2O_4$ agree well with the calculations in [110] on Jahn-Teller distorted systems using GGA. All calculations are performed with spin-polarization. As discussed later, the total energy of a given structure depends critically on the magnetic state of the metal ions, and high-spin states are favored by the DFT+U scheme we use. The ordering of the spin on the ions in different magnetic structures (i.e. ferromagnetic, anti-ferromagnetic or more complicated ordering) results in difference in the total energy of the order 10–60 meV per formula unit. From the total energies, the average lithiation potential can be calculated through Eq. 2.3.

Figures 2-8 and 2-9 show respectively the average Li insertion voltage as function of U in the olivine, and in the layered and spinel structure. The horizontal short line indicates the experimentally measured voltage. Three calculated points for each system are marked on the curve: the small open circles indicate respectively the voltage one would obtain using the calculated U for the most reduced and most oxidized TM-state in each structure (e.g. Fe^{2+} and Fe^{3+} in $LiFePO_4$). The large filled circle corresponds to the voltage for the averaged U. The results for each system are discussed in more detail below.
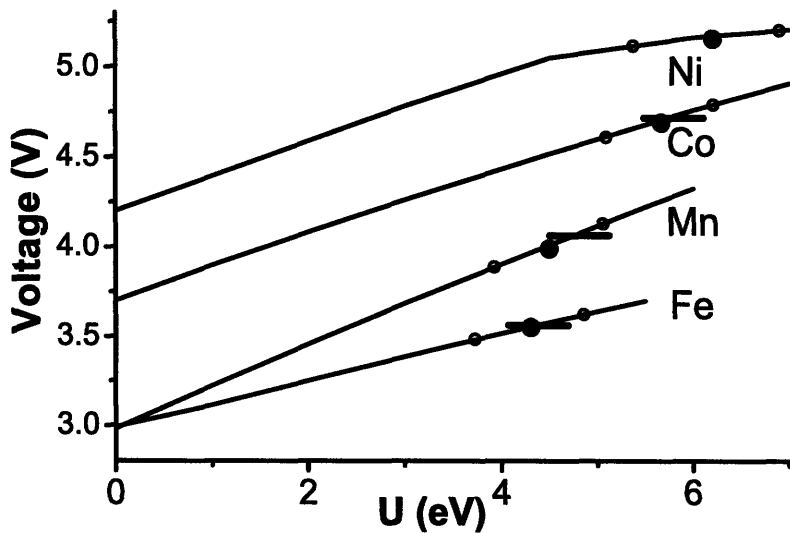
Figure 2-8: Voltage as a function of U for the $LiMPO_4$ materials in the olivine structure. The short horizontal lines on the curves indicate the experimental voltage of the each material. The two small open circles on a curve represent the voltage for U calculated in the oxidized (delithiated) or reduced (lithiated) states. The big solid circle represents the voltage at the average of the two U values.
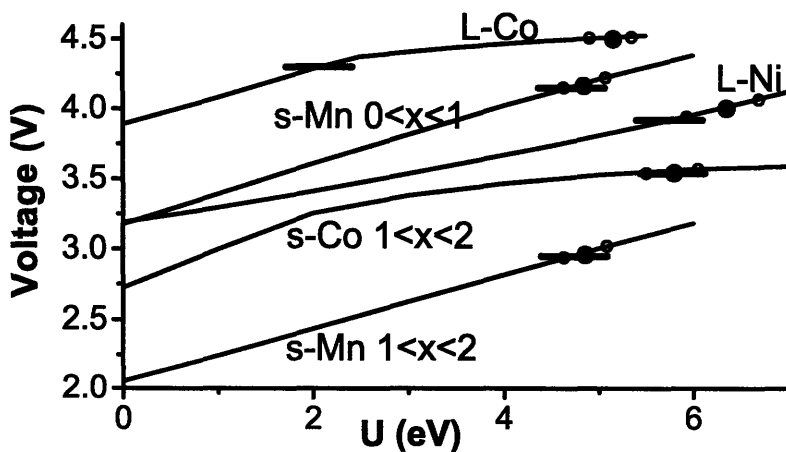
Figure 2-9: Voltage as a function of U for the layered and spinel structures. Legend the same as in Fig. 2-8.

2.5.4 Olivine phosphates Li$_x$MPO$_4$ (M=Mn, Fe, Co, Ni)

Fig. 2-8 have been calculated with AFM spin configuration in both end members. The calculated and experimental cell parameters, as well as the electronic occupation of the TM ions are listed in Table 2.4.

		a (Å)	b (Å)	c(Å)	V(Å^3)	TM ion
LiMnPO$_4$	GGA	10.55	6.13	4.78	309.13	$t_{2g}^3 e_g^2$
	GGA+U	10.62	6.17	4.80	314.52	$t_{2g}^3 e_g^2$
	Exp. [105]	10.44	6.09	4.75	302.00	
MnPO$_4$	GGA	9.92	6.01	4.93	293.92	$t_{2g}^3 e_g^1$
	GGA+U	9.98	6.07	4.96	300.47	$t_{2g}^3 e_g^1$
	Exp.[105]	9.69	5.93	4.78	274.67	
LiFePO$_4$	GGA	10.39	6.04	4.75	298.09	$t_{2g}^4 e_g^2$
	GGA+U	10.42	6.07	4.76	301.07	$t_{2g}^4 e_g^2$
	Exp. [132]	10.33	6.01	4.69	291.39	
FePO$_4$	GGA	9.99	5.93	4.90	290.28	$t_{2g}^3 e_g^2$
	GGA+U	9.99	5.88	4.87	286.07	$t_{2g}^3 e_g^2$
	Exp. [132]	9.82	5.79	4.79	272.36	
LiCoPO$_4$	GGA	10.30	5.93	4.75	290.13	$t_{2g}^5 e_g^2$
	GGA+U	10.33	5.97	4.76	293.55	$t_{2g}^5 e_g^2$
	Exp. [7]	10.20	5.92	4.70	283.90	
CoPO$_4$	GGA	9.71	5.48	4.59	244.24	t_{2g}^6
	GGA+U	9.98	5.78	4.74	273.42	$t_{2g}^4 e_g^2$
	Exp. [7]	10.09	5.85	4.72	278.66	
LiNiPO$_4$	GGA	10.09	5.91	4.74	282.66	$t_{2g}^6 e_g^2$
	GGA+U	10.12	5.90	4.73	282.42	$t_{2g}^6 e_g^2$
	Exp. [68]	10.03	5.85	4.68	274.49	
NiPO$_4$	GGA	9.66	5.72	4.71	260.25	$t_{2g}^6 e_g^1$
	GGA+U	9.92	5.82	4.84	279.43	$t_{2g}^6 e_g^1$

Table 2.4: Cell parameters of the olivine LiMPO$_4$ and MPO$_4$, as well as the corresponding electron configuration at the transition metal ions.

Mn Both Mn^{2+} and Mn^{3+} are high-spin ions in GGA and GGA+U calculations. Attempts to constrain them to lower spin states lead to much higher energy. FM ordered magnetic structures are 10 - 30meV higher in energy than the AFM ordered magnetic structure as U is varied. A strong collective Jahn-Teller distortion is observed in MnPO$_4$, where Mn^{3+} is in the high-spin $t_{2g}^3 e_g^1$ state, in GGA(+U). The experimental voltage for the MnPO$_4$/LiMnPO$_4$ redox couple

has been obtained from Ref. [105]. The voltage predicted with GGA+U (4.04 V at $U = (U_{\text{Mn}^{2+}} + U_{\text{Mn}^{3+}})/2$) is within a few percent of the experimental voltage (4.1 V), and in sharp contrast to the large error made by GGA ($V_{\text{GGA}} = 2.98$ V).

Fe Both Fe^{2+} and Fe^{3+} are high-spin in GGA(+U) calculations, and the AFM ordering is more stable than FM ordering. Using $U_{\text{Fe}^{2+}}$ and $U_{\text{Fe}^{3+}}$ we calculated a voltage of 3.39 and 3.55 V respectively. The voltage calculated with the average $U= 4.30$ eV is 3.47 V, which agrees very well with the experimentally measured value of 3.5 V [208]. This is a substantial improvement over the GGA predicted value of 2.97 V. Previously, the localization of electrons induced by U was also shown to qualitatively affect the phase behavior in this system [220].

Co In LiCoPO$_4$ Co^{2+} is stable in the high-spin $t_{2g}^5 e_g^2$ state. In the delithiated CoPO$_4$, Co^{3+} is stable as non spin-polarized with GGA, but more stable by several eV with GGA+U in the high spin $t_{2g}^4 e_g^2$ configuration at the calculated U value of 6.34 eV. As shown in Table 2.4 the cell parameters of CoPO$_4$ calculated with non spin-polarized Co^{3+} in GGA is appreciably smaller than experimental values, while GGA usually slightly overestimates cell parameters. With GGA+U and high-spin Co^{3+} the calculated parameters are close to experimental values. While there is only limited electrochemical data on this material [132], the predicted voltage of 4.73 V at U_{average} is within a few % of the result 4.8V established by Anime *et. al* [7], compared to the poor GGA prediction of 3.70 V. The high voltage of this material makes it particularly attractive for high-energy density applications.

Ni At $x = 1$ Ni^{2+} is stable as high-spin $t_{2g}^6 e_g^2$. At $x = 0$ Ni^{3+} occurs in the low spin state $t_{2g}^6 e_g^1$ for both GGA and GGA+U , but the high spin state $t_{2g}^5 e_g^2$ is less unstable in GGA+U than in GGA. Note that low-spin Ni^{3+} is a weak Jahn-Teller ion, and no appreciable collective distortion is observed in our relaxed unit cell. With U_{average}, a voltage of 5.07 V is obtained. This prediction was later confirmed by experimentally measured lower bound of 5 V [201].

2.5.5 Layered Li_xMO_2 (M=Co, Ni)

For the layered and spinel structures AFM spin ordering on transition metal ions is topologically frustrated, and their actual magnetic ground states are not always clear in experiment. But as the energy associated with different magnetic orderings is small, the simple FM ordering is used in the following calculations.

Co In $LiCoO_2$ Co^{3+} is stable in the non spin-polarized state for the calculated $U_{Co^{3+}} = 4.91$eV. At $x = 0$, Co^{4+} is almost degenerate in either non spin-polarized or spin-polarized t_{2g}^5 in GGA, but more stable with spin-polarization in GGA+U at the calculated $U_{Co^{4+}} = 5.37$eV. While GGA+U still improves the agreement of voltage with experiment [127] over pure GGA, the error for this system is larger than in the other systems we calculated. This might be related to the fact that the GGA result is already closer to experiment than for all other systems.

Ni In $LiNiO_2$ Ni^{3+} is most stable in the low-spin $t_{2g}^6 e_g^1$ state and is a weak Jahn-Teller ion. With GGA a distorted unit cell is found with the short and the long Ni-O bond length being 1.92Å and 2.13Å, respectively, compared to experimental values of 1.91Å and 2.14Å [110], and a stabilization energy relative to an undistorted cell of only -2meV, within the range of numerical errors, compared to -11 meV in Ref. [110]. With GGA+U no appreciable distortion is observed. Experimentally there is no cooperative Jahn-Teller distortion in $LiNiO_2$ though the Ni-O octahedra are locally Jahn-Teller distorted [155], suggesting a very small stabilization energy,consistent with both GGA and GGA+U results. At $x = 0$, Ni^{4+} is stable as a non spin-polarized ion. The GGA+U voltage value of 3.92V agrees well with the experimental average voltage of 3.85V [54], and is substantially better than the GGA result of 3.19V.

2.5.6 Spinel $Li_xM_2O_4$ (M=Mn, Co)

For the spinel $Li_xMn_2O_4$ there are two distinct plateaus in the voltage profile, between $0 < x < 1$ and $1 < x < 2$, respectively. For $0 < x < 1$ Li enters tetrahedral sites,

while the reaction from $LiMn_2O_4$ to $Li_2Mn_2O_4$ occurs through a two-phase process whereby the $LiMn_2O_4$ phase with only tetrahedral Li disappears at the expense of the $Li_2Mn_2O_4$ phase with all Li octahedral. Calculations were done for $x = 0$, 1 and 2 structures to get separate average voltage values for the two processes. For M = Co the $0 < x < 1$ reaction potential curve is difficult to obtain accurately in experiments. Therefore only the average voltage for the $1 < x < 2$ reaction is shown in fig. 2-9.

Mn Both Mn^{4+} and Mn^{3+} are high-spin. Mn^{3+} is a strong Jahn-Teller active ion. In GGA, the calculated Mn-O short and long bond lengths 1.94Å and 2.40Å agree with Ref. [110]; in GGA+U they become 1.96Å and 2.32Å, respectively. Experimental values are 1.94Å and 2.29Å, respectively [81], showing that the good structural prediction of GGA is retained in GGA+U. Coexistence of distinct Mn^{4+} and Mn^{3+} is found in GGA+U in the LiM_2O_4 compound. The GGA+$U_{average}$ results (4.19V and 2.97V respectively, for the first and second plateaus) is in excellent agreement with the experimentally measured values of 4.15V and 2.95V [127].

Co Like in the layered structure, Co^{3+} in $Li_2Co_2O_4$ is non spin-polarized, and at $x = 0$ Co^{4+} is more stable as spin polarized t_{2g}^5 in GGA+U. The GGA+U voltage (3.56V at $U_{average}= 4.84$eV) agrees very well with experimental data available for the $Li_1Co_2O_4$ to $Li_2Co_2O_4$ reaction (3.5V [36]).

Note that in the $x = 1$ structure of the spinel materials $Li_xM_2O_4$ we find distinct M^{3+} and M^{4+} ions in GGA+U instead of ions of intermediate valence. The same phenomenon was observed in the intermediate structures Li_xFePO_4 of the iron phosphate [220]. This is a direct consequence of the E_U correction term to the total energy in Eq. 1.1 which penalizes the non-integral occupation of the d-orbitals. Such charge ordering is necessary for correctly predicting the $0 < x < 1$ and $1 < x < 2$ average voltage values of $Li_xMn_2O_4$ simultaneously, as well as the $1 < x < 2$ voltage of $Li_xCo_2O_4$, and is not present in pure GGA unless localization is assisted by a strong polaronic contribution such as the Jahn-Teller distortion around Mn^{3+}.
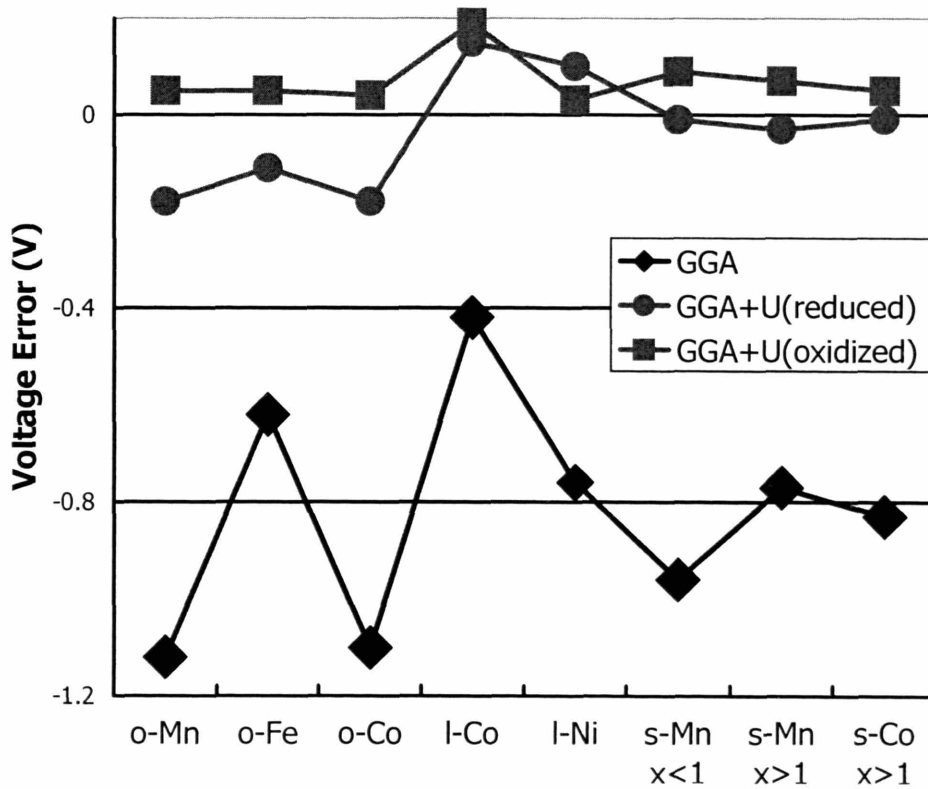
Figure 2-10: Difference between calculated and experimental voltage [5-9], for GGA and GGA+U, at the calculated U of the oxidized (delithiated) and reduced (lithiated) states, respectively (l=layered, s=spinel). For the spinel structures two voltage values for the $0 < x < 1$ and $1 < x < 2$ plateaus are calculated separately. Olivine $LiNiPO_4$ is not shown here because the voltage is unknown.

2.5.7 Discussions and conclusions

Introduction of on-site correlations in GGA through the GGA+U clearly improves predicted lithiation potentials considerably over the use of pure GGA (or LDA for that matter). The errors of GGA+U and pure GGA on all systems for which we have experimental data are summarized in Fig. 2-10. Pure GGA consistently underestimates the lithiation voltage, which is a measure of the energy lowering when Li is transferred from Li metal (the anodic reference) to a Li^+ ion and electron in the TM oxide or phosphate. The contribution of the Li^+ ion to the reaction energy is largely electrostatic, and one would expect this effect to be well captured in GGA or LDA. Hence, the large voltage error in LDA/GGA must arise from the electron transfer from Li metal to the TM cation. Since the voltage is always underestimated in LDA/GGA these approximations clearly penalize the energy of the electron on the TM. It seems reasonable to attribute this to the poor treatment of electronic correlations in LDA/GGA. In metallic lithium the electron is affected by a small self-interaction in LDA/GGA as its charge density is delocalized. On the TM ion, however, the electron occupies a much more localized d-orbital and will experience a much larger self-interaction. The lack of cancellation between the self-interactions contributions to the energy, which are related to an improper description of the correlation effects in LDA/GGA, leads to a systematic error in the prediction of the redox potential. In the direction in which the electron is transferred from a delocalized to a localized state, the reaction energy is penalized (not negative enough), making the potential too small. The use of GGA+U allows for a better description of the electronic correlation and, by discouraging fractional occupations of the Kohn-Sham orbitals, removes the spurious self-interaction thus producing a much more accurate prediction of the redox voltage. While we demonstrate the GGA/LDA problem and improvement obtained with DFT+U on Li-insertion materials, we believe that a more accurate description of correlation effects within the DFT+U scheme is also necessary in the study of other redox processes in which electrons are transferred between states of different kind (e.g. catalysis of organic molecules on TM surfaces). In fact,

as explained in Ref. [39], a better description of the electronic correlation (which enforces the independence of the single electron energy eigenvalues of the partially occupied states on their occupation, thus leading to the elimination of the spurious self-interaction) is needed to reproduce the physical difference among the ionization potential and the electronic affinity (or the band gap in crystalline solids) which plays a very important role in the energetics of processes involving electron transfer.

In our calculations high-spin TM ions are always energetically favored by GGA+U over low-spin or non spin-polarized states. In $CoPO_4$ the non spin-polarized Co^{3+} in GGA leads to cell parameters inconsistent with experiment. In GGA+U Co^{3+} becomes high spin, improving agreement with experiment. For the other systems the GGA and GGA+U cell parameters are rather close, though GGA+U seems to lead to volumes that are slightly too high. Jahn-Teller distortions predicted by GGA are also reproduced in GGA+U for Mn^{3+}.

In summary, we have shown that the under-estimation of the lithium intercalation voltage in LDA/GGA can be corrected by using GGA+U with a self-consistently calculated parameters U, without sacrificing properties that are already accurately predicted by GGA (e.g. Jahn-Teller effect, cell parameters, magnetic ordering). Voltages for most systems are predicted within a few % of experimental values.

We believe that DFT+U will significantly improve the accuracy of voltage prediction for candidate materials can be predicted, and therefore enhance the capability of screening new materials for their ability to be good cathodes.

2.6 Low temperature phase stability in Li_xFePO_4

2.6.1 Introduction

Previously, Section 2.4 confirms that the well-documented band-gap problem can be treated by the introduction of SIC-like LDA+U corrections for the Li_xMPO_4 systems, resulting in more accurate electronic structures. Section 2.5 establishes the method's improvement in predicting TM redox potential and other properties in lithium metal

oxides. Effects of electron correlation (or errors of LDA/GGA) in predicting structure and phase stability in oxides are less well characterized. In this section we focus on a dramatic and qualitative failure of LDA/GGA in olivine type $LiFePO_4$. At room temperature electrochemical Li removal proceeds through a two phase region between the end members $LiFePO_4$ and $FePO_4$, and no intermediate compound Li_xFePO_4 exists in between [132]. Such instability of Li_xFePO_4 is not seen in many other intercalated TMOs, e.g. Li_xCoO_2, Li_xNiO_2, Na_xCoO_2 [189, 179, 48]. In spinel $Li_xMn_2O_4$ there is a two-phase region in $1 \leq x \leq 2$ due to strong electrostatic repulsion between face-sharing tetrahedral and octahedral lithium sites, e.g. a specific geometry factor of the lattice. The reason for phase separation in Li_xFePO_4 is not obvious at all. The role on-site correlations play in phase stability is tentatively probed in this section. A thorough investigation of phase stability in Li_xFePO_4 can be adequately addressed only when electronic excitations out of the ground state are accounted for, therefore will be presented in the next chapter when we deal with thermodynamics involving localized electrons.

2.6.2 Method

To investigate the stability of compounds between the composition $FePO_4$ and $LiFePO_4$ different Li/vacancy arrangements are considered in the primitive cell. All possible symmetrically distinct decorations of the 4 Li sites give seven structures, including two end members (x=0, 1), one structure at each of x=0.25 and 0.75, and three at x=0.5 (named 0.5a,b,c). The structures 0.5a, 0.5b and 0.5c have Li remaining at sites 1 and 2, 1 and 3, and 1 and 4, respectively (see Table 2.5). All the five intermediate structures have lower symmetry than the end members, and are monoclinic or triclinic. Total energy calculations were performed for the seven structures in GGA (or LDA where explicitly stated) with the same computational settings as those in previous sections. All the atoms and cell parameters are fully relaxed at each structure. For x=0.25 and 0.75 the remaining S_2 point group symmetry has to be removed to get the electronic ground state (see below). The results in this paper represent the ferromagnetic (FM) spin-polarized configurations unless stated explicitly, for reasons

	Li 1	Li 2	Li 3	Li 4	Fe 1	Fe 2	Fe 3	Fe 4
x	0	0.5	0.5	0	.28	.22	.78	.72
y	0	0	0.5	0.5	.25	.75	.25	.75
z	0	0.5	0.5	0	.98	.48	.52	.02

Table 2.5: Fractional positions of the four Li and four Fe atoms within the $LiFePO_4$ unit cell.

x	0.25	0.5a	0.5b	0.5c	0.75
LDA	-155	-255	-247	-136	-168
GGA	-135	-209	-197	-129	-138

Table 2.6: LDA and GGA formation energy (meV/$LiFePO_4$) at different Li concentrations.

stated in previous sections.

Here we define $\Delta E(x)$, the formation energy per $Li_x FePO_4$ as

$$\Delta E(x) = E(x) - (x E (x = 1) + (1 - x) E (x = 0)) \tag{2.4}$$

where $E(x)$ is the ground state total energy per formula unit for $Li_x FePO_4$. A negative formation energy means compound formation is energetically favorable. In order for phase separation to occur at room temperature, all intermediate structures should have positive formation energy, large enough to overcome the mixing entropy. LDA results of $\Delta E(x)$ for all five structures, shown in Table 2.6, are negative. Although GGA slightly increases the formation energy, the prediction remains in qualitative disagreement with experiment.

Given that the true formation energies should all be positive, these errors are large and somewhat surprising, since formation energies are properly weighted energy differences between similar structures, and as such usually benefit from significant error cancellations. For example, in many binary alloys formation energies are only 100~200meV/atom in magnitude, and hence large errors such as those found here would make them completely unreliable, which, based on the good agreement of many LDA/GGA studies with experiment, is not the case [49].

To investigate whether on-site correlation effects could be related to this substantial failure of LDA/GGA, we have performed GGA+U calculations with one effective
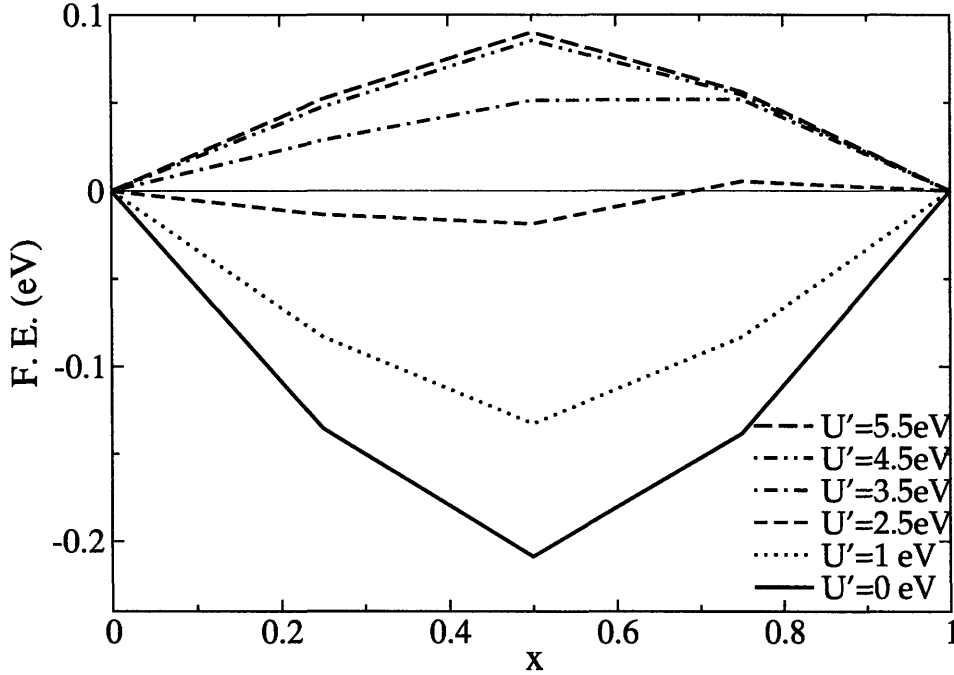
Figure 2-11: Formation energy of Li_xFePO_4 at different x and U' values. The points at $x = 0.5$ correspond to structure 0.5a.

parameter, $U' = U - J$. We evaluate all results as functions of U', spanning the range from 0 to 5.5eV *. When calculating formation energies for a given U', we assume U' to be composition independent.

2.6.3 Results and discussions

In Fig. 2-11 formation energies at different U' are shown as a function of Li concentration x. At each concentration ΔE increases with U' and becomes positive at intermediate U'($\approx$2.5-3.5eV). The formation energies saturate to a nearly constant value around U' $\approx$3.5-4.5eV. The effect of the E_U term is to drive the Fe-$3d$ orbital occupation numbers to integer (0 or 1) values. As a result, the Fe ions tend to have integral occupancy even in the partially lithiated structures, and charge ordering occurs: we see distinct Fe^{3+} and Fe^{2+} in DFT+U instead of the uniform $Fe^{(3-x)+}$ seen in LDA/GGA. For low U' values ($U' \lesssim$ 1eV) the four Fe ions in the unit cell have sim-

---

*Even though we have calculated the parameter self-consistently, we want to show the trend with respect to it since the phase stability of Li_xFePO_4 is a serious qualitative failure when $U = 0$. We will argue that the correct physics is obtained within a reasonable range of U'.

54

ilar $3d$ electron occupancy and Fe-O bond lengths for all the intermediate structures. Therefore, little charge ordering occurs in this limit, even though the Fe ions occupy symmetrically distinct positions. We will call these Fe cations (3-x)+ like. They are stable with respect to small perturbations in initial charge distribution. In the high limit of $U'(\gtrsim 3.5$ or 4.5 eV) there are 2 types of Fe ions, one very similar to those in $FePO_4$ (which we call Fe^{3+} like) and the other similar to those in $LiFePO_4$ (called Fe^{2+} like). The designation 3(2)+ is only meaningful in that the Fe ions are similar to those in $FePO_4(LiFePO_4)$. The Fe-O hybridization gives them less than nominal charge. For x=0.25(0.75) calculations imposing the symmetry of the structure on the charge density leads to two 3(2)+ like and two 2.5+ like Fe ions. Only when symmetry is broken does a lower energy state with three 3(2)+ like and one 2(3)+ like ions form. In these structures the charge density has lower symmetry than what would be expected from the ionic positions and, hence, charge ordering occurs. As the analysis for all five structures is similar we choose x=0.5a as a typical intermediate structure for further discussion.

In Fig. 2-12 ΔE(x=0.5a) is shown as a function of U'. We investigated AFM spin configurations in x=0, 0.5a and 1 and found them to give only slightly lower total energies. The AFM ΔE (dotted line) is almost equivalent to the FM one with charge ordering (solid line). We also studied a 'restricted' FM system at x=0.5a where all four Fe ions have the same initial magnetization, ending up 2.5+ like. Charge ordering is absent in this metastable state, which has higher total energy than the charge-ordered ground state. From Fig. 2-12 we can compare ΔE with and without charge ordering. Note that the curve with charge ordering levels off for $U \gtrsim 4.5$eV, which is explained below.

To study quantitatively the change in formation energies and electron distribution as U' is increased, the contributions to ΔE are separated into the LDA energy, ΔE_{LDA}, and the correction term, ΔE_U, with definitions analogous to ΔE in Eq. 2.4. The occupancy of the most occupied of the five minority-spin $3d$-orbitals of iron is displayed in the lower part of Fig. 2-13. This orbital is most relevant because its occupation makes the difference between Fe^{3+} and Fe^{2+}. When charge ordering is
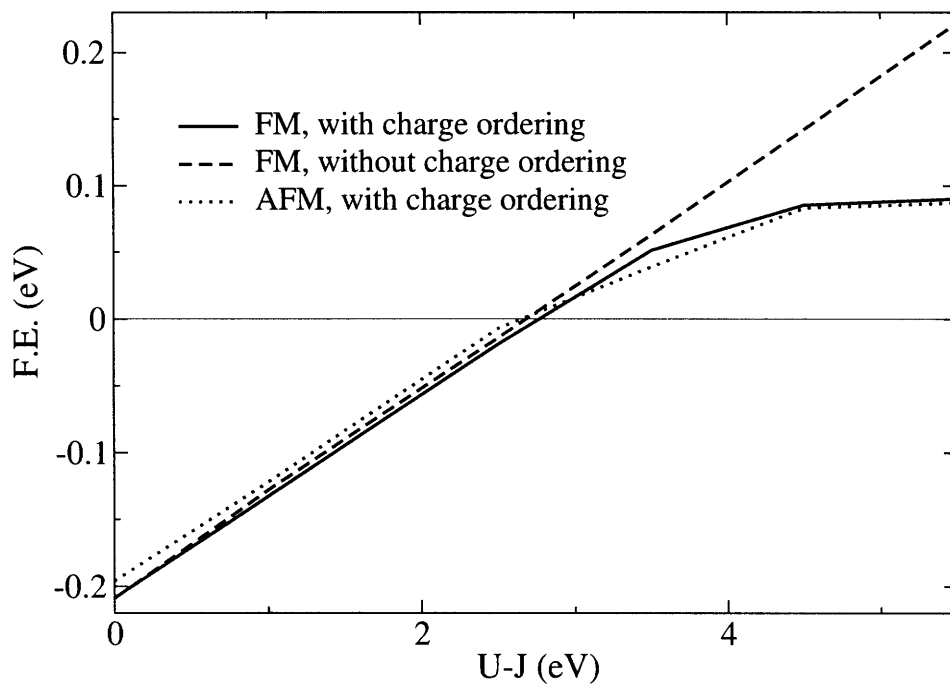
Figure 2-12: Formation energy of structure 0.5a versus U'. The solid line correspond to ferromagnetic states with charge ordering, the dashed line to ferromagnetic states without charge ordering and the dotted line to anti-ferromagnetic states with charge ordering.
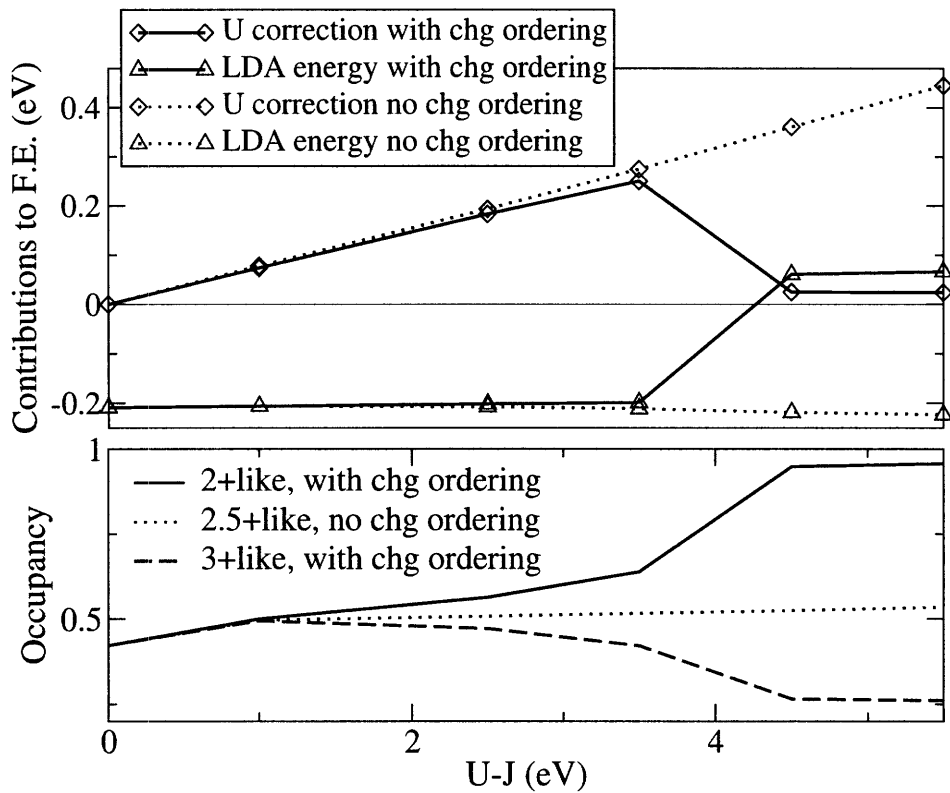
Figure 2-13: Upper part: LDA (triangle) and U correction term (diamond) contributions to ΔE(x=0.5a) *vs.* U'. Solid/dotted lines indicate presence/absence of charge ordering. Lower part: occupancy of the most occupied minority-spin orbital versus U', for Fe 2+ (solid line) and 3+ (dashed line) in the charge-ordered state and for 2.5+ (dotted line) in the state without charge ordering.

absent, the occupation number does not change much with U' and stays near 0.5, as expected of a 2.5+ like Fe cation. In contrast, the curves in the charge-ordered state separate beyond $U' \approx 1eV$, with half of the ions becoming 2+ like and the other half 3+ like. These occupancies can help to explain ΔE in the upper part of the diagram. When charge ordering is absent (dotted lines) the four Fe cations in the x=0.5a structure are equally affected by U' in terms of $3d$ occupation, as they are in x=0 and 1, and the changes in $\text{Tr}(\hat{n}(1 - \hat{n}))$ in $Li_{0.5a}FePO_4$ are canceled by the weighted average of those in x=0 and 1 structures. As a result, the correction term ΔE_U is almost proportional to U', explaining its linear behavior in Fig. 2-13. When the symmetry is sufficiently broken, ΔE_U will make Fe-$3d$ charge density order so as to create, as much as possible, orbitals with integer occupation. This comes at a cost to ΔE_{LDA}, which changes from large negative values at $U' = 0$ to positive values. We see two possible reasons why ΔE_{LDA} increases when charge ordering occurs. Localization of the minority spin electrons into half of the Fe sites as Fe^{2+} obviously leads to an increase in kinetic energy. Additionally, since Fe^{2+} and Fe^{3+} have different Fe-O bond lengths, their coexistence in one structure comes with a penalty in elastic energy. As the increase, relative to the LDA values, in ΔE_U is much smaller than in ΔE_{LDA}, the latter can be identified as the cause of phase separation.

The ground state electronic structure is also affected. The 0.5a compound is a Mott insulator when charge ordering occurs in LDA+U, while it is metallic in LDA without charge ordering.

We have further confirmed that the positive formation energies obtained in Fig. 2-12 are not an artifact of using a single unit cell by calculating the energy of four other structures (x=0.25 or 0.75) with a doubled unit cell. We found all these formation energies to be within ±10 meV of the results shown in Fig. 2-12.

2.6.4 Conclusions

In summary, we find that both LDA and GGA qualitatively fail to reproduce the experimentally observed phase stability and mixing energetics in the Li_xFePO_4 system. For $U - J > 3.5eV$, LDA+U calculations give positive ΔE, in agreement with

experiments.

2.7 Conclusions

We have confirmed that the LDA+U method leads to better description of the electronic properties (band gap, magnetization, localization of d-electrons) in lithium metal compounds. This in turn gives more accurate energetics, improving prediction of lattice constants, redox potential and formation energies. Indeed, the method has already helped in research of potential transition metal compounds as lithium battery electrode [46, 47]. In the next chapter we continue to investigate the effects of electron correlations on finite temperature thermodynamics.

Chapter 3

Electronic Entropy and Phase Diagram of Li_xFePO_4

3.1 Introduction

In the previous chapter, we have focused on the ground state or zero temperature (zero-T) properties of transition metal compounds. Electron correlation effects, treated with the GGA+U method, have been shown to play a crucial role in accurate characterization of the materials. The GGA+U method is able to capture the main (static) correlation effects among localized d-electrons and removes the spurious self interaction present in LDA or GGA. The method accurately predicts total energy properties (voltage, formation energy) and correctly determines the basic trends in the electronic structure (band gap). An increasing number of authors have also supported the GGA+U approach as a positive step towards description of the ground states of strongly correlated electron systems.

Finite temperature thermodynamical properties are equally important for scientific interest and for practical applications of transition metal compounds. Good description of the ground states does not automatically guarantee accurate thermodynamics, unless all the relevant (or at least the most significant) degrees of freedom and excitations are accounted for properly. We show in this chapter that qualitatively different thermodynamical properties can occur as a result of correlated elec-

trons [221]. To the best of our knowledge, this is the first time a phase stabilized by configurational electronic entropy has been identified and studied.

Entropy is central to the understanding of phase transitions and reactions of materials. The statistical mechanical origin of entropy is Boltzmann's postulate,

$$S = k_B \ln \Omega, \tag{3.1}$$

where Ω is the number of microscopic configurations, and k_B is Boltzmann's constant. In other words, from a microscopic perspective, the entropy is a measure of the number of microscopic states that are accessible by the macroscopic system. The microscopic configurations that contribute to Ω have different origins, e.g. configurational disorder (as in metallic alloys) or vibrational states (as in liquids). Using a Monte Carlo simulation with explicit ionic and electronic degrees of freedom we demonstrate that in Li_xFePO_4 a high temperature solid solution is stabilized by electronic rather than ionic configurational entropy.

The binary phase diagram of Li_xFePO_4 is very distinctive from many other Li-intercalated transition metal oxides in that the material phase separates into end members $FePO_4$ and $LiFePO_4$ at low temperature and exhibits a eutectoid transition at higher temperature. It was shown in Chapter 2 that the ordering of electrons/holes on the Fe sub-lattice , localized due to electron correlation effects, co-exists with the Li/vacancy ordering on the lithium sub-lattice. We attempt to explain the unexpected phase behavior of Li_xFePO_4 by treating the configurational electron degrees of freedom explicitly. The competition of attractive and repulsive interactions, of charge carriers between and within sub-lattices respectively, is found the microscopic mechanism of the unusual phase separation. Through Monte Carlo simulations using the obtained interaction parameters, we also demonstrate that configurational electronic entropy plays a crucial role in driving the eutectoid transition substantially more than Li/vacacy ordering. The calculated phase diagram agrees very well with experiment. A potential meta-stable intermediate phase is predicted by our simulation.

The consequences of our results on other mixed-valence oxides are discussed.

3.2 Electronic entropy

A pure DFT approach is applicable to zero-T. To study finite-T phase diagram, one has to be particularly careful and identify all the excitations and degrees of freedom involved in creating entropy. Typically in alloy theory the focus is on the configurational disorder (substitution of different elements or vacancies (V)) while the electronic degrees of freedom are, in the spirit of the adiabatic approximation, integrated out [49, 28]. * For example, many phase diagrams can be satisfactorily reproduced by considering the configurational entropy of two elements (Most examples are in alloys [14, 160, 159, 169, 29, 28]. For review, see [49] and references there in) or element and vacancy (see, for example, [189, 48, 33], and for review [194]).

The electronic degrees of freedom is another entropy-generating mechanism. Electronic entropy comes from different electronic states accessible at a given temperature. It is conventionally thought of as a small quantitative correction.

3.2.1 Band (kinetic) entropy

In some cases electronic entropy becomes important and can be calculated from the band structure: [203, 126]:

$$S_e^{\text{band}} = -k_B \int n(f \ln f + (1 - f) \ln(1 - f)) dE, \tag{3.2}$$

where n and f are the density of states and Fermi distribution function, respectively. Only electrons within $\sim k_B T$ to the Fermi level participate in the excitations, so S_e^{band} is usually small. This is a reason why in finite-T thermodynamics calculation of materials electronic entropy is ignored more often than not.

In the insulating $LiFePO_4$ the band gap is larger than 3 eV (see previous chapter). Up to 1000 K, the number of electrons that can be excited over this band gap is still very limited. Therefore we expect the band entropy of electrons to be negligible.

---

*phonon contributions may give quantitative corrections [197, 187, 14, 202], but they are relatively composition insensitive and will not be discussed here

3.2.2 Configurational entropy

A different type of electronic entropy could arise if electrons/holes (e/h) are localized and contribute to the total entropy in the same fashion as the ordering of atoms. One would expect such configurational electronic entropy to be particularly important in mixed-valence transition metal oxides. Many technologically important materials, such as doped manganites, high-T superconductors, Na- and Li-metal oxides, and mixed conductors, fall in this category. Very little is known about the contribution of localized e/h to the finite-T phase stability of such materials, though previous evidence exist in doped YBCO superconductors [165, 180] and perovskites [97] that a configurational electronic entropy term (assuming random e/h distribution) is needed to explain the entropy of oxidation/reduction, given by:

$$S_{\mathrm{e}}^{\mathrm{loc,rand}} = -k_B \left[x \ln x + (1 - x) \ln(1 - x) \right], \tag{3.3}$$

where x is the concentration of localized electrons or holes. Note that $S_{\mathrm{e}}^{\mathrm{loc,rand}}$ can potentially be as significant as the configurational entropy of ions. Nevertheless Eq. 3.3 only relies on the composition x and does not incorporate explicit configuration dependence, which is essential for any non-trivial treatment of finite-T thermodynamics associated with configurational disorder. Presently there exists no clear demonstration that electronic entropy can qualitatively modify finite-T phase diagram.

In this chapter we investigate the effects of configuration dependent electronic entropy. We go beyond a random model such as Eq. 3.3 and sample electron configurations explicitly. We focus on the $Li_x FePO_4$ system. While its high intrinsic Li^+ mobility makes it of interest as the next-generation cathode for rechargeable Li batteries [132], it is also crucial to ensure good phase equilibration, even at room temperature (RT). So $Li_x FePO_4$ is a particularly good system to benchmark theory against. We find that excellent agreement with the experimental phase diagram can only be achieved by taking into account configurational electronic entropy, and qualitative discrepancies exist if the electron degree of freedom is ignored.
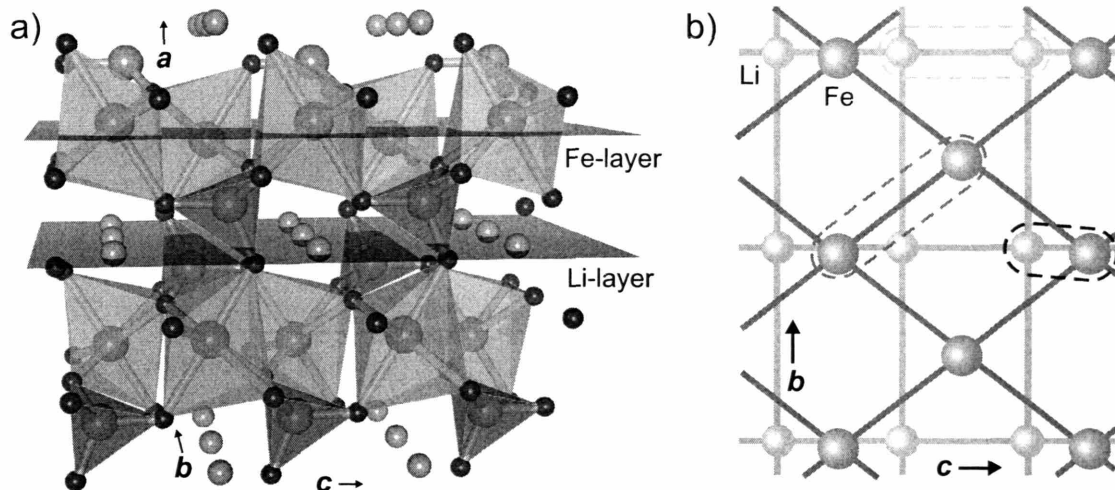
Figure 3-1: The LiFePO$_4$ structure shown with: **a**) PO$_4$ and FeO$_6$ polyhedra as well as Li atoms **b**) adjacent layers on Li and Fe sub-lattices, projected along axis a, with nearest-neighbor (NN) inter- and intra-lattice pairs highlighted.

3.3 Experimental phase diagram of Li$_x$FePO$_4$

In Li$_x$FePO$_4$ the electrons accompanying insertion of Li$^+$ ions are mainly localized on the Fe ion sites. Charge separation due to correlation effects stabilizes distinct Fe^{2+}/Fe^{3+} (or electrons/holes) states, i.e. formation of small polarons [220, 112]. The crystal structure of LiFePO$_4$, already presented in the previous chapter, is reproduced in Fig. 3-1a for convenience. The picture of coexisting Li/vacancy and electron/hole sub-lattices is emphasized in Fig. 3-1b. The Li$^+$ ions and vacant sites sit on an orthorhombic lattice, of which one layer is shown in Fig. 3-1b (large green points). On each side of this Li layer is a plane of Fe sites (only one plane shown in small brown points).

The Li$_x$FePO$_4$ phase diagram is critical for understanding and improving the material's performance. As described in Section 2.6, room temperature (RT) electrochemical Li removal exhibits a miscibility gap between triphylite (T) LiFePO$_4$ and delithiated heterosite (H) FePO$_4$, and no intermediate compound Li$_x$FePO$_4$ exists between T and H [132]. Both phases having a very limited amount of solubility (vacancies + holes (Fe^{3+}) in T and Li+ ions + electrons (Fe^{2+}) in H) [211]. Recently the high-T phase diagram was investigated by Delacourt et al [51] and by Dodd et al [58].
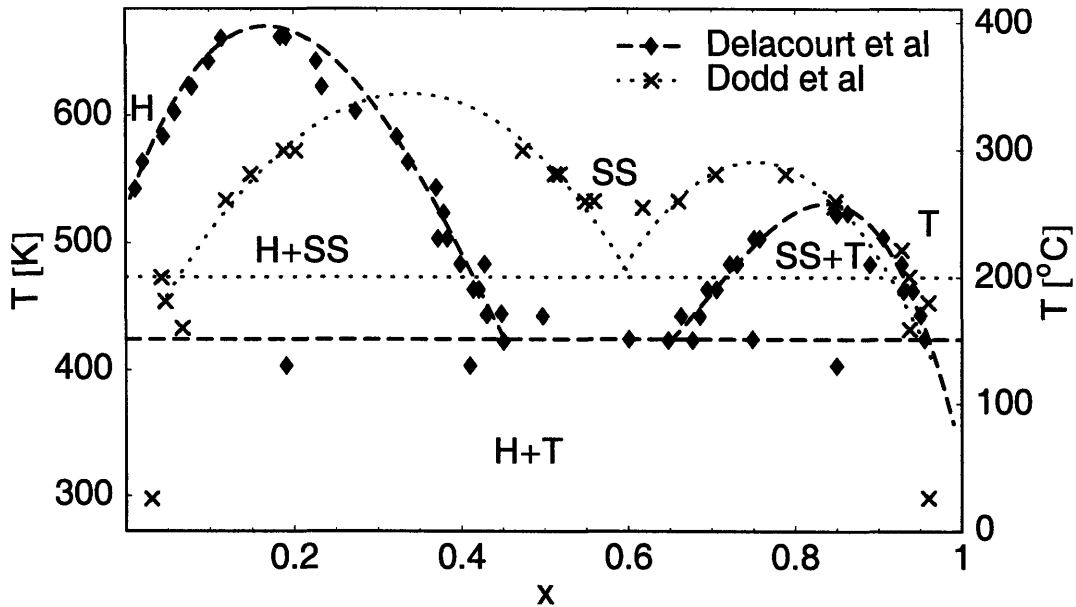
Figure 3-2: Experimental phase diagram of Li_xFePO_4. The boundary data points are taken from Delacourt et al [51] and from Dodd et al [58]. The dashed lines, also taken from these references, are just fitted to the data points, not real data.

For convenience of comparison with our theoretical work, the experimental phase boundaries are presented here in Fig. 3-2. The low-T immiscibility was confirmed in both works, but an unusual eutectoid behavior was also reported: above the transition temperature of about 150°C [51] or 200°C [58] a disordered phase (D) emerges around $x \approx 0.45 - 0.65$; above congruent T of about 300 − 400°C SS dominates all compositions (Fig. 3-2).

This phase diagram is quite unexpected from a theoretical point of view. First, why does the system phase separate at all at low-T? In a simplified picture of a generic oxide Li_xMO_n, the Li^+ ions repel each other due to electrostatics so that ordered intermediate compounds are energetically favorable over phase separation, i.e. segregation of Li^+ (vacancies) into Li-rich (deficient) regions. This is indeed the case in many other materials, in which mobile ions and vacancies coexist, e.g. Li_xCoO_2, Li_xNiO_2, Na_xCoO_2 [189, 179, 48]. To answer this question one has to go beyond the above intuitive argument. Secondly, what is the origin of the complex high-T behavior? Transitions from a two-phase coexistence state to a solid solution are typically driven by the configurational entropy of the ions in the SS, with a maximum

transition T near equiatomic A/B composition. The experimentally established phase diagram, shown in Fig. 3-2, is unlikely to come from such ionic configurational entropy unless the effective Li-V interactions are unusually strongly composition dependent. We notice that there has been considerable experimental effort [52, 210, 34, 171] following the phase diagram works, but these questions remain to be answered.

As discussed earlier, the electronic degrees of freedom are usually integrated out for metallic system. However in a insulating material with localized electrons and holes such as Li_xFePO_4, such omission cannot be quite justified. We demonstrate that the topology arises from electron degrees of freedom which stabilizes the SS near $x \approx 0.5$.

3.4 Energy calculations

There are at least three levels of theoretical approach to our problem. The works presented in the previous chapter show that the GGA+U method provides a reliable description of the ground states of Li_xFePO_4. In this chapter, we will couple it with explicit treatment of electron/hole configurations.

3.4.1 Computational schemes

Li_xFePO_4 has been previously studied by DFT with the local density or generalized gradient approximation (LDA/GGA) [177, 166, 124, 206]. Due to its lack of treatment of electron correlation in transition metal oxides, LDA or GGA fails to correctly reproduce some key properties of $LiFePO_4$. It predicts very small or even vanishing band gap [177, 166, 206], contradicting experimentally suggested 3.8–4.0 eV [219, 16]. Even worse, intermediate Li_xFePO_4 is found stable at $T = 0$ [220], a qualitative error. Therefore LDA/GGA cannot be used to study the phase diagram. Note that compared to Fe^{3+} in $FePO_4$ (high spin $t_{2g}^3 e_g^2$) and Fe^{2+} in $LiFePO_4$ (high spin $t_{2g}^4 e_g^2$), the electrons donated by Li^+ ions in Li_xFePO_4 are smeared out to all iron sites, resulting in an average valence of $2 + x$ (high spin $t_{2g}^{3+x} e_g^2$) [220].

The second level of approach is the GGA+U method [13], which treats on-site

electron correlation of localized d or f electrons explicitly and has much success in electronic structure calculations of transition metal oxides [10]. Its application in Li_xFePO_4 has been very fruitful [220, 217, 219, 216, 103]. We have shown that electronic structure properties, such as band gap, can be accurately predicted with this method [219] (and verified by experiment [219, 16]). Good agreement with experiment has also been found in total energy related properties, e.g. the lithium intercalation voltage of 3.5 V [217, 216], and the instability of Li_xFePO_4 with respect to the end members [220]. In this approach, the Fe ions charge-separate into 2+ and 3+ valence states, and small polarons can be formed [220, 112]. Following conventions in the literature the Fe 2+ and 3+ states will later be referred to as (localized) electron and hole, respectively. Recently the migration of small polarons in Li_xFePO_4 and their association with lithium ions and vacancies have been studied using GGA+U [112].

The successes of GGA+U for Li_xFePO_4 ground-state properties, including confirmation of zero-T phase separation, assumes that electrons follow ions and always adopt their lowest energy configuration. To deal with finite temperature properly we have to consider excitations. In this chapter we propose the third level of approach: GGA+U plus explicit electron, as well as lithium, configurational degrees of freedom. The distribution of e^-/holes, as well as Li^+/vacancies, are expected to significantly affect the energetics of the system. Both the Fe and the Li sub-lattices may contribute composition-dependent terms as large as Eq. 3.3 to the total configurational entropy. Therefore the free energy landscape, hence the phase diagram, would be profoundly influenced by both degrees of freedom.

3.4.2 Details of the calculation method

GGA+U calculations are performed using self-consistent parameter $U = 4.3$ eV [217] and the projector-augmented wave (PAW) method [23] as implemented in VASP [93]. An energy cut-off of 500 eV and appropriate k-point mesh were chosen so that the total ground state energy is converged to within 3meV per formula unit. All atoms and cell parameters are fully relaxed, with no symmetry conditions imposed. Our GGA+U calculations are initiated with (and relaxed to) ferromagnetic high-

spin Fe ions. Although the ground-state magnetic structure of $(Li)FePO_4$ is anti-ferromagnetic [161, 156], previous studies showed very small magnetic couplings [220, 44], and we expect the entropy associated with the magnetic degree of freedom to be ($\approx k_B[x \ln 5 + (1 - x) \ln 6]$), linear in x at RT, therefore negligible in phase diagram calculations.

It is straightforward to obtain a desired Li configuration in the calculation. To deal with an electron configuration, we begin a preparatory step first with potential energy corrections similar to those used in standard GGA+U:

$$\Delta V = \sum_i V(n_i) = \sum_i (n_i - n_i^0)^2, \qquad (3.4)$$

where n_i is the occupancy variable at Fe site i and n_i^0 represents the occupation number of desired Fe valence 3+ or 2+. * This first step makes sure that the electrons are localized appropriately. The resulting charge density and atomic coordinates become initialization for the next standard GGA+U calculation. This procedure works in all but very few configurations, where the electrons are relaxed to a charge configuration different from the input, or to a state where they is not completely localized (some 3.5+ like iron sites). The energies of these failed cases are always higher than the normal results corresponding to the same Li/vacancy ordering (but different initial charge ordering) and therefore discarded.

3.4.3 Results

We included 245 distinct Li_xFePO_4 ($0 \leq x \leq 1$) configurations of Li/V and e/h with super-cells of up to 8 primitive cells (32 formula units) in our calculations. Figure 3-3 shows the calculated formation energies (Eq. 2.4), which characterize the structures' stability relative to the end members, of as a function of Li composition x (black dots). Unless otherwise stated, all extensive quantities (energy, entropy, etc) are presented per formula unit. Several features are obvious from the plot. All 243 intermediate

---

*Due to metal-oxygen rehybridization, the occupancy of the "ideal" 3+/2+ states $n_i^0 = 5.5$ or 6.1 is fractional and different from the nominal electron occupation number 5 or 6.
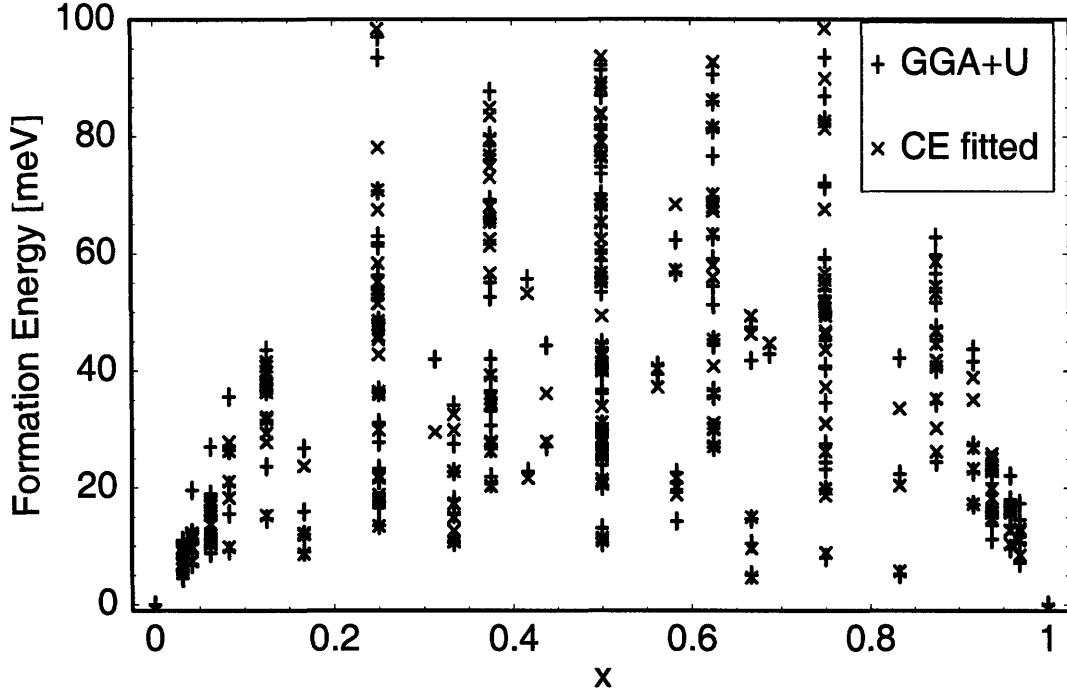
Figure 3-3: Li$_x$FePO$_4$ formation energy per formula unit. The calculated (black cross) values are compared to fitted (red crosses) in a cluster expansion model. Only $\Delta E < 100$ meV shown.

$\Delta E(x)$ ($0 < x < 1$) values are positive, meaning that all corresponding structure are unstable with respect to FePO$_4$/LiFePO$_4$, in agreement with the low-T phase separation phenomenon. Formation energies at certain concentrations (e.g. $x = \frac{1}{2}, \frac{2}{3}$) are considerably lower than at others, indicating the composition dependence of energy. Significant scattering of data points at the same concentration points to the configuration dependence of energy. In the dilute solution limit, for example $x \approx 0$, the concentration of defects (Li$^+$ + e$^-$) is low and the scattering is mainly due to different separation distance between the Li$^+$ and e$^-$: the closer the pair, the lower the energy in general. Slight energy asymmetry between concentrations x and $1 - x$ can also be observed: the lowest ΔE close to $x = 1$, or the defect energy in LiFePO$_4$, is larger than that close to $x = 0$, or the defect energy in FePO$_4$. We therefore expect the solubility limit to be lower in LiFePO$_4$ than in FePO$_4$.

Fig. 3-4 shows the lowest energy configurations in Fig. 3-3 at $x = 1/2$, $3/4$ and $2/3$, respectively. They are similar in that each is obtained from the fully lithiated

70
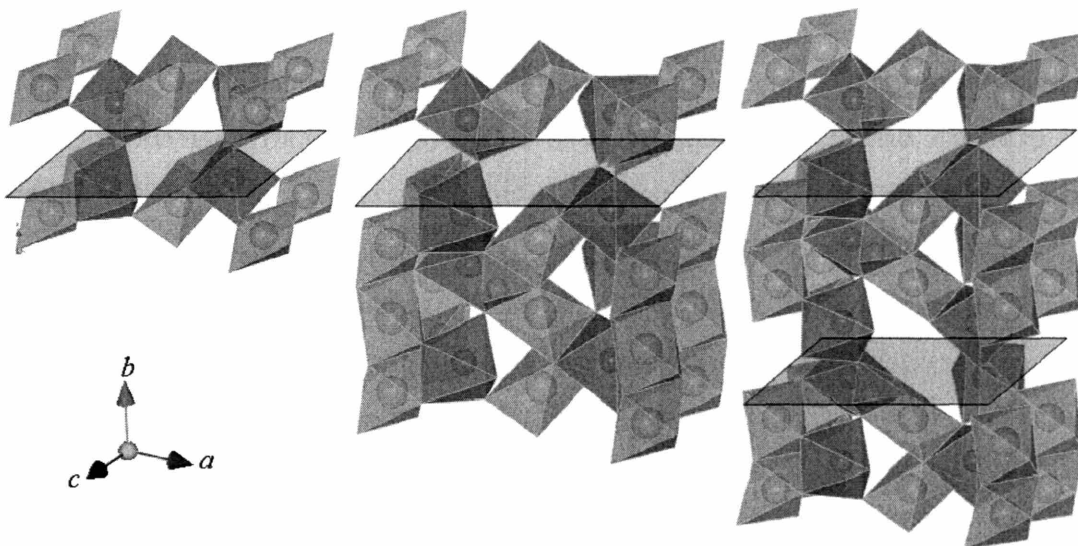
Figure 3-4: Low energy configurations in Fig. 3-3 at $x = 1/2$, $3/4$ and $2/3$, respectively. Green polyhedra= LiO_6, brown= FeO_6 and purple tetrahedra= PO_4. The locations of vacancy sites are indicated with green planes.

structure by removing one layer of lithium atoms (the plane shown in Fig. 3-4, not the ones in Fig. 3-1b) from a plane perpendicular to the lithium diffusion channel (axis b), out of every 2, 4 and 3 Li layers, respectively. The relatively stable configurations in figure 3-4 can be partly explained by the fact that many of the nearest neighbor Li-Li and e^--e^- repulsive interactions can be avoided by taking out a lithium layer, as discussed in the next section.

3.5 Lattice energy model

The first principles calculations are quite expensive. A single structure show in Fig. 3-3 takes up to hundreds of CPU hours. To sample a large amount of Li/V + e/h configurations we simplify the energy evaluation to a lattice model of interactions.

3.5.1 Cluster Expansion

In systems with configurational degrees of freedom, it is convenient to describe properties that depend on configuration, such as energy, with a cluster expansion (CE)

[159]. To model the energy dependence on both the Li/V and the e/h configurations, we have applied the coupled cluster expansion method [178], which is a Hamiltonian that explicitly describes the dependence of the energy on the arrangement of Li^+/V and Fe^{2+}/Fe^{3+} (e/h), i.e. both ionic and electronic degrees of freedom. Representing with $\lambda_i = \pm 1$ occupation of site i by a Li^+ or vacancy and with $\epsilon_a = \pm 1$ the presence of Fe^{2+} (electron) or Fe^{3+} (hole) on site a, the energy can be expanded without loss of generality in polynomials of these occupation variables [159, 178]:

$$E[\vec{\lambda}, \vec{\epsilon}] = J_\emptyset + J_i \lambda_i + J_{ij} \lambda_i \lambda_j + J_{ia} \lambda_i \epsilon_a + J_{ab} \epsilon_a \epsilon_b + \ldots \qquad (3.5)$$

The expansion coefficients J are called effective cluster interactions (ECI), essentially coupling constants in a generalized Ising model. Charge neutrality requires that the number of Li ions be always equal to the number of localized electrons. Therefore in Eq. 3.5 we only need to include the point term for lithium $J_i \lambda_i$, and the term for electron $J_a \epsilon_i$ is linearly dependent and does not appear. In its untruncated form, Eq. 3.5 is exact and includes all multi-body terms within one sub-lattice (Li/V or e/h) and between sub-lattices though some truncation takes place in practice.

3.5.2 Cluster Optimization

The calculated formation energies in Fig. 3-3 are used to train our cluster expansion model. The aim is to construct a well-converged set of clusters that is capable of accurately reproduce configuration dependence of energy. General observations in the application of CE to solid state materials has led to some general rules based on the underlying physics, including

1. priority on low-order clusters (e.g. pair favorable over triplet),

2. priority on short range clusters, on the assumption of decaying interaction magnitude over distance [222],

3. priority on low energy structures, since in many cases a cluster expansion cannot accurately reproduce the energetics on all energy scale and the physically

meaningful ground states need more accuracy [67],

4. inclusion of all sub-clusters of every cluster to prevent attributing a cluster's energetic contribution to its parent clusters [186, 212, 170].

Cross validation (CV) score provides a good mathematical estimation of the predictive power of the cluster expansion. It is specifically designed to estimate the error made in predicting the energy for structures not included in the least-squares fitting. The so-called "leave-one-out" formula gives the CV as [186]:

$$CV^2 = \langle (E_i - \hat{E}_i^P)^2 \rangle, \tag{3.6}$$

where $\hat{E}_i^P$ is the predicted value of structure i obtained from a least-squares fit to all other structures. By optimizing CV with cluster construction, we can steadily test and improve the cluster expansion model, in a mathematical fashion. However, when number of the training structures is limited, as is often the case in first-principles calculations, optimization solely by minimizing CV score can lead to over-fitting to these structures while neglecting the nature of the interactions.

In this chapter, we choose the clusters "manually", aiming to reduce the CV score while observing the above rules. While constructing the model all pair clusters within 11 Å, triplet clusters within 8 Å and quadruplet clusters within 7.2 Å were tested (the "size" of a cluster is determined by the maximum distance between any two sites in it).

More detailed discussions on cluster expansion and optimization of the model can be found in the next chapter when we apply it to the protein design problem.

3.5.3 Optimized ECIs

Our optimized cluster expansion model has a cross validation score of less than 6 meV per formula unit and consists of 29 distinct ECIs: the constant and the point terms which respectively reference the total energy and the Li chemical potential and which do not affect the phase diagram; 7 small triplet terms, which mainly represent slight

73
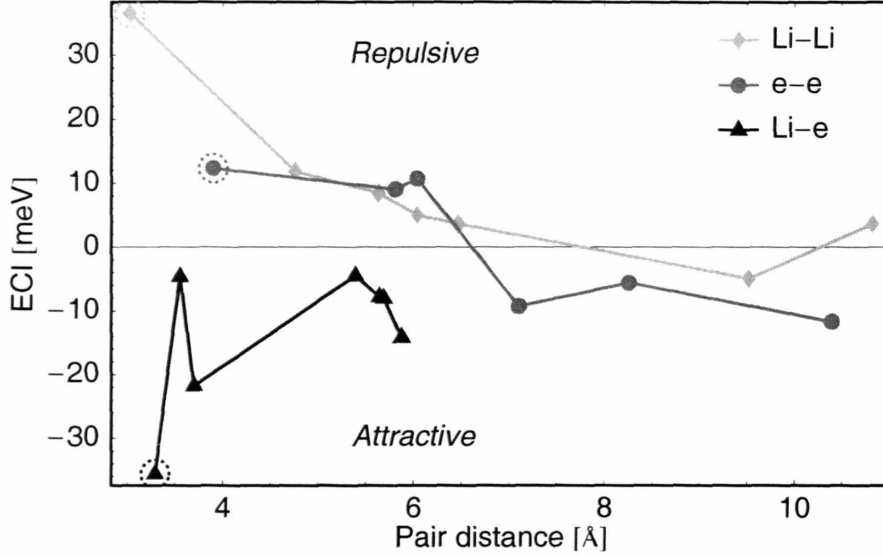
Figure 3-5: Pair ECI versus the sites' distance (measured from their ideal coordinates in LiFePO$_4$). The circled points correspond to NN Li-Li, e-e and Li-e pairs in Fig. 3-1b, respectively.

asymmetry between FePO$_4$ and LiFePO$_4$; and most significantly 20 pair interactions shown in Fig. 3-5. Note that these are *effective* interactions including the effects of many physical factors: electrostatics, screening, relaxation, covalency, etc.

Without calculating finite-T thermodynamics, much information can already be extracted from Fig. 3-5. The Li-Li ECI (diamond) is largest for nearest-neighbor (NN) Li$^+$ ions, which repel each other strongly for electrostatic reasons. As the pairs are separated further, the repulsion is screened considerably. The small negative J_{Li-Li} at large distance indicates some mediation of the effective interactions by lattice distortions. Roughly the same trend is observed for J_{e-e} (circles), being repulsive at short range and attractive at long range. On the contrary, the Li$^+$-e$^-$ inter-lattice ECIs are strong short-range attractions that generally become weaker at longer distance. They quantitatively explain the energy scattering near $x=0$ (1) in Fig. 3-3 and the fact that structures with closely bound Li-e (vacancy-hole) pair defects in the dilute limit are relatively stable. The trend in the three curves is not monotonic, since the ECIs contain complex lattice factors beyond isotropic electrostatics. The low-T phase separation can be explained by considering the dominating short-range

terms. The Li^+ ions repel each other and so do electrons, while Li-e attractions compete to bind them together: if Li^+ ions are together then the e^- can bind to more of them. This is a delicate balance of competing terms. The Li-e attractions prevail partly because of the host's geometry: the multiplicity of the NN Li-e ECI, the strongest attraction, is two per formula unit, while that of NN Li-Li ECI, the strongest repulsion, is one (see Fig. 3-1). We therefore conclude that *phase separation in $Li_x FePO_4$ is mainly driven by Li^+-e^- attractions in competition with Li^+-Li^+ and e^--e^- repulsions.* This is fundamentally different from a system where the electronic mixed valence becomes delocalized, as in metallic $Li_x CoO_2$, thereby making the Li-e coupling independent of the Li/V distribution.

3.6 Phase diagram

3.6.1 Free energy integration

Following the prescription in [197, 186, 185, 184], we make use of the following properties of the grand potential $\Phi \equiv E - TS - \mu x$:

$$
\begin{aligned}
\mathrm{d}(\beta\Phi) &= (E - \mu x)\mathrm{d}\beta - \beta x \mathrm{d}\mu \\
\beta\Phi(\beta,\mu) &= \beta_0\Phi(\beta_0,\mu_0) + \int_{\beta_0,\mu_0}^{\beta,\mu} (E - \mu x, -\beta x)\mathrm{d}(\beta,\mu),
\end{aligned}
\tag{3.7}
$$

to calculate potential Φ through integration over the (β,μ) parameter space that stabilizes a single phase P. In Eq. 3.7 μ represent the chemical potential of a Li^+ + e^- pair in $Li_x FePO_4$, and $\beta = 1/T$. The following equation helps trace the phase boundary between two phases at equilibrium.

$$
\frac{d\mu}{d\beta} = \frac{\Delta E}{\beta \Delta x} - \frac{\mu}{\beta},
\tag{3.8}
$$

where the Δ sign represents the finite difference in a first-order phase transition.

The starting equilibrium point (β_0, μ_0) can be chosen such that the potential Φ approaches analytical expressions. A convenient starting point is when $T \to 0$ and

one of the ground states g ($x = 0, 1$) is stabilized, so Φ can be approximated by the low temperature expansion (LTE) with a single pair of flippings:

$$\Phi(T \to 0, \mu) = E_g - \mu x_g - k_B T \sum_e \exp\left[-\beta(\Delta E_e - \mu \Delta n_e)\right], \qquad (3.9)$$

where E_g is the formation energy per FU of ground state g $\text{Li}_{x_g}\text{FePO}_4$, ΔE_e and $\Delta n_e = \pm 1$ are the changes in energy and Nx associated with flipping a given Li site and electron site e from g, and the last term is summed over "short-range" electron sites. In the low T region (up to 300 K) only 2 terms with degenerate lowest energies were retained in our calculation. When deriving Eq. 3.9 one makes the assumption that all flipped pairs of sites are well separated and the pair is bound within a distance r_{bound} small compared to separation distance D_{pair}. We found that it suffices to limit the summation r_{bound} to the maximum calculated value 15 Å.

The high temperature expansion (HTE) of the grand potential Φ provides a mean-field like expression as another reference point:

$$\begin{aligned}
\Phi(T \sim \infty, \mu) &= E(x) - TS(x) - \mu x \\
&= m_\alpha J_\alpha (2x - 1)^{n_\alpha} + 2T\left[x \log x + (1 - x) \log(1 - x)\right] - \mu x \quad (3.10)
\end{aligned}$$

where the summation over all calculated clusters α is implied, and m_α, J_α and n_α represent the multiplicity (number of symmetrically equivalent clusters), ECI and order (number of sites) of cluster α, respectively. The $x(T, \mu)$ parameter is obtained by minimizing Eq. 3.10 with respect to x: $\partial \Phi(T, \mu, x)/\partial x = 0$. Note that the factor 2 of entropy in Eq. 3.10 includes contributions from both Li/V and e/h disorder.

3.6.2 Monte Carlo simulation

Monte Carlo (MC) simulations were carried out on a lattice model with the calculated ECIs and a $6 \times 12 \times 12$ super-cell. The solid solution phase proved hard to obtain convergence on. Data points in the SS typically take more than 10^6 MC passes. For the FePO_4 and LiFePO_4 phases, no more than 10^5 passes provide sufficient accuracy.
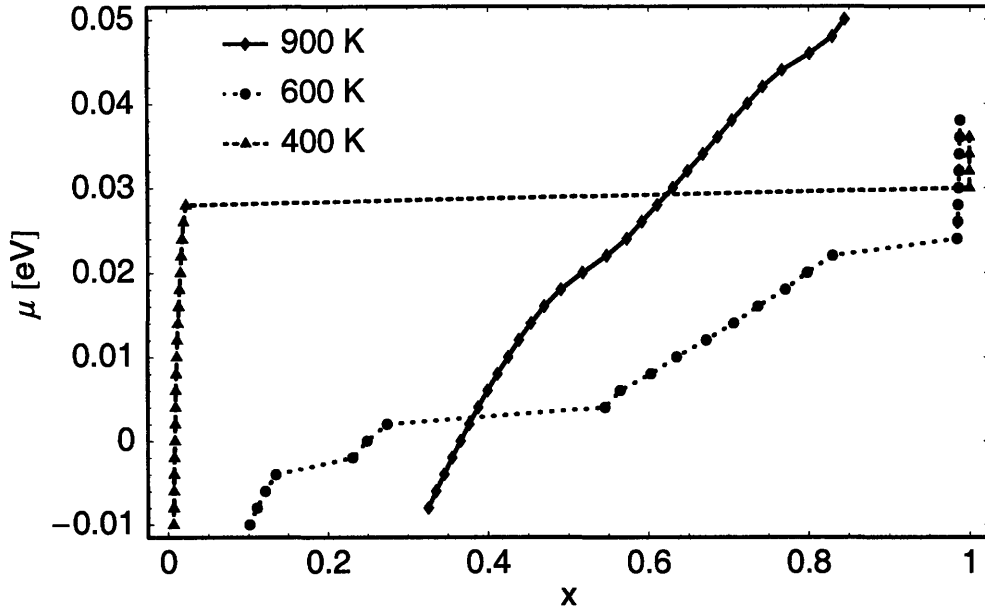
Figure 3-6: Chemical potential (μ, arbitrary reference) scan in the MC simulations at $T = 300$, 600 and 900 K by increasing μ.

To expedite convergence we attempted with 20% chance a canonical move (swapping two opposite variables on the same sub-lattice, e.g. exchanging a Li ion and a vacancy), which conserves Li concentration, and 80% chance a grand canonical move (flipping together one Li and one electron occupation variable of the same sign to keep charge neutrality, e.g. transforming a $Li^+ + e^-$ pair into a vacancy + hole pair), which changes the number of Li atoms by one. The simulation is driven by the standard Metropolis algorithm [117]. Thermodynamical quantities are calculated at given temperature T and lithium chemical potential μ.

Fig. 3-6 shows chemical potential μ versus x when μ is increased gradually at given temperatures. A sudden jump in x signifies a first-order phase transition. At low-T (300 K) the transition goes directly from $FePO_4$ to $LiFePO_4$ while at elevated temperature (600 K) the transition passes an intermediate phase. Note that the $x \approx 0.1 \rightarrow 0.6$ transition splits into two steps, with a new phase appearing at $x \approx 0.25$ This $x \approx 0.25$ will be discussed in more detail later. At high-T (900 K) there is no sign of phase transition, i.e. there is only one single solid solution phase. Due to hysteresis the values of x and μ at the phase boundaries cannot be found directly in

the scan. The exact phase boundaries are obtained with free energy integration.

3.6.3 Results

We present the calculated phase diagram in Fig. 3-7. In the (T, μ) space, the phase boundary between two phases at equilibrium follows the following relationship according to Eq. 3.7:

$$\frac{d\mu}{dT} = -\frac{\Delta E}{T \Delta x} + \frac{\mu}{T} = -\frac{\Delta S}{\Delta x}, \tag{3.11}$$

where ΔE, Δx and ΔS are the finite difference in energy, concentration and entropy of the two phases, respectively. In Fig. 3-7a the phase boundary starts with equilibrium between the triphylite $LiFePO_4$ phase T and the heterosite $FePO_4$ phase H at room temperature at small positive μ and continues with large positive slope. According to Eq. 3.11 the entropy in H is slightly larger than in T. At $T_{eu} = 405$ K, $x = 0.47$ a eutectoid point occurs with an intermediate solid solution phase SS between H and T. At higher temperature the branch of the phase equilibrium between SS and T has positive slope, i.e. the disordered phase has higher entropy. The SS-T branch ends at a congruent point around $T = 810$ K where the $LiFePO_4$ phase merges into the solid solution. On the other hand, the slope of the H-SS phase boundary is negative until about 570 K, meaning that SS also has larger entropy than $FePO_4$. However at higher temperature the slope becomes positive, meaning that the opposite is true, and the solid solution phase *has to be far from a random solid solution to have relatively small entropy*. The unusual crossover of entropy one again reflects the complex nature of the Li_xFePO_4 phase diagram. The H and SS phases emerges at a congruent point of about 810 K.

Fig. 3-7b depicts the same phase diagram in the (T, x) space, which can be directly compared with experimental data in Fig. 3-2. Excellent agreement is indeed found: two-phase region at low temperature, appearance of the intermediate D phase at the eutectoid point, and dominance of the disordered phase above two congruent points. The solubility is less than 1% at room temperature, and slightly larger in $FePO_4$ than in $LiFePO_4$. The eutectoid temperature is only 20–70 K off experimental
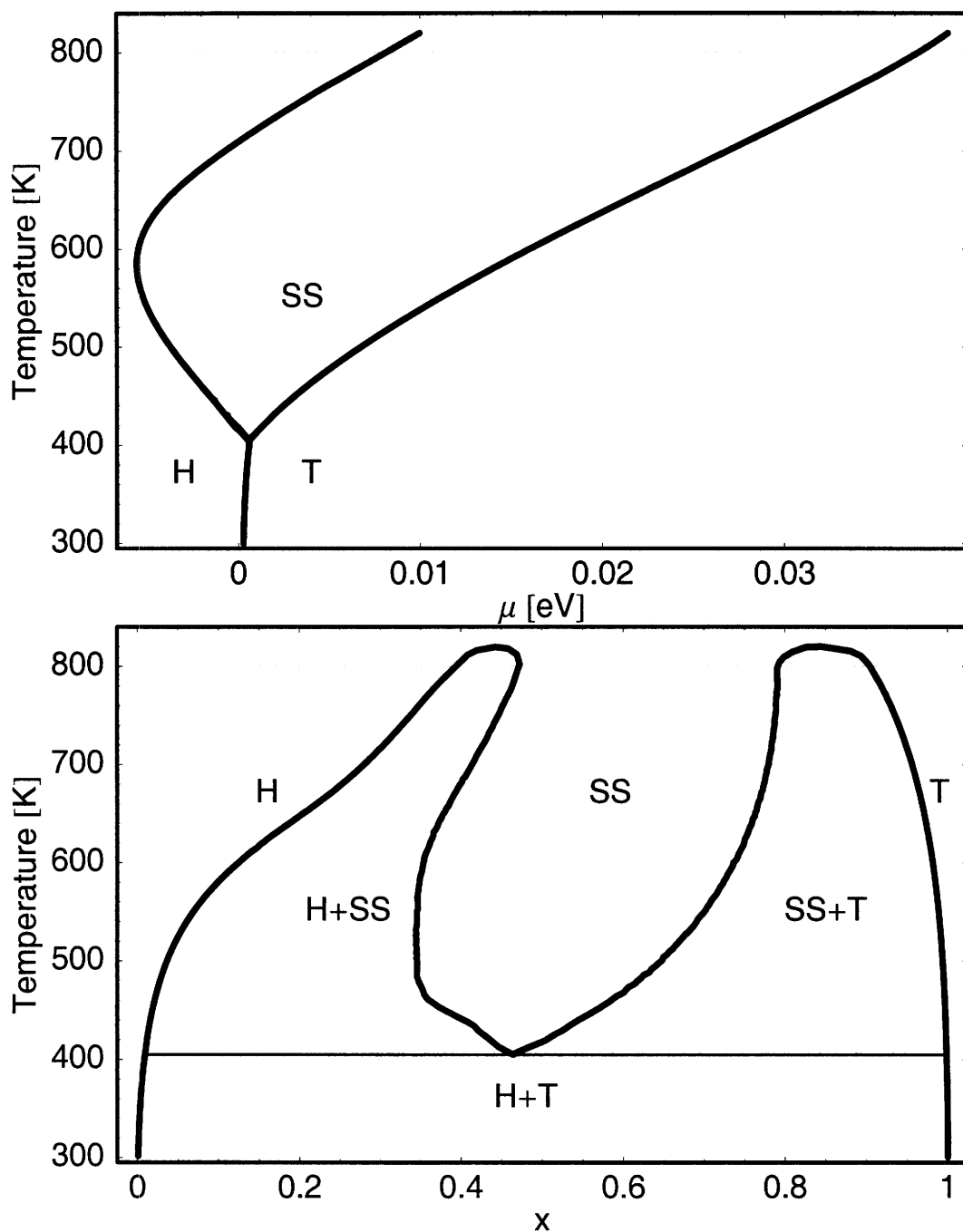
Figure 3-7: Calculated Li_xFePO_4 phase diagram taking into account both Li/v and e/h **a**) in the $T - \mu$ (arbitrary reference) space, **b**) $T - x$ space.

reports [51, 58], and the congruent temperature about 100–150 K off.

We predict the enthalpy of mixing at the eutectoid point to be 8.6 meV/formula unit, consistent with the measured lower limit 700 J/mol= 7.3 meV/formula unit for an $x = 0.47$ sample [58].

3.6.4 Phase diagram without explicit electrons

To understand better which physics determine the shape of Fig. 3-7b, we have performed calculations in the "traditional" way, i.e. to consider only the Li/V configurations as the entropy generating mechanism, assuming electrons always occupy the lowest energy state for each Li/V configuration.

Out of the set of 245 energies in Fig. 3-3, 124 represent the lowest-energy electronic state of their corresponding Li/V configurations. To enhance the predictive power of the CE model, 14 additional Li/V configurations have been evaluated with the Li+electron ECI in Eq. 3.12 to find the lowest energy e/h distribution. These 14 energies, as well as the 124 direct energies, are used to construct a lithium-only cluster expansion model. A lithium-only cluster expansion can be written as:

$$E[\vec{\lambda}] = J_\emptyset + J_i\lambda_i + J_{ij}\lambda_i\lambda_j + J_{ijk}\lambda_i\lambda_j\lambda_k + \dots \qquad (3.12)$$

The optimized Li-only CE has a cross validation of about 7 meV per formula unit, slightly larger than CV=6 meV in the Li+electron model Eq. 3.5. It consists of 29 ECIs: 1 constant, 1 point, 12 pair, 6 triplet and 9 quadruplet terms. Unlike Li+e pair ECIs in Eq. 3.5, a clear and physical picture of the Li-only ECIs in Eq. 3.12 cannot be established. The 9 quadruplet terms also make the Li-only CE model more complicated. Therefore we see two more advantages of explicitly treating the electrons:

1. that by add more degrees of freedom, less information is coarse grained, leading to better accuracy,

2. the interaction information of hidden variables (e/h in this case) obscured by
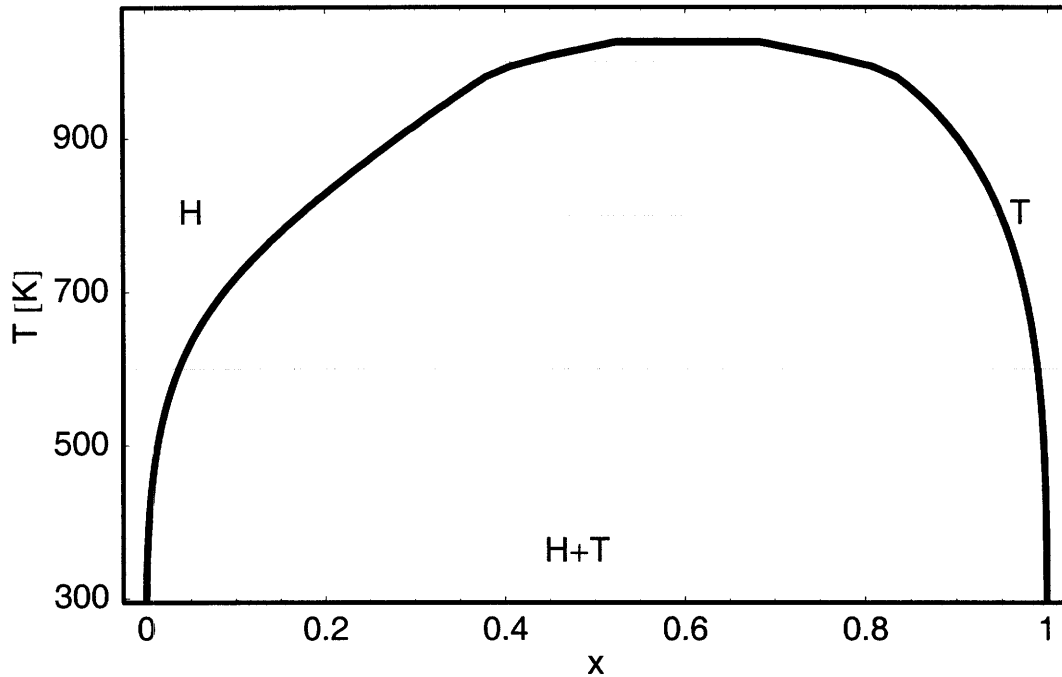
Figure 3-8: Same as Fig. 3-7b, except that explicit electron degrees of freedom are not considered.

coarse graining can be revealed to yield a physically more meaningful picture.

The calculated $T - x$ phase diagram (Fig. 3-8) shows a simple two-phase region, qualitatively different from experiment but similar to typical immiscible systems. The striking difference between the high-T portions of Fig. 3-7b and 3-8 points to the crucial importance of explicitly treating the electron degrees of freedom in excitations and finite-T thermodynamics of these mixed-valence systems. The low-T portions of Fig. 3-7b and 3-8 are similar in shape. The solubility limits at a given temperature is smaller in the Li-only phase diagram of Fig. 3-8 than in the Fig. 3-7b. Actually by re-scaling the temperature in Fig. 3-8 by a factor of about 0.85 the low-T portions of Figs. 3-7b and 3-8 match reasonably well. This indicates that the entropy in the Li-only simulations is smaller than in the Li+electron case.

3.7 Analysis of entropy contributions

3.7.1 Joint, conditional and marginal entropies

A deeper analysis of the phase diagram in Fig. 3-7b requires us to look into the entropy that drives the phase transition. The total (joint) configurational entropy $S(\text{Li}, \text{e})$ of the electronic+ionic system can be calculated through free energy integration in the simulation. To partition the entropy into ionic and electronic contributions, we note that

$$S(\text{Li}, \text{e}) = S'(\text{Li}) + S'(\text{e}) + I(\text{Li}, \text{e}), \tag{3.13}$$

where I is the mutual information of the two degrees of freedom, and

$$S'(X) \equiv S(X|Y) = \sum_y P(y)S(X|y) \tag{3.14}$$

is the conditional entropy from the X (Li or electron) degree of freedom, i.e. the entropy contribution of X with fixed Y, thermal averaged over the marginal probability distribution function (PDF) $P(Y)$. $S'(X)$ measures how random X can be when Y is frozen. If X and Y are independent, S' is exactly the entropy contribution from one degree of freedom, a direct physical meaning. But when the two degrees of freedom are coupled as lithium and electron are through Li-e (and triplet) interactions, no exact separate contribution may be defined. Nevertheless S' would still in some way preserve its characteristics as a separate entropy term.

One may also use the marginal entropy, i.e. the entropy associated marginal PDF $P(x) \equiv \sum_y P(x, y)$:

$$S(X) \equiv -k_B \sum_x P(x) \ln P(x) = S'(X) + I(X, Y), \tag{3.15}$$

to represent separate entropies. It differs from S' by the same value $I(X, Y)$ for both X and Y contributions. In this thesis we use S' to compare different entropy contributions, as it can be directly calculated in MC simulations.

```
{
    // Outer loop: normal grand-canonical MC
    grand canonical steps
    {
        // Inner loop: fix sub-lattice Y, perform canonical moves
        // within sub-lattice X and gradually cool down the system from
        // given temperature to (nearly) zero-T and integrate dS
        canonical steps within X
        calculate dE
        S_i=S_i+ dE/T
        reduce temperature till T_min
    }
}
// the averaged S obtained from the above loop is conditional entropy
average conditional entropy S_i for X
```

Figure 3-9: Algorithm to calculate conditional entropy S'_X with MC simulations.

3.7.2 Calculation of conditional entropy

To calculate conditional entropy S'_X, we have used the algorithm outlined in Fig. 3-9. According to Eq. 3.14, S'_X can be evaluated by (1) sampling the marginal distribution $P(Y)$ and (2) for each fixed configuration y calculating the entropy associated with conditional PDF $P(X|y)$. The distribution P_Y is reached with normal Monte Carlo steps in the outer loop of Fig. 3-9 to reach equilibrium at any given (T, μ). Then the current y_i configuration is frozen and in the inner loop entropy integration $\int dS = \int dE/T$ is carried out from equilibrium temperature T down to 10 K, resulting in the entropy term $S(X|y_i)$ associated with $P(X|y_i)$. Since y_i is sampled according to Y's marginal distribution, the average of all calculated $S(X|y_i)$ gives S'_X, while the weight $P(y)$ is naturally incorporated.

Entropy integration in the inner loop of Fig. 3-9 is computationally expensive, so the number of calculated $S(X|y_i)$'s is limited. A typical conditional entropy value presented in this section takes 20 iterations, with 100 temperature steps per iteration, and $\sim 10^5$ MC passes per step.

One may be tempted to calculate the total entropy $S(X, Y)$ with canonical simulations and entropy integration. However, by cooling down the whole system, the cell

usually goes across phase boundaries and this brings in unnecessary complications. In the S' calculations for Li_xFePO_4, one fixed sub-lattice locks the system in a single phase without phase transitions during cooling. The total entropy can be easily evaluated through free energy integration discussed in Section 3.6.1.

3.7.3 Results

In Fig. 3-10 we show the total and separate entropy along the solubility limits of the H and T phases (leftmost and rightmost phase boundaries in Fig. 3-7b, respectively), as well as along $x = 0.5$ in SS. At low-T the total entropy (bold lines in Fig. 3-10a) is small, slightly larger in H than in T. Larger solubility and entropy in the H phase can be explained by smaller defect energy in it (see formation energy in the dilute limit of Fig. 3-3 and discussions in Section 3.4.3).

The solid solution phase is far from random:

1. when it first appears at the eutectoid point, its entropy is a mere 0.3 k_B,

2. the total entropy of the H phase exceeds that of SS above about 570 K even though its composition is lower,

3. up to 900 K, well above the congruent points, the total entropy 1.1 k_B of SS($x = 0.5$) is still smaller than (complete random) $2S_e^{loc,rand}(0.5) = 1.39$ k_B.

The difference between $S(Li, e)$ and $S'_{Li} + S'_e$ (thin dashed curve of Fig. 3-10a) is the mutual information $I(Li, e)$, indicating how correlated the two degrees of freedom are.

Fig. 3-10b shows separate S'_{Li} and S'_e in dashed and dotted curves, respectively. It is noteworthy that in all but the T branches S'_e is noticeably larger than S'_{Li}; surprisingly, S'_e dominates the SS phase and contributes much more than S'_{Li}. At the eutectoid point the mixing entropy driving the transition into SS is overwhelmingly electronic: 0.19 k_B from S'_e and 0.05 k_B from S'_{Li}. A qualitative explanation for the larger S'_e is that the leading J_{e-e} terms are weaker than the leading J_{Li-Li}, and the electron excitation spectrum at a fixed Li configuration is lower in energy than the opposite. We therefore conclude, to the extent S' represents a separate entropy,
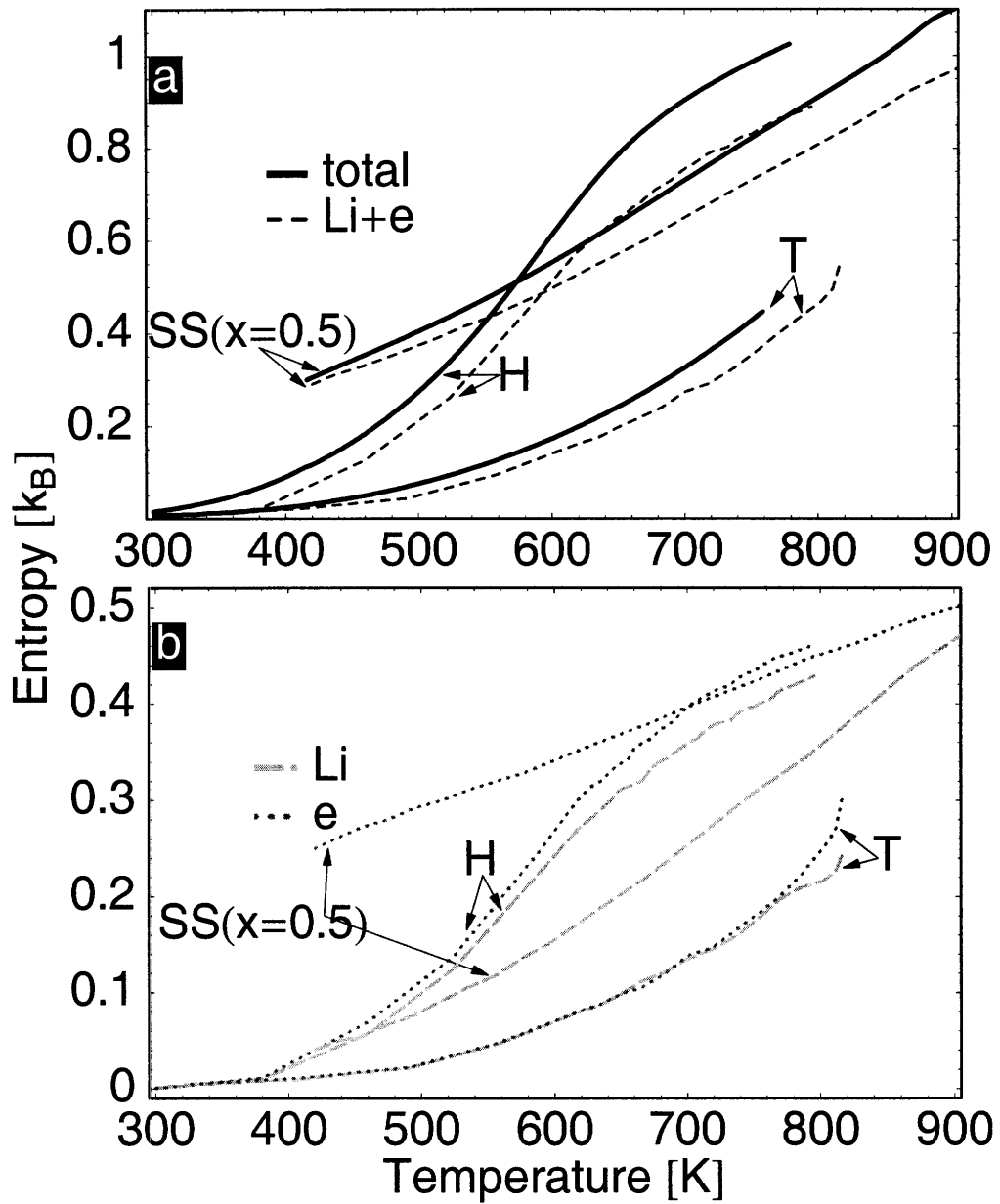
Figure 3-10: Configurational entropy per formula unit. a) total entropy and the sum $S'_{\mathrm{Li}} + S'_{\mathrm{e}}$ for comparison; b) separate conditional entropy S'_{Li} and S'_{e}.
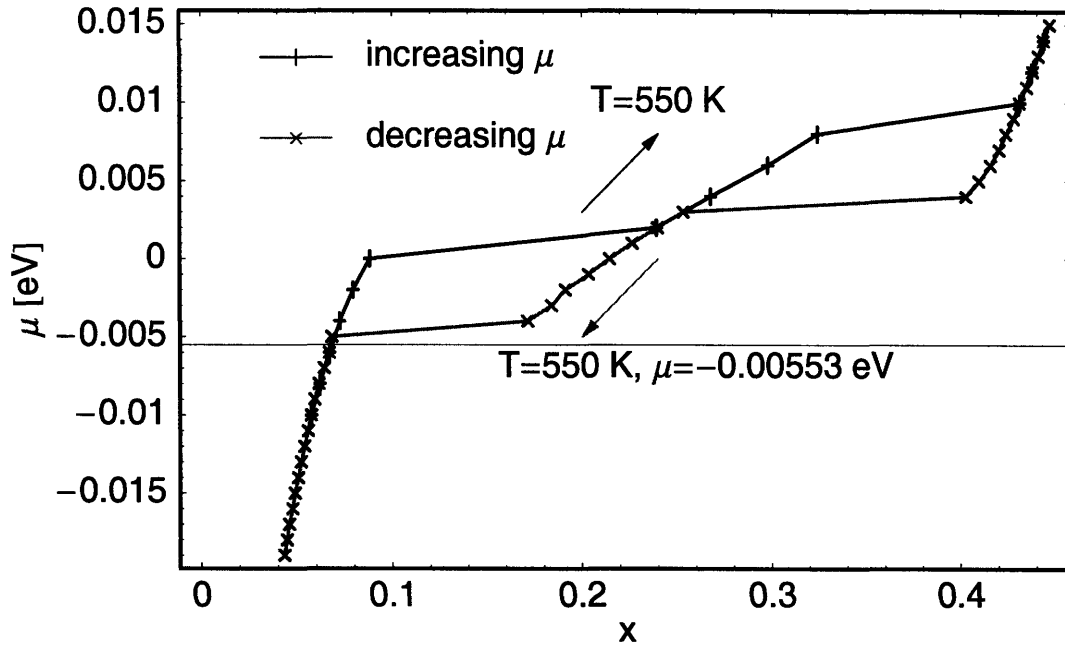
Figure 3-11: Chemical potential versus x at $T = 550$ K by increasing and decreasing μ, respectively

that the electron degree of freedom contributes at least as much as the Li ions to disordering of the system, and that the formation of the solid solution state is driven by e/h disorder. To our knowledge, no other examples of electronic entropy-driven solid solution have been identified, though electronic entropy driven modification of ordering interactions through band entropy has been proposed for Ni_3V [85].

3.8 Meta-stable phase

In section 3.6.2 another intermediate phase (M) other than the solid solution was observed at $T = 600$ K when the chemical potential is increased gradually in the MC simulations. Here we study this phase more closely by scanning the chemical potential in both directions at $T = 550$ K. The results are shown in Fig. 3-11. The thermodynamical equilibrium chemical potential at 550 K, according to Fig. 3-7a, is labeled by the horizontal line. Due to hysteresis the M–H transition chemical potential is between those observed in the two runs (increasing and decreasing of μ), which are both above the horizontal line. If M were a thermodynamically stable phase at 550

K somehow missed in our previous MC simulations, its range of composition had to be larger than the lower bound of SS's composition in Fig. 3-7b at that temperature, or about 0.35. However, according to Fig. 3-11 M appears around $x = 0.25$. *We therefore concluded that the phase M is meta-stable.* According to our simulations phase M appear around $0.15 < x < 0.3$ and $470 < T < 575$ K. Recently, there is evidence for a new $Li_x FePO_4$ phase around $x = 0.1$–0.2 [73, 104].

3.9 Discussions and conclusions

Our calculations depend on relatively small energy terms in delicate balance. For example, slight ΔE asymmetry introduces small triplet ECIs that in turn result in larger solubility limit and entropy in $FePO_4$, as well as the crossover of entropy in the H and D phases above 570 K. We believe the delicacy is a reflection of the nature of the $Li_x FePO_4$ binary phase diagram, not numerical artifact. In deed, the disordered phase is stabilized by mixing enthalpy as small as 7 meV. Consequently it takes a very accurate energy model to get reliable results. Our GGA+U approach has been bench-marked over other properties of the material as a robust tool. With reasonable fine-tuning of the cluster expansion model construction, our ECIs are fairly stable. As a result details in the calculated phase diagram, such as transition temperatures and concentration at the eutectoid point, can be slightly variable, while the bulk behavior persists.

In oxides both electron localization and delocalization can occur. For example, a system such as $Li_x CoO_2$ is metallic for $x < 0.9$ [116] and explicit e/h entropy is less crucial. LDA and GGA in which mixed valence states are delocalized will be an adequate treatment for such system [189]. On the other hand materials in which carriers localize require more careful treatment both for their energy calculation (e.g. in GGA+U, SIC methods or DMFT [69]), and for their contribution of the electronic degree of freedom to the entropy as demonstrated in the present work. A simple model can be setup to estimate the effects of configurational e/h ordering based on Δn, the d-electron occupancy difference at variable valence transition metal. This number is 0.35

in Li_xFePO_4. The e-e coupling scales as $(\Delta n)^2$, and the coupling to Li scales as Δn. In the limit of large Δn, or well localized electrons, where the strong Li-e coupling could overshadow the Li-Li repulsion, an ordered intermediate structure would be penalized. Li_xFePO_4 may be close to this case. However, in the limit of very strong electron-ion coupling (e.g. Al^{3+} instead of Li^+), the localized electrons are closely bound to the ions, and their independence as a degree of freedom maybe diminished. In the limit of small Δn, the Li-el binding is weak and Li ordering is favored. This include the scenario where the electrons are more delocalized. Li_xCoO_2 with $\Delta n = 0.15$ may be closer to the second limit. It should be noted that in our MC simulations, e/h are treated as classical particles (but not in the DFT energy calculations). If hopping becomes so fast that electron wavefunctions overlap, the notion of localized electrons becomes meaningless, and it becomes difficult to enumerate the eigenstates over which to sum excitations, until one reaches the nearly free-electron limit where the band picture is applicable. It is up to further investigation to establish quantitative effects of the localized electron degrees of freedom in thermodynamics of other transition metal oxides.

Chapter 4

Cluster Expansion for Protein: Theory

4.1 Introduction

Proteins are the active units that drive living cells and, ultimately, organisms. The building blocks of proteins are called *amino acids* (AA). Twenty different types of amino acids are involved in making most proteins. They all share a similar backbone (main-chain) structure. The difference between them is the so-called side-chain "R" group. As shown in Fig. 4-1, amino acids are strung together in particular *sequences*, which can be designated by the species of side-chains $R_1, R_2, \cdots$. A given sequence will fold up into a specific structure. A short protein sequence, usually less than 100 residues, is also called a peptide, although there is no exact distinction between the two names.

Protein folding and protein design, including the prediction and design of macro-molecular interactions, stand among the most formidable and significant challenges in contemporary computational biology. The function of proteins is dictated by their structure. The 3-D structure of a protein is uniquely encoded in its 1-D sequence. Enormous theoretical and computational research effort has been devoted to under-standing this encoding [91, 146, 196]. The problem can be posed two ways: (1) protein *folding* deals with predicting the final 3-D structure of a protein given its
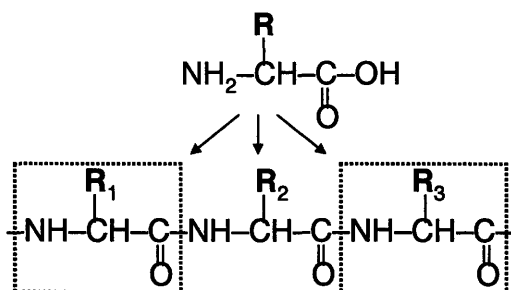
Figure 4-1: Schematics of an amino acid (top) and bonding of AAs into a peptide (bottom). Side-chain = group **R**, backbone = other atoms.

AA sequence. Because the number of proteins with known sequence far exceeds the number with known structure, an ability to predict structure from sequence would be extremely valuable. (2) Protein *design* is concerned with finding an optimal sequence to fold to a pre-defined structure. Protein design is useful both because it allows for the engineering of macromolecules with desired properties [63, 5, 96], e.g. the usefulness of proteins as reagents and therapeutics [100], and because the development of computational design methods deepens our general understanding of protein folding and stability. In this thesis, we focus on the protein design problem with physical interaction models.

As shown in Fig. 4-1, all atoms in a protein can be classified as either "*backbone*" or "*side-chain*". The backbone atoms are the same for each AA and represent the overall structure or "fold" of a protein, as shown for two examples in Fig. 4-2A. The side-chain atoms are different for different AAs, and give rise to additional degrees of freedom termed "side-chain conformations" or, when discretized, "*rotamers*" (see Fig. 4-2B-C).

Many applications in computational structural biology involve evaluating the energy of a protein adopting a **specific** structure. Unless stated otherwise, we limit our study to problems with **fixed** protein backbone. A variety of functions are used for this purpose. Statistical potentials are fast to evaluate but do not have a clear physical basis, whereas physics-based functions consist of well-defined terms that can be costly to compute. This chapter describes how the theory of cluster expansion, originally developed to describe the energies of alloys, can be applied to generate a
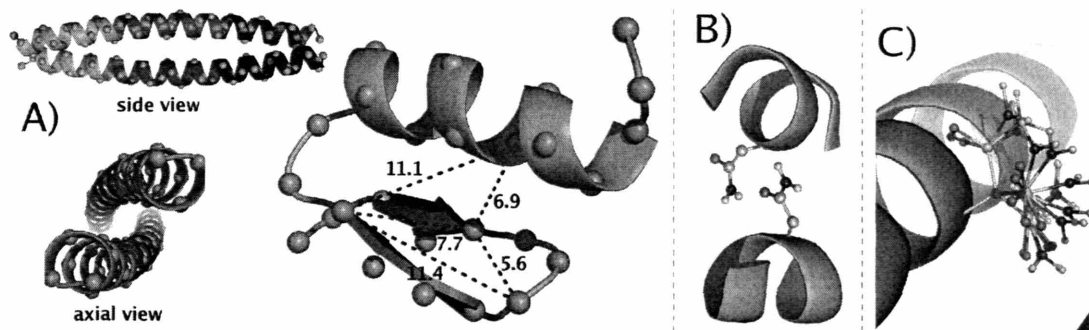
Figure 4-2: Small protein molecules in ribbons and in ball-and-stick representations. A) The coiled coil (left) top - side view; left - helix axis view) and the zinc-finger (right) protein folds. Orange spheres are backbone atoms and the ribbons are a cartoon representation of the backbone geometry. The coiled-coil unit cells are highlighted. Futher discussions about the coiled coil and the zinc finger folds can be found in Chapter 5. B) The optimal rotamers for two AA's in an all-atom ball-and-stick representation. C) A set of common rotamers for one AA shown superimposed.

physical potential function for proteins that is extremely fast to evaluate. The theoretical framework of cluster expansion is presented, with emphasis on the adaptations to the very high dimensional nature of biological systems. Results of applications of the method will be presented in the next chapter.

4.2 Energy evaluation

At the heart of any computational approach to protein design or structure prediction lies the problem of determining the *fitness* (*effective energy*) of a particular protein in a **given** conformation or state. Depending on the relevant application, this *effective energy* may correspond to different physical quantities, e.g. stability, solubility, binding affinity, catalytic efficiency or a combination thereof. In protein design, the goal is to optimize this fitness in the large space of possible amino-acid sequences. In the fold-recognition approach to structure prediction (also called threading), the goal is to identify the most suitable structure for a particular sequence, given a library of known folds. In both cases the complexity of the problem imposes two sometimes conflicting requirements on the energy function used: physical accuracy and computational efficiency.

91

4.2.1 Protein Energy models

There are two major classes of fitness functions used in the fields of structure prediction and design. Lazaridis and Karplus [102] refer to these as statistical effective energy functions (SEEFs) and physical effective energy functions (PEEFs).

SEEFs are derived from databases of proteins with known structures and describe the distribution of residues (or atoms) at different distances, solvent exposure, and sometimes more complicated measures, such as local atom density or relative orientation of secondary structure elements [157]. These terms are treated as effective potentials for calculating the energy of a protein in a given conformation. Most statistical energy functions include up to pair interactions [213, 70, 168]. However, it has been suggested that pairwise statistical energy functions may not be suitable for protein design or fold prediction [195, 118], so some SEEFs include higher order terms [118, 27, 154]. The advantages of SEEF methods lie in their computational efficiency, simplifying abstraction from details, and ability to implicitly capture effects such as desolvation, loss of entropy, and the hydrophobic effect, which are hard to account for explicitly. To achieve these benefits, accuracy and physical interpretability are compromised.

Physically motivated effective energy functions use atomic-level representations to capture underlying physical phenomena and approximate the free energy of the studied system. These express the energy of a protein sequence adopting a specified structure in terms of atomic coordinates. Let the variable $\sigma^i = 1 \ldots m$ indicate which of the m AAs is present at site i. A sequence is then expressed by $\vec{\sigma} = \{\sigma^1, \ldots, \sigma^L\}$. The energy of a protein $E[\vec{\sigma}, \vec{\tau}]$ depends on this sequence and on the other microscopic information $\vec{\tau}$ (e.g. coordinates of all atoms on the protein and solvent molecules). Some of the terms commonly included in PEEFs are van der Waals interactions, electrostatic interactions, hydrogen bond energies, dihedral angle torsion energies, atomic desolvation energies and solvent-accessible-surface area or volume-dependent estimates of the hydrophobic effect [102, 72, 91, 146].

$$E[\vec{\sigma}, \vec{\tau}] = E^{\text{vdW}} + E^{\text{elec,wat}} + E^{\text{h-bond}} + E^{\text{solv,sc}} + E^{\text{torsion}} + \cdots, \qquad (4.1)$$

Some attempts have also been made to model side-chain entropy [41]. The advantage of PEEFs is that they have the potential to provide a more comprehensive and interpretable understanding of the observed phenomena. The disadvantages are that much of the underlying physics is difficult to account for quantitatively, and when it is possible to do so, it is usually computationally expensive. In this thesis we focus on physics base energy models.

4.2.2 Energy coarse-graining

The relevant thermodynamic function, $F[\vec{\sigma}]$, is the free energy as a function of sequence $\vec{\sigma}$ at finite temperature, and can be calculated by phase space (usually non-sequence degrees of freedom $\vec{\tau}$) sampling and integration. This is simplified in many cases to finding the energy minimum over $\vec{\tau}$:

$$F[\vec{\sigma}] \approx E_{\min}[\vec{\sigma}] = \min_{\vec{\tau}} E[\vec{\sigma}, \vec{\tau}]. \tag{4.2}$$

Practical evaluation of Eq. 4.2 can be quite expensive. Aside from the complexity of the atomistic energy model, calculating the physically meaningful energy (or free energy) of a sequence involves a search through the conformational space. To reduce the phase space, the degrees of freedom of the solvent molecules are often treated implicitly using continuum dielectric models, leaving only the side-chain degrees of freedom. Still, this is a combinatorial optimization problem. Even for a relatively small protein fold of 100 AAs and 10 common rotamers per AA there are $\sim 10^{100}$ rotamer configurations in the search space. The computational complexity of high-quality physics-based scoring functions makes optimization over all sequences and rotamers infeasible.

An optimal energy function would have the simplicity and computational efficiency offered by SEEFs while retaining the theoretical rigor and physical interpretability of PEEFs. Because sequence determines the properties of a protein given a defined environment, a function should exist that coarse-grains Eq. 4.2 and maps sequence directly to properties (energy in particular), regardless of the complexity of the underlying

93

physics that determines that energy. A sufficiently accurate and computationally tractable approximation of this function would find widespread use in computational studies of protein structure. In the next sections we present a general method by which the energy of a protein on a fixed backbone, based on an arbitrary energy function, can be accurately expressed as a simple function of its sequence. In principle, this method can be applied in conjunction with any energy function, the only limitation on the complexity being that energies for enough training sequences can be generated, at reasonable computational effort.

4.3 Cluster Expansion: from alloys to proteins

Mapping sequence to energy is similar to the configurational problem in alloy theory [159, 49, 28], where the the lattice configurations (occupations of the lattice sites by different atoms) approximately label the possible states of the system and specify the energy. Using this simplified picture, the cluster expansion maps the true alloy Hamiltonian onto a very simple Ising-like form. The cluster expansion (CE) [159, 49] is a method for representing a property (energy in our case) that depends on discrete and topologically ordered degrees of freedom in a system. The method finds its origin in alloy theory, where very expensive quantum-mechanical calculations are required to accurately capture material properties, and only computations on a small number of atomic arrangements with relatively small unit cells are possible. The cluster expansion is essentially a parameterization of the energy in terms of discrete variables that give the occupancy of each lattice point in the crystal. When the occupation variable is a spin variable ($\sigma = \pm 1$), the CE takes on the form of a generalized Ising model. The success of the cluster expansion formalism has been extraordinary in studies of metals, semiconductors, and ceramics, and its ability to reproduce experimental thermodynamic data has been confirmed repeatedly in binary systems [67, 189, 15, 185] (in analogy to the HP model [98] for protein), and in multi-component systems [49, 115, 204]. In essence, the lattice model form of the alloy Hamiltonian can be seen as the result of integrating out faster degrees of

94

freedom such as small lattice displacements (static or dynamic) and electronic degrees of freedom. The CE is essentially a technique to expand this Hamiltonian in easy to identify basis functions. Formally, CE is an expansion of the energy in a set of linearly independent basis functions that span the relevant configuration space (e.g. all possible distributions of atoms A and B on a crystal lattice, or all possible amino-acid sequences on a protein backbone). In most forms, the basis set of the cluster expansion is mathematically complete by construction, and a full expansion will result in a perfect representation of the energy. Truncated expansions may have practical utility, however. The use of a truncated cluster expansion to model the energy is analogous to using any truncated expansion in basis functions (e.g. plane waves or spherical harmonics) to represent a complex unknown function. The goal in developing an effective CE is to identify a truncated expansion that, when fit to a training set of data, provides an accurate mapping between degrees of freedom and energy using a minimal number of parameters.

Under this general perspective of cluster expansion, we find striking connections between solid-state alloys systems and proteins, illustrated in Fig. 4-3. One can

Alloys

- Built of elements
- ~2-10 elements per alloy
- Different **orderings** → distinct properties (energy, conductivity, hardness, *etc.*)
- Cluster expansion efficiently models ordering dependence

Proteins

- Built of amino acids
- ~20 distinct amino acids
- Different **sequences** → distinct properties (energy, folds, docking, etc.)
- Cluster expansion efficiently models sequence dependence

Figure 4-3: Connection between the alloy and the protein systems. Colors designate different elements in alloy or amino acids in protein.

make a correspondence between an alloy lattice and a protein backbone and between alloy constituent elements and amino acids. While in alloys one typically treats

binary distributions (two possible species per site) or on rare occasions ternaries [6, 114], the complete collection of natural amino acids requires twenty species per site. Such a dramatic increase in phase space requires some reformulation of the CE implementation typically used for alloys. The general idea is to define a set of basis functions that correspond to the energetic contributions of single amino acids at single sites, pairs of amino acids at pairs of sites, triples of amino acids at sets of three sites, and so on. If intuition holds, the lower-order terms in this expansion will be more important than the higher-order ones, and a truncated expansion will be sufficient to represent the energy. In practice, given a set of training sequences and their energies, the CE is derived by starting with lower order terms and successively considering higher order contributions until a fit of the expansion to the data gives adequate performance when tested under cross validation. A formal mathematical description of the theory of CE as we have applied it to protein energetics follows.

4.3.1 Basic formalism

Given a discrete variable σ that can take on M different values ($\sigma = u_0, \ldots, u_{M-1}$), all real functions of it $f : \{u_0, \ldots, u_{M-1}\} \longmapsto \mathbb{R}$ form an M dimensional Hilbert space:

$$\langle f, g \rangle = \sum_{\sigma} f(\sigma) g(\sigma). \tag{4.3}$$

To facilitate further discussions we expand the definition of f naturally to linear combinations of allowed values *

$$f(\sum_{x} c_x u_x) \equiv \sum_{x} c_x f(u_x). \tag{4.4}$$

---

*Linear combination or discrete values may become meaningful when, for example, one is dealing with a alloy site with equal chance of being occupied by element A or B, or an ambiguous residue in a protein that can be either Glutamine or Glutamic acid. If we are interested in the statistically averaged sequence property, the occupation variable may be expressed in combinations of multiple allowed species

Function f can be expanded using a basis set of M linearly independent functions $\Phi = \{\phi_0 \equiv 1, \phi_1, \ldots, \phi_{M-1}\}$:

$$f(\sigma) = \sum_a J_a \phi_a(\sigma) \tag{4.5}$$

where coefficients J are constants and functions ϕ are called *point functions* (PF). A similar statement can be made about any function f of N discrete variables $\vec{\sigma} = \{\sigma^1, \ldots, \sigma^N\}$, because $\vec{\sigma}$ can be thought of as a discrete variable with M^N possible values. [†] Thus, to expand $f(\vec{\sigma})$ exactly, a basis set with M^N functions is needed. According to the cluster expansion formalism [159], a particularly convenient basis set for expanding f can be obtained by considering all the possible tensor products between point functions in the N point basis sets Φ^i ($i = 1, \cdots, N$), each completely describing the sequence space at site i. Thus, a basis set suitable for expanding f is defined in the product space of the point functions:

$$\Phi^1 \otimes \cdots \otimes \Phi^N = \left\{ \begin{array}{ccc} 1 & & \\ \phi_1^1, & \ldots, & \phi_{M-1}^N \\ \phi_1^1 \phi_1^2, & \ldots, & \phi_{M-1}^{N-1} \phi_{M-1}^N \\ \cdots & & \\ \phi_1^1 \cdots \phi_1^N, & \ldots, & \phi_{M-1}^1 \cdots \phi_{M-1}^N \end{array} \right\} \tag{4.6}$$

where in row k, the subscripts indexing point functions independently run through $1 \ldots M - 1$ and the superscripts indexing variable sites take on all possible combinations of $k - 1$ different values in $1 \ldots N$. Each basis function in set 4.6 depends on the value (amino-acid identity in the case of proteins) at either no site (constant function 1 on the first row), one site, two sites and so on. We call a set of specific sites a *cluster*. Each size-k cluster $\{i_1, \cdots, i_k\}$ has $(M - 1)^k$ basis functions, or *cluster functions* (CFs), associated with it: $\phi_{a_1}(\sigma^{i_1}) \ldots \phi_{a_k}(\sigma^{i_k})$, because each subscript a can indepen-

---

[†]Without loss of generality we assume the same number of allowed values M for each discrete variable for simplicity. It takes a simple extension to rewrite a more general formalism with site-dependent number of allowed values, or even site-dependent sets of point functions, resulting in merely more book-keeping of indices in the expressions. This is sometimes called a coupled cluster expansion in alloy theory. The CE applications in the next chapter do deal with different types of sites.

dently be assigned 1 through $M - 1$. Since there are C_N^k different size-k clusters, the total number of k-body cluster functions (in row $k + 1$ of Eq. 4.6) is $C_N^k(M - 1)^k$. For example, the second row of Eq. 4.6 contains N clusters corresponding to N sites, each with $M - 1$ cluster functions. Note that point function index a of ϕ_a should not be zero in any k-body cluster function, because $\phi_0 \equiv 1$ would reduce the order of the CF by one. The total number of linearly independent cluster functions in the basis set is therefore

$$\sum_{k=0}^{N} C_N^k(M - 1)^k = M^N,$$

the same as the number of possible values of the discrete variable $\vec{\sigma}$, and the set 4.6 forms a complete basis for the functional space of f. Given the constructed basis set, we can exactly expand f as:

$$f(\vec{\sigma}) = \sum_I \sum_A J_A^I \phi_A^I \tag{4.7}$$

where $I = \{i, \ldots, j\}$ is a cluster, $A = \{a, \ldots, b\}$ is the corresponding set of point function indices, $\phi_A^I = \phi_a(\sigma^i) \ldots \phi_b(\sigma^j)$ is the cluster function associated with cluster I and index A, and the coefficients J_A^I are referred to as *effective cluster interactions* (ECI). The expansion 4.7 is in principle exact, if all terms are included, though in practice the expansion has to be truncated.

4.3.2 Correlation matrix

Given sequences $\vec{\sigma}_1, \cdots, \vec{\sigma}_L$ we call the table of corresponding cluster function values a *correlation matrix*:

$$M = \begin{pmatrix} 1 & \phi_1(\sigma_1^1) & \cdots & \phi_1(\sigma_1^1)\phi_1(\sigma_1^2) & \cdots \\ & & \cdots & & \\ 1 & \phi_1(\sigma_L^1) & \cdots & \phi_1(\sigma_L^1)\phi_1(\sigma_L^2) & \cdots \end{pmatrix}, \tag{4.8}$$

where in each row are cluster functions evaluated at a given sequence. Matrix M characterizes the input sequences with correlation information among different sites.

It plays a central role in the linear equation between sequence and property

$$\vec{f} = M\vec{J},$$
(4.9)

where $\vec{f}$ is the list of target properties from input sequences. The list of cluster functions in $\vec{J}$ and in each row of M is usually a subset of all CFs, as in a truncated form of expansion 4.7.

4.3.3 Symmetry

Cluster expansion of alloy systems can be simplified when the host lattice possess symmetry. A *symmetry* operation $\mathcal{A}$ associated with the lattice can be defined as a mapping of the lattice to itself such that property f is invariant under $\mathcal{A}$:

$$f(\sigma) = f(\mathcal{A}(\sigma)).$$
(4.10)

It can be shown that two symmetrically equivalent cluster functions ϕ_A^I and ϕ_B^J (i.e. $\mathcal{A}(\phi_A^I) = \phi_B^J$) have identical ECI in the expansion 4.7. This allows one to rewrite 4.7:

$$f = \sum_{I,A} J_A^I \phi_A^I = \sum_{(I,A)_s} J_A^I \sum_{I',A'} \phi_{A'}^{I'}$$
(4.11)

where the first summation is over all symmetrically distinctive cluster function indices $(I, A)_s$, and the second over all CFs $\phi_{A'}^{I'}$ equivalent to ϕ_A^I. The set of equivalent cluster functions are said to form an *"orbital"*. Note that the empty cluster always forms an orbital of itself. In a system of high symmetry, the number of symmetrically inequivalent CFs is greatly reduced. An important example is the translational symmetry operators in a crystalline alloy, where the ECI of any non-constant CF is unchanged by space translation, reducing N (the number of unit cells in the system) otherwise independent coefficients to just one ECI of an orbital containing N CFs. For another example, the symmetry operators in the point group O_h of a simple cubic lattice render the 6 nearest-neighbor pair clusters equivalent.

99

Using the above symmetry considerations, one can often greatly reduce the complexity of the cluster expansion model. Fewer independent coefficients also mean less numerical noise in their calculation and a smaller training set.

Symmetry could be relatively scarce in biological systems. Most biological macromolecules do not possess any symmetry operators. Nevertheless there are some well-known exceptions in nature, including translational symmetry in ideal α-helices and translation as well as screw axis in DNAs and the coiled coil (see Chapter 5, where we take advantage of the symmetrical geometry of the coiled coil to reduce the number of independent ECIs).

4.3.4 Effective interactions

In a simple binary system there is only one ECI associated with each cluster, which can be interpreted easily as a measure of coupling energy. For multi-component ($M > 2$) cases, the ECIs for a given cluster are labeled by point function indices and their physical meaning is obscured. We define the *effective interactions* (EI) of a cluster in terms of occupation variables by summing ECIs over their point function indices:

$$V(\sigma^i \ldots \sigma^j) = \sum_{a \ldots b} J^{i \ldots j}_{a \ldots b} \phi^i_a(\sigma^i) \ldots \phi^j_b(\sigma^j). \tag{4.12}$$

The EIs of cluster $\{i, \cdots, j\}$ in Eq. 4.12 are identified directly by the variable species at that cluster. For example, a pair cluster's EIs are in the form of an $M \times M$ interaction table, similar to the 20×20 statistical contact energy table for amino acids [120, 121]. Detailed discussion of the interpretability of EIs will be presented later. The M^k EIs of a size-k cluster are derived from $(M-1)^k$ independent ECIs, hence they are linearly dependent.

The total target property in Eq. 4.7 can be rewritten with EIs as

$$f(\vec{\sigma}) = \sum_I V(\vec{\sigma}^I), \tag{4.13}$$

where I represents all clusters. Eq. 4.13 further reduces the energy calculation to

several energy-table look-ups and a final summation, obviously providing a dramatic speed-up over the full energy optimization scheme in Eq. 4.2.

4.4   Point functions

So far we have not discussed the effects of different point function basis sets. The choice of point functions $\Phi = \{\phi_0 \equiv 1, \phi_1, \ldots, \phi_{M-1}\}$ is in principle arbitrary, but it has significant practical consequences, since the ECIs are dependent on the choice of point functions. It is not a particularly noticeable problem for low-dimensional cluster expansion problems such as in binary alloys. Conventionally $\phi_1 = \pm 1$, a simple "switch" function, is used to denote the two possible states. The set $\{1, \phi_1\}$ constitutes an orthorgonal basis according to definition 4.3. The problem becomes more obvious in applications for proteins because: (1) biologists are interested in the interaction energy between amino acid species, which gives them intuition about the protein structure being studied. If reasonable and consistently defined interaction energy can be extracted from cluster expansion, the cluster expansion method would see more application in the field. (2) In a high-dimensional cluster expansion the large number of parameters (ECI's) is a serious practical problem for numerical evaluation and stability. It is desirable to construct a cluster expansion model with as few ECIs as possible and keep only the "essential" degrees of freedom.

The above argument is similar to the choice of electron wavefunction basis for quantum mechanical calculations: as long as the basis is complete and linearly independent, the physics would not be affected. In practice, the orthogonal planewaves are mathematically elegant, but sometimes suffer from slow convergence and lack of direct physical interpretation. Localized functions like atomic/molecular orbitals or Gaussian functions, though difficult to handle, may converge faster and be interpreted easily. In this section and the next we wish to construct a set of non-constant point functions taking into account the nature of the interactions between the amino acids, with emphasis on convergence and interpretability.

4.4.1 Background component

To facilitate further discussions we make some definitions here. First we define the orthogonal-complement function ϕ_B with respect to all non-constant point functions ϕ_a ($a = 1, \cdots, M-1$), i.e. $\phi_B \perp \phi_a$, or

$$\langle \phi_B, \phi_a \rangle = \sum_\sigma \phi_B(\sigma)\phi_a(\sigma) = 0. \tag{4.14}$$

Because non-constant functions ϕ_a and ϕ_0 ($\equiv 1$) constitute a complete basis set for a M-dimensional functional space, one is always able to uniquely define ϕ_B up to a prefactor. This definition also specifies a unique linear combination of allowed components $\sigma_B = \sum_{x=0}^{M-1} \phi_B(u_x)u_x$, which satisfies for any non-constant ϕ_a:

$$\phi_a(\sigma_B) = \sum_x \phi_B(u_x)\phi_a(u_x) = \langle \phi_B, \phi_a \rangle = 0 \tag{4.15}$$

$$V(\cdots, \sigma_B, \cdots) = \sum_{\cdots} \sum_{a=1}^{M-1} \cdots \phi_a(\sigma_B) \cdots = 0. \tag{4.16}$$

Eq. 4.16 shows that σ_B does not take part in any effective interaction, and we call it a *background component*. So whenever we choose the point functions at a certain site in cluster expansion, a background naturally emerges eliminating any EI involving it. If the point functions $\{1, \phi_a\}$ are orthogonal, ϕ_B is identical to 1 and σ_B becomes the simple average of all allowed components.

4.4.2 Effects of basis set on EI

For simplicity consider a change of point basis functions at site n. (To incorporate changes at multiple sites we just need to apply the following argument repeatedly.) Any EI involving site n is given by

$$V(\cdots, \sigma^n, \cdots) = \sum_{\cdots} \sum_a \sum_{\cdots} J_{\cdots a \cdots}^{\cdots n \cdots} \phi^{\cdots} \phi_a^n(\sigma^n)\phi^{\cdots}.$$

If ϕ_B, the orthogonal-complement of non-constant point functions, is unchanged, the transformation of point functions only involves the non-constant ones. New point functions $\{\phi'_1 \cdots \phi'_{M-1}\}$ are related to the original $\{\phi_1 \cdots \phi_{M-1}\}$ by $\phi'_a = A_{ab}\phi_b$ or $\vec{\phi}' = A\vec{\phi}$. Consider ECIs with all but the point function index for site n fixed: $\vec{J} = J^{\cdots,n,\cdots}_{\cdots,a,\cdots}$. As coefficients of cluster functions, $\vec{J}$ transforms as $\vec{J} = A\,\vec{J}'$. Following these transformations we have

$$V' = \vec{J}' \cdot \vec{\phi}' = \vec{J}' \cdot A \cdot \vec{\phi} = \vec{J} \cdot \vec{\phi} = V. \qquad (4.17)$$

Therefore the effective interactions are invariant to the change of basis as long as the functional space spanned by the $M - 1$ non-constant basis functions, or equivalently the orthogonal-complement function ϕ_B, is fixed. Linear transformation within $\{\phi_1, \cdots, \phi_{M-1}\}$ has no effect on EI. The EI does depend on ϕ_B, because it vanishes at the corresponding background component (Eq. 4.16).

4.4.3   Interpretation of the expansion

The background component, determined by the choice of point functions, effectively sets a reference state for the cluster expansion model with respect to which we can measure the effective contributions of different species. Now we may attempt to associate physical interpretation with the effective interactions. The constant ECI/EI has a clear meaning of being the energy of a hypothetical system that is completely (and statistically) composed of the background component, out of which only the constant term $J_\emptyset$ of Eq. 4.7 survives. An arbitrary EI should be taken as energetical deviation contributed by the involved components, measured against a non-interacting background.

For example, if in a A/B binary alloy $\phi_1(\sigma) = \pm 1$, we have $\phi_B \equiv 1 = \phi_0$, $\sigma_B = (A + B)/2$ since $\phi_1((A + B)/2) = (1 - 1)/2 = 0$. Physically it corresponds to a completely random $A_{0.5}B_{0.5}$ at high temperature. In a multicomponent system the background species $\sigma_B = \sum_\sigma \sigma/N$ if all point functions are orthogonal to 1.

For the purpose of interpretability, one should try to identify the background

component according to the characteristics of the system. The unweighted average component as the background is probably suitable when no particular component stands out among others in their physical property. There are situations where *a certain component may be naturally taken as the background*, such as vacancy in a multi-component alloy. In proteins, alanine is often used as a reference to calculate mutation energies. If we take σ_B =alanine, EIs in the expansion represent contributions to the energy of a sequence relative to poly-alanine. The point EI V_a^i corresponds to the point contribution of amino acid u_a at site i relative to alanine. This is the sequence independent portion of an alanine-mutation energy. The context-dependent effects are captured by higher-order terms. Similarly, V_{ab}^{ij} represents effective interaction between the two amino acids, point contributions subtracted. This is conceptually identical to a double mutant coupling energy, a measure well known to biochemists [95, 1].

Even though the physical model for a protein in solution often consists only of single-atom energies and pairwise atomic interactions, higher order contributions may arise if one integrates out some degrees of freedom. For example, when modeling molecular solvation, higher order interactions are necessary to accurately describe electrostatics as a function of conformational changes in the solute [80]. Similarly, integrating out side-chain degrees of freedom and expressing energy as a function of sequence can lead to higher order interactions between sequence variables, even if on the atomic level no more than pairwise interactions are present. As shown in Eq. 4.7, the CE formalism allows for arbitrarily high order interactions.

As we know the interaction distance cutoff depends on the size and length of the side-chain, i.e. amino acid species. Long side-chains can still see each other even when they are as far as 12 Å away (for details see the next chapter). It is desirable to be able to manipulate the expansion terms on an interaction $E(a@X, b@Y, \cdots)$, not an interaction site or cluster $E(X, Y, \cdots)$ basis. As explained below, a physically natural selection of point function basis is suitable because of the above concerns.

4.5 Delta functions as point basis set

When a natural background component exists among the allowed ones (say $\sigma_B = u_0$), we have $\phi_B(u_0) = 1$ and $\phi_B(u_a) = 0$ for $1 \le a \le M - 1$, i.e. $\phi_B(u_x) = \delta(x)$. The $(M - 1)$-dimensional functional space perpendicular to ϕ_B can be spanned by any $M - 1$ independent functions with no effects on the effective interactions. A particularly simple choice of $\{\phi_a\}$ is to use delta functions for all the non-constant point functions ($1 \le a \le M - 1$):

$$\phi_a(u_x) = \delta(x - a). \tag{4.18}$$

The point functions in Eq. 4.18, plus $\phi_B = \delta(x)$, are delta functions centered from 0 to $M - 1$ and hence orthogonal to each other. (but NOT to $\phi_0 = 1$). With delta functions as point basis the calculation of EIs from ECI in Eq. 4.12 becomes trivial:

$$V(\sigma^i = u_x, \cdots, \sigma^j = u_y) = \begin{cases} J_{x \cdots y}^{i \cdots j} & \text{if all of } x, \cdots, y > 0 \\ 0 & \text{if any of } x, \cdots, y = 0. \end{cases} \tag{4.19}$$

In other words, EI is equal to ECI unless the "non-interacting" background component is involved. This makes the discussions about the interpretability of ECIs in Section 4.4.3

4.5.1 Computational efficiency

Note that cluster function $\phi_a^i \cdots \phi_b^j$ vanishes unless $\sigma^i = u_a, \cdots, \sigma^j = u_b$. Given a sequence, at most one of the M^k CFs associated with a size-k cluster is non-zero. This leaves most of the elements of the correlation matrix (Eq. 4.8) to be zero. The larger the number of components, the sparser the matrix is. Therefore with delta functions we obtain a huge numerical advantage in that the correlation matrix is a sparse matrix, resulting in much less memory usage. Computer programs are able to utilize optimized sparse matrix subroutines to achieve faster speed. *Memory and speed* are the limiting factors in scaled-up applications of CE for large protein systems
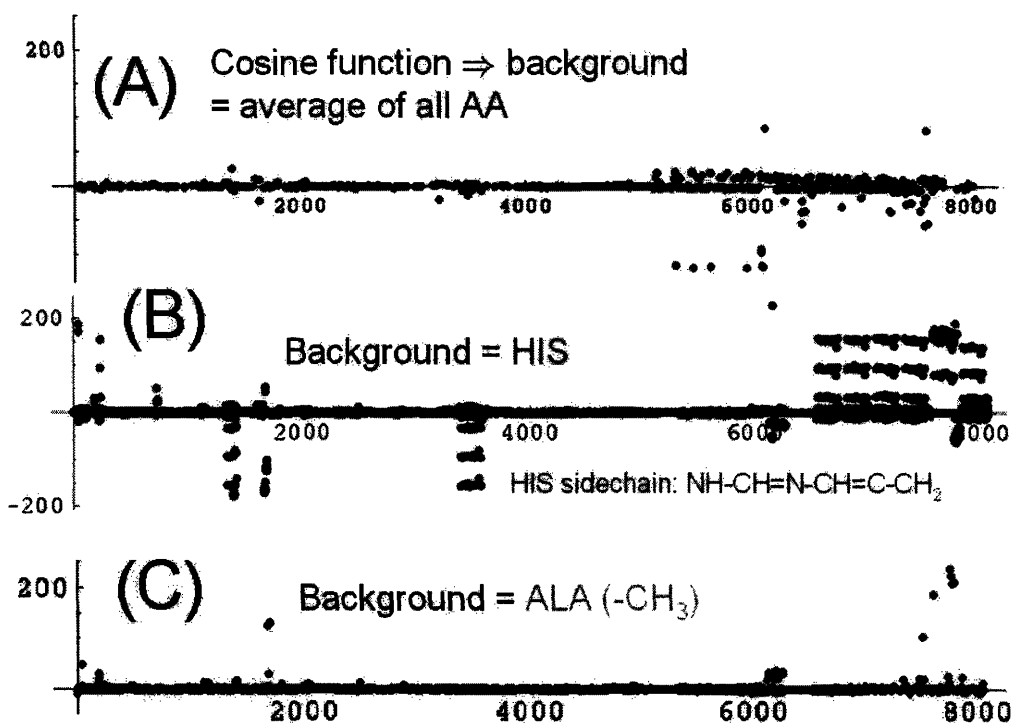
Figure 4-4: ECIs in kcal/mol of coiled coil with 16 AA species. The ECIs are mainly for pair interactions

*. Without delta functions the correlation matrix would be a dense one, and dealing with large proteins would be very difficult. To deal with a multi-component system like protein, we have found that delta functions in Eq. 4.18 are the only practical choice except for very short sequences.

4.5.2 Expansion convergence

To see the effects of different point functions on convergence, we present an example in Fig. 4-4. Long side-chain amino acids on certain sites cause steric collisions, resulting in a few large unphysical pair interactions ($\sim$ 100 kcal/mol), where the other pair interactions are of the normal magnitude ($\sim$ 1 kcal/mol) In panel A the ECIs are obtained with ϕ_a = cosine functions, which are orthogonal to each other and to $\phi_0 \equiv 1$, i.e. the background is the average of all 16 AAs. A lot of unphysically large negative effective interactions are obtained. In panel B the delta functions are used, though

---

* More details about scalability can be found in Section 4.7

a "bad" choice of background species, Histidine (with a long sidechain). By forcing Histidine not to interact with other AAs, we get even more unphysically large (positive and negative) EI's that cancel each other. In panel C we use the delta function as point basis set while the "correct" background Alanine is selected. Only few large interactions appear, and most EIs are of a reasonable magnitude of few kcal/mol. So unless an appropriate set of point functions (such as delta functions that vanish at alanine, panel C) is used, the few large pair interactions will be propagated to other terms, resulting in an unnecessary increase in the number of ECIs to include, i.e. slow convergence.

4.6 Cluster expansion construction

4.6.1 ECI Calculation

The essence of a successful cluster expansion is to truncate the series in Eq. 4.7 and construct an 'optimal' set of clusters at which the variables interact considerably.

There are several techniques to determine ECI for a truncated expansion [186]. A common one is the Structure Inversion Method (SIM) or the Collony-Williams [40] method, in which ECI are fitted to the calculated or experimental energy of a relatively small number of sequences. A least-square (LS) fit for the linear system (Eq. 4.9)

$$\vec{f} = M\vec{J}$$

minimizes the root mean square (RMS) of the fitting error

$$RMS^2 = \langle (E_i - E_i^P)^2 \rangle, \tag{4.20}$$

through matrix (pseudo)inversion

$$\vec{J} = M^{-1} \cdot \vec{f} \equiv (M^T M)^{-1} M^T \cdot \vec{f}, \tag{4.21}$$

where E_i and E_i^P are the true and predicted or fitted energies for sequence i, respec-

tively.

In order to more accurately fit the important low energies, the fitting can be weighted by $\max(e^{-(E-E_0)/K}, w_0)$, where E_0 is the lowest energy in the data set, K is approximately the range of interest above E_0 and w_0 is the minimal weight at large E to avoid numeric instability.

4.6.2 Cross validation

Given $E_{\min}$ for enough sequences, Js can be extracted by standard fitting procedures (Eq. 4.21). Determining which Js to keep in the fit is not always obvious. While one may be guided by the idea that point terms are larger than pairs, which in turn are larger than triplets, this is not always true. Here we employ the *cross-validation* (CV) as a systematic and quantitative measure of the quality of the CE model. Essentially, the CV score is the average error with which each sequence is predicted when left out of the fitting, and as such is a good measure of the prediction power. Our procedure consists of fitting a selected set of candidate clusters and ordering them by the average $|J|$. Clusters for which the J value largely arises from numerical noise increase the CV score, and are excluded. The so-called "leave-one-out" formula [186]

$$CV^2 = \langle (E_i - \hat{E}_i^P)^2 \rangle, \qquad (4.22)$$

where $\hat{E}_i^P$ is the predicted value of sequence i obtained from a least-squares fit to all other sequences. By optimizing CV with cluster construction we can steadily test and improve the EI approach to protein energetics.

4.6.3 Statistics-based analysis

As exemplified in Fig. 4-4 and explained in more detail in the next chapter, the ECIs in a typical protein can be envisioned as dominated by terms up to the pair level with potentially just a few high-order contributions from long side-chain amino acids. It is possible to pinpoint these high-order terms based on the characteristics of the poorly fitted sequences $\mathcal{A} : |\Delta E| > \Delta E_{\max}$ in a pair-level fitting to locate important
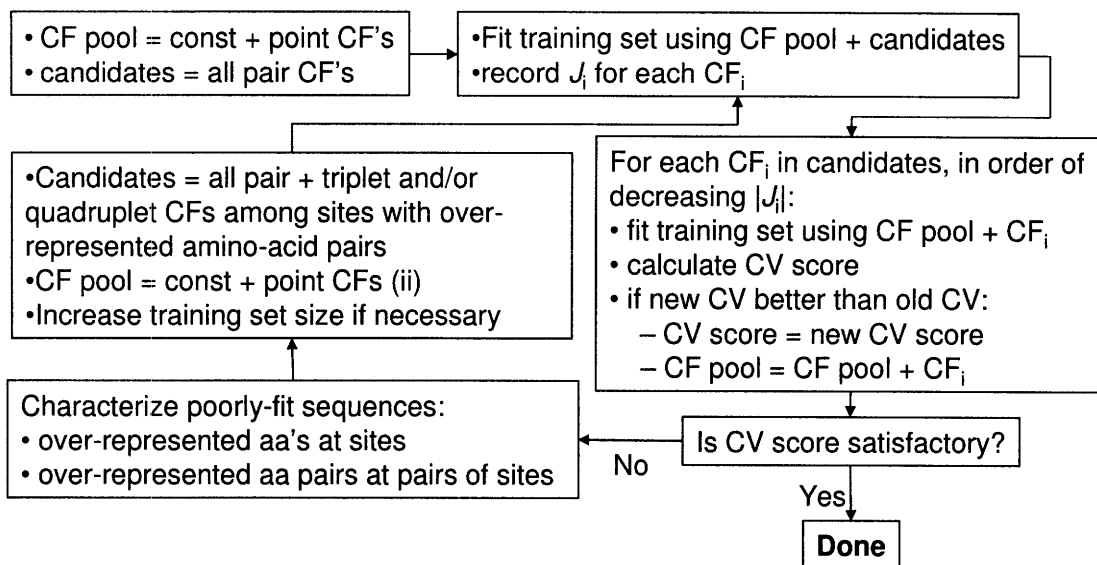
Figure 4-5: The procedure for fitting a cluster expansion.

multi-body terms ($\Delta E_{\max}$ is a relatively large error). We calculate the information content

$$I^i = \ln m^i - S(p(\sigma^i|\mathcal{A})) \tag{4.23}$$

$$I^{ij} = \ln m^i m^j - S(p(\sigma^i \sigma^j|\mathcal{A})) - I^i - I^j$$

for each site i and for each pair $\{i, j\}$ out of the AA distribution in $\mathcal{A}$ ($S(p) = -\sum_{\{p\}} p \ln p$ denotes entropy). Although it is possible in principle to calculate three-body information content and above, the set $\mathcal{A}$ is usually quite limited and does not provide enough sequences for reliable statistics when we go to high order. Nonetheless, with Eq. 4.23 we are usually able to identify only a few high-information terms. Then multi-body cluster functions associated with the these sites and the high-frequency AAs in $\mathcal{A}$ can be easily tested.

4.6.4 Heuristic construction method

A heuristic cluster construction procedure appropriate for proteins is illustrated in Fig.4-5. The selection criteria is based on minimization of the CV score. It incorporates the constant and all point terms, and implicit favors pair clusters. As for

109

higher order cluster functions, the statistics based method is used to pinpoint a limited number of those. Additional sequences may have to be computed when we follow this algorithm in order to reach higher prediction accuracy. This procedure will be used in the next chapter.

4.7 Scalability

Scalability is an important issue for the application of cluster expansion. For general fold recognition applications (e.g. recognition of 100–200 residues domains), it will be important to be able to expand larger systems. We can estimate the scalability as follows. If the type of clusters to be included is pre-determined (e.g. capped at triplets), then the number of parameters that need to be fit will scale linearly with the size of the protein. This is because strong interactions in proteins are of finite range, and it is never necessary to include all pairs or triplets, just those that interact within the range. The number of residues that interact with a given site, in a large protein, is approximately constant.

Currently the CV score is obtained through matrix inversion, and this limits the size of tractable problems N by the amount of memory available and the time it takes to invert a matrix N^3. When N is very large, the CV score can be approximately obtained with fast linear-equation solving for sparse matrices. On the other hand, energy evaluation for large proteins and enough training sequences can be extremely expensive, but that is not a limitation of CE.

4.8 Conclusion

In summary, we have established the framework of cluster expansion suitable for protein energetics. The nature of the amino acid interactions enables us to use appropriate basis functions to quickly expand their energy. In the next chapter we show concrete examples for the cluster expansion's applications in proteins.

Chapter 5

Cluster Expansion for Protein: Applications

5.1 Introduction

In the previous chapter we have established the theoretical framework of cluster expansion for protein. In this chapter, cluster expansion is applied to explore the energetics of specific proteins. Not surprisingly, the CE method does not work well when the protein fold is allowed to change. When the backbone is fixed, very accurate energy prediction can be achieved. This is demonstrated in the case studies of three small protein folds - the α-helical coiled coil, the zinc-finger and the WW domain. The three case studies show that protein sequence can be mapped directly to energy using surprisingly simple expansion terms with high accuracy. The expansion retains, and in certain ways enhances, physical interpretability of the underlying energy models. The promising results on these small systems suggest that the method may have be useful for larger molecules.

As the first "real-world" application of CE for proteins, we use cluster expansion to expedite design of an α-helix sequence that optimally binds to the surface of a carbon-nanotube (CNT).
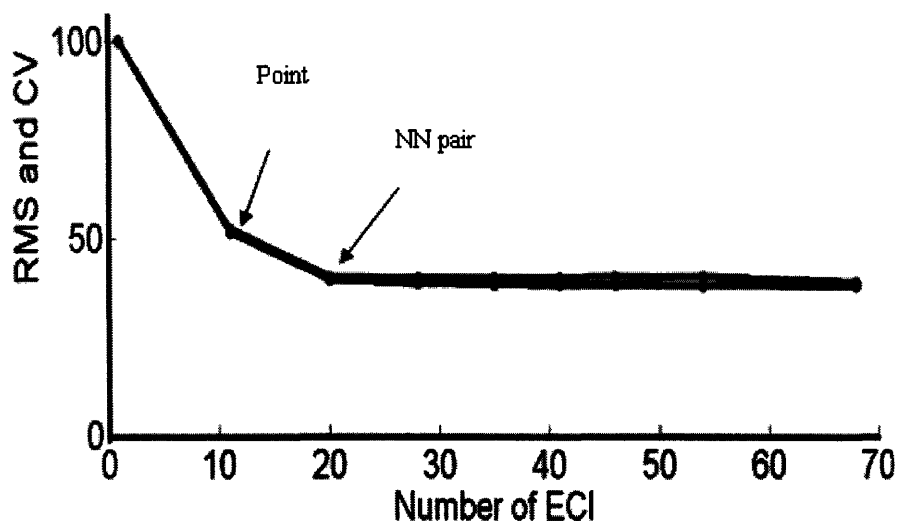
Figure 5-1: RMS and CV scores versus the number of clusters included for RISMC energies.

5.2 Case study with flexible protein backbone

The analogy between the alloy host lattice and the protein backbone (Fig 4-3) implies the importance of the two as the central stage for cluster expansion. In alloy applications the problem of whether to relax the lattice is not trivial. Usually if there is only small atomic displacement without change of lattice topology after relaxation (as is the case in chapter 3), it is possible to obtain an accurate expansion. In this section we show the results of cluster expansion in a case where the protein backbone is allowed to change.

Monte Carlo simulations were carried out with the Rotational Isomeric State Monte Carlo (RISMC) code [108]. To avoid the enormous complexity of simulating whole peptides, the code ignore all degrees of freedom but the torsional angle of the amide bond, which connects neighboring amino acid backbone atoms C and N and solely determines the backbone conformation. The energy model is based on amide bond torsional energy, electrostatics and simple volume exclusion penalty to represent vdW interactions.

We tested various binary combinations of amino acid species: polar + non-polar, hydrophobic + hydrophibic, acidic + acidic, etc. Sequence lengths from 6, 8, to 10

were tested. Because we were doing binary cluster expansion and the scale of the problem is not too large, all possible sequences and all clusters, without imposing any symmetry constraint, were used for the expansion. The cross validation scores typically drop to 50-80% at the minimum, and sometimes become larger than 1. Here we show a relatively good fitting example in Fig.5-1, where the main prediction power is from point terms with very small contributions from nearest-neighbor pair interactions. The reason of the poor predictive power of cluster expansion when the backbones of peptides are free to move is that different sequences relax to different structures and it is almost impossible to find a set of simple effective interactions that apply to any structure. By fixing the backbone and optimizing the sidechain degrees of freedom, much better results can be obtained. Therefore all other protein CE results presented in this thesis are based on the fixed backbone approach except in the above example.

5.3   Case studies with fixed protein backbone

This section demonstrates concrete results obtained with the cluster expansion method on protein sequences. We show the power of the CE by testing it on three different protein folds, mimicking the protein design problem. The three case studies are the α-helical coiled coil, the zinc finger and the WW domain [218, 75]. These proteins systems represent relatively small, while significant in functionality, segments of larger molecules. By focusing on smaller problems, we are able to refine and adjust the implementation of CE to better describe the nature of protein interactions. For each system, we show that CE can derive useful yet highly simplified energy expressions. In the case of the coiled coil, where experimental data is available, good agreement between calculated effective interaction parameters and measured coupling energies is found.

5.3.1 Energy evaluation methods

The folding energy is defined as the energy difference between the folded and the unfolded protein states: $E^{\text{folding}} = E^{\text{folded}} - E^{\text{unfolded}}$. Although the CE can in principle be used with any energy model, we test it here with a physically meaningful but relatively simple expression similar to Hamiltonians commonly used in the design field [4]:

$$E[\vec{\sigma}, \vec{\tau}] = E^{\text{vdW}} + E^{\text{elec,wat}} + E^{\text{solv,sc}} + E^{\text{torsion}}, \tag{5.1}$$

where E^{vdW} is the van der Waals energy (with atomic radii scaled to 90%) modeled as a 6-12 Lennard-Jones potential, $E^{\text{elec,wat}}$ is the total electrostatic energy given by a distance-dependent dielectric model (excluding intra-sidechain interactions), $E^{\text{solv,sc}}$ is the solvation energy of all backbone and sidechain atoms [101], and E^{torsion} is the dihedral angle torsional energy. All energy terms are calculated using the CHARMM package [25] with parameter set 19. Rotamers were taken from the Dunbrack 2002 rotamer library [62]. For the zinc finger and WW domain, the unfolded state is modeled by retaining only sidechain self energies and local interactions between sidechains and their surrounding penta-peptide backbone. For the coiled-coil system, the unfolded state is modeled according to an energy model previously shown to perform well in recognizing coiled-coil dimerization preferences (model HP/S) [74].

Because $E[\vec{\sigma}, \vec{\tau}]$ in Eq. 5.1 is pairwise-decomposable, we are able to apply the dead-end-elimination (DEE) algorithm [56, 71] followed by a branch-and-bound search to arrive at the optimal rotamers corresponding to $E^{\text{folding}}_{\text{repack}}$. Thus, in a CE derived from these $E^{\text{folding}}_{\text{repack}}$, the Js, and hence EIs, parameterize optimized energies whereby all the side-chain degrees of freedom are coarse-grained out. The EI, defined at the sequence level, may include higher order terms even though the initial energy expressions at the conformational level are pairwise decomposable. The advantage of this procedure is an enormous reduction in the search space, from $(20m)^L$ to m^L, where 20 is the average number of rotamers considered per AA.

A main disadvantage of the above energy calculation method is that the opti-

mized state is selected directly from those of discretized protein conformations and steric sidechain–backbone clash may be present even in the optimized state. Steric clashes result in large $E_{\text{repack}}^{\text{folding}}$ due to unphysical vdW interactions. The following three case studies show that even after careful selection of the backbone and the side-chain rotamers, steric clashes still occur. To remove this artifact and to better treat electrostatics, the solutions obtained with DEE then undergo continuous side-chain minimization in CHARMM (ten cycles of steepest-descent minimization and ten cycles of adopted-basis Newton-Raphson minimization). The resulting structures are evaluated with an energy function in which 100% radii are used for van der Waals calculations, and screening of electrostatic interactions is modeled using the Generalized Born model with "perfect" Born radii [129] computed using the program PEP [22]. The calculated energy $E_{\text{min,GB}}^{\text{folding}}$ is therefore no longer based on a pair interaction model.

In the following three case studies, both direct repacking and continuous minimization methods are used. The energy calculations were performed by our collaborators [75].

5.3.2 Coiled coil
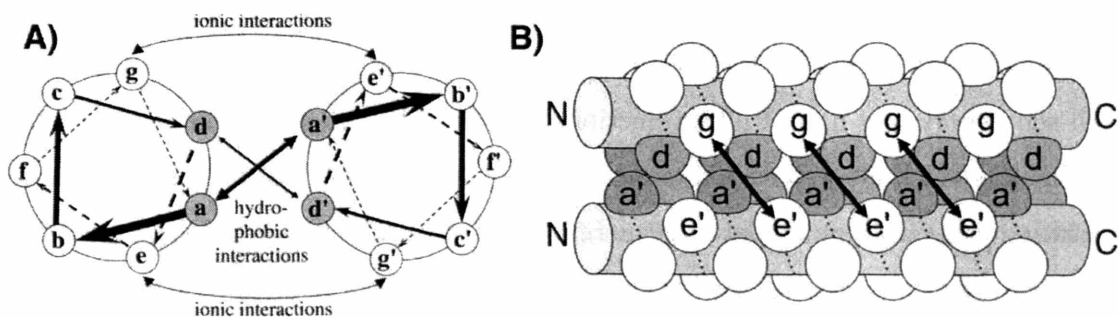
Figure 5-2: Schematic of a parallel dimeric coiled coil. A) axis view. Opposing a and d residues interact in the core while opposing e and g residues frequently participate in electrostatic interactions. B) Cartoon representation viewed from the side. Residues are represented as spheres.

Our first case study involves the coiled coil, a common and well-characterized protein interaction interface (see Fig. 5-2). An ideal coiled-coil backbone possesses a
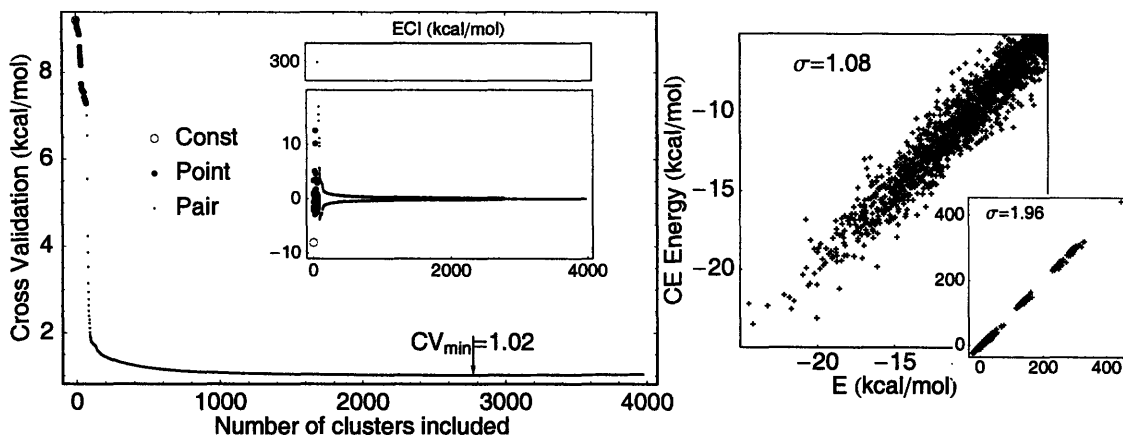
Figure 5-3: Cluster expansion of coiled-coil with direct repacking calculations. Left: evolution of the CV score when more cluster functions are included, and the inset shows the ECIs. Right: Comparison of CE predicted versus input energies.

screw axis (in the middle of Fig. 5-2A) with a repeating unit every 7 residues (see Fig. 5-2B) as well as C2 symmetry about the coil axis [42]. To derive a scoring function for coiled coils of arbitrary length, we expanded the energetics of a repeating structural element (unit cell). We postulated that interactions between amino acids more than one heptad apart in a coiled coil would not be appreciable and so did not include clusters corresponding to these interactions in the CE. The unit cell was chosen to be a two-heptad dimeric parallel coiled coil (see Fig. 5-2). Additionally, to avoid edge effects, we used a periodic boundary condition for the backbone structure and sequence. CE was fit to just the energy of the central unit cell (all of the unit cell self energy and half of all interactions between the unit cell and the rest of the molecule), which allowed each interaction type to be counted exactly once. Thus the resulting ECI map exactly onto the energies of the corresponding interactions and can be applied to non-periodic sequences. Only 4 sites in each heptad are each modeled as one of 16 selected AA species (the 3 remaining sites are set to Ala). These 4 sites have been shown, in many cases, to be sufficient to determine coiled-coil dimerization preferences and other properties [130, 66].

We first show the results with the direct repacking energies in Fig. 5-3. The training set consists of $\sim 30,000$ randomly chosen sequences weighted by $\max(e^{-(E+26)/120}, 0.01)$. Truncating the CE at the pair level is sufficient to accurately reproduce the energet-

116

ics of the system. The structural symmetry reduces all 137 clusters up to pairs to 1 constant, 4 point and 36 pair-level independent cluster (7741 independent Js). We are therefore able to include all of them as candidate clusters in the fitting. Fig. 5-3 left panel shows the weighted CV scores of the least square fitting versus the number of included cluster functions, following the algorithm described in Fig. 4-5. The CV score reaches a minimum at 2770 clusters functions, and fluctuates (mostly increases) slightly afterwards. We thus come to an "optimal" set of 2770 ECIs for energy prediction. The most significant ECIs are found to correspond to residues that mediate contacts between different helices, in agreement with biologists' intuition about the system.

To test the predictive character of the CE we compare its prediction for 3995 random sequences not included in training to the directly calculated energy (Fig. 5-3 right panel). The unweighted RMS error is 1.96 kcal/mol for all energies and 1.08 kcal/mol for -26$< E_{\min} <$-5 kcal/mol. The error is sufficiently small for such applications as sequence optimization, and is comparable with the accuracy of the underlying energy model. We trade such a small error for being able to predict the optimal energy of any sequence by summation of EIs for 41 clusters, as opposed to performing global optimization over an average of 5.9×10^{55} rotamers. Even compared to the highly efficient DEE method for sidechain positioning, the time to calculate $E_{\min}$ of a sequence is reduced from $\sim$200 sec to $\sim$1 μs with our coarse-grained Hamiltonian, a 2×10^8-fold acceleration.

The CE can work with any energy model, and when the model is refined, the CE results are also improved. Fig. 5-4 displays the results with continuous minimization on the sidechains, which alleviate the problem of steric clashes and unphysical vdW interactions that appear as large energies in Fig. 5-4's right panel. Comparison of the ECIs in both figures clearly indicate that the minimization procedure can remove the sidechain clashes.

Fig. 5-5 shows the agreement between experimentally measured g-e'+ coupling energies [95] (the prime designates the opposite strand and + sign indicates the next heptad) and the corresponding pair ECIs. The right panel is for ECIs after
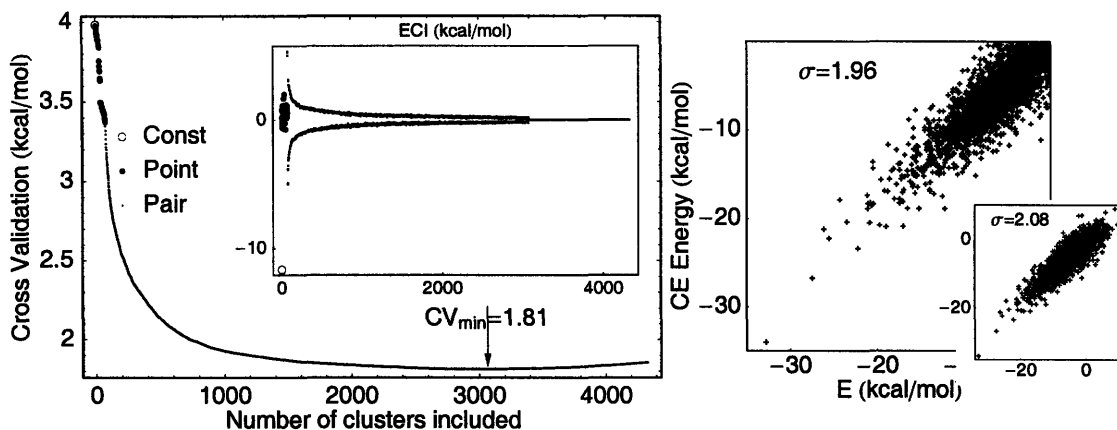
Figure 5-4: Cluster expansion of coiled-coil with continuous minimization.
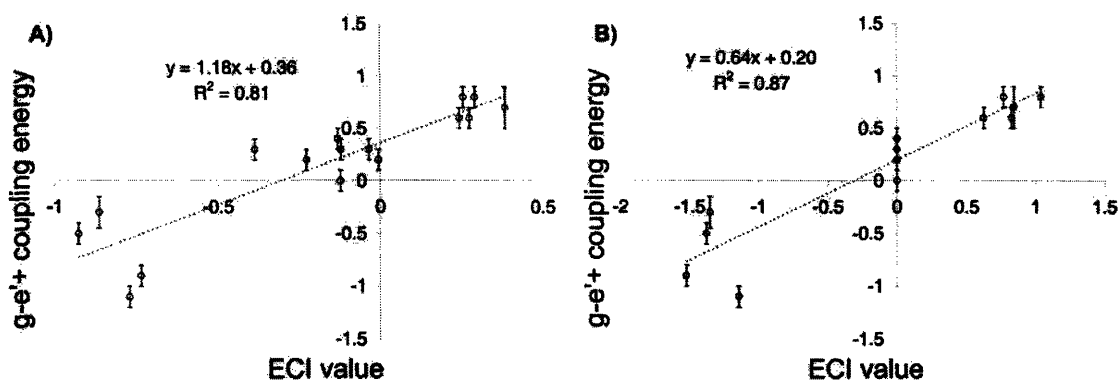
Figure 5-5: Agreement between experimentally measured double-alanine coupling energies for residues E, Q, R and K at g-e+ [95] and corresponding pair ECI from the cluster expansion (in kcal/mol). The energies are from (A) from direct repacking (B) continuous minimization.

sidechain relaxation and shows better agreement with experiment than the left panel without relaxation. The agreement illustrates the physical interpretability of the cluster expansion.

5.3.3 Zinc finger

As a second application we consider the zinc-finger, a common DNA-binding fold (Fig. 5-6). The backbone of Zif268 (PDB ID 1ZAA) residues 33-60 is used as a model Zn-finger structure. Following Mayo et al. [43], we consider a sequence space in which 2 sites are fixed, 1 site has 7 candidate species, 18 sites have 10 and the other 7 sites have 16. The training set consists of $\sim 60,000$ random sequences. Again following the
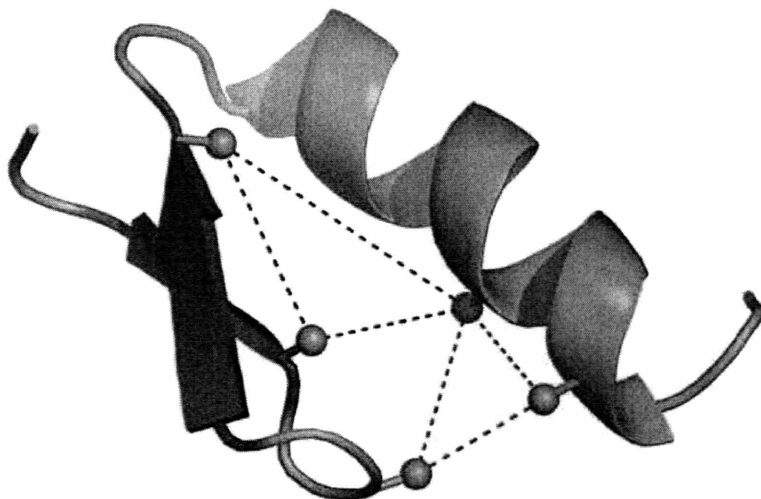
Figure 5-6: A ribbons representation of the zinc-finger structure with important triplet clusters for the cluster expansion model of zinc-finger. Orange balls represent the location of the C_α atoms of side chains. Two clusters are shown, one in red and one in blue.

statistics-based method in section 4.6.3, we have identified some triplet and quadruplet clusters functions among certain sites * that are almost exclusively occupied by aromatic sidechains W, H, Y and F. Such clusters represent close-range interactions of bulky residues. This demonstrates the existence of complex correlations in a globular protein, and the CE provides a systematic, quantitative way of identifying such correlated sites.

Fig. 5-7 shows the fitting and prediction results for zinc finger. Larger prediction errors (~ 2 kcal/mol) than in coiled coil are obtained with sequences in the low energy regime, indicating that the more complicated geometry of the domain may make the energy a more complex function of sequence. Such error is traded for a remarkable reduction in search space: from 1.4×10^{60} to 1.9×10^{27} states.

5.3.4 WW domain

The WW domain is a protein-protein interaction motif composed of 35-40 residues. It forms the smallest known independently stable triple-stranded antiparallel β-sheet. A defining feature of this motif, from which its name is derived, is the presence of

---

* Some of these sites are indicated in Fig. 5-6.
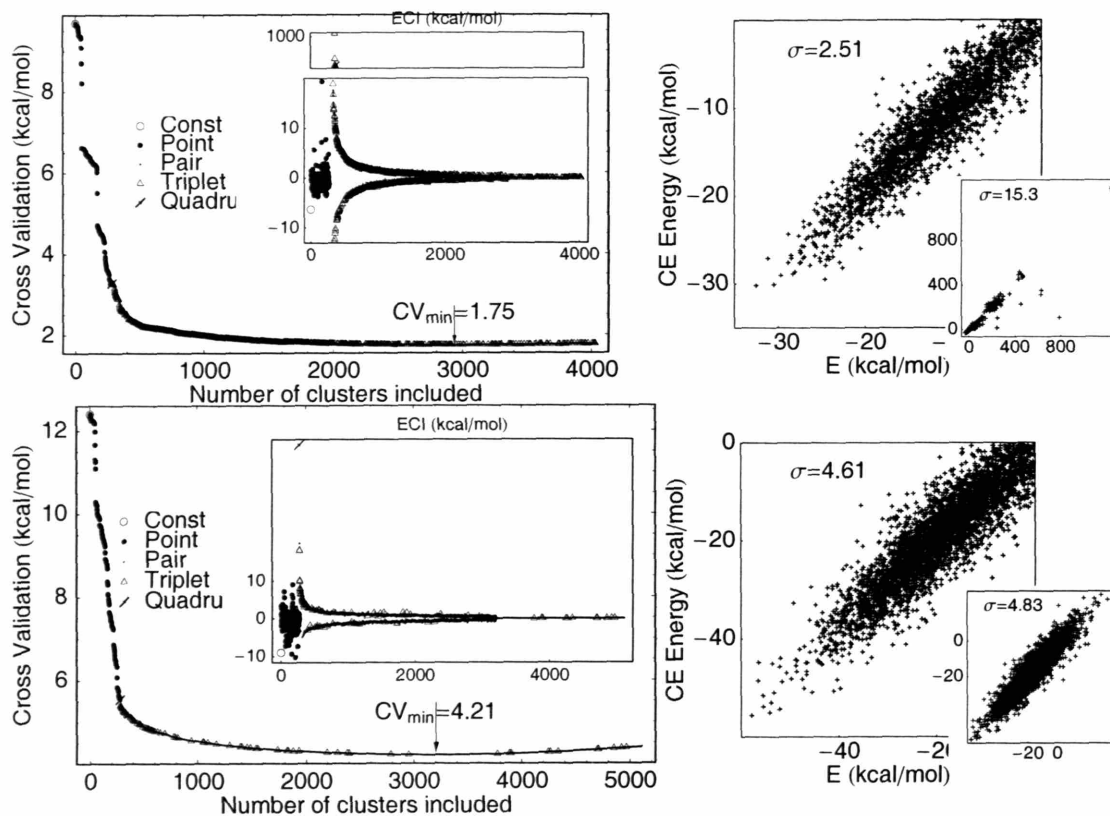
Figure 5-7: Cluster expansion of zinc-finger energies from direct repacking (upper panel) and from sidechain relaxation (lower panel).

Figure 5-8: Structure of the WW domain with important higher-order clusters. Orange balls represent the location of the C_β atoms of side chains. A) A structurally compact cluster corresponding to short-range interactions. B) A more disperse cluster arising from long-range electrostatic interactions. C) Quadruplet cluster with many contributing cluster functions corresponding to a wide range of amino-acid types.
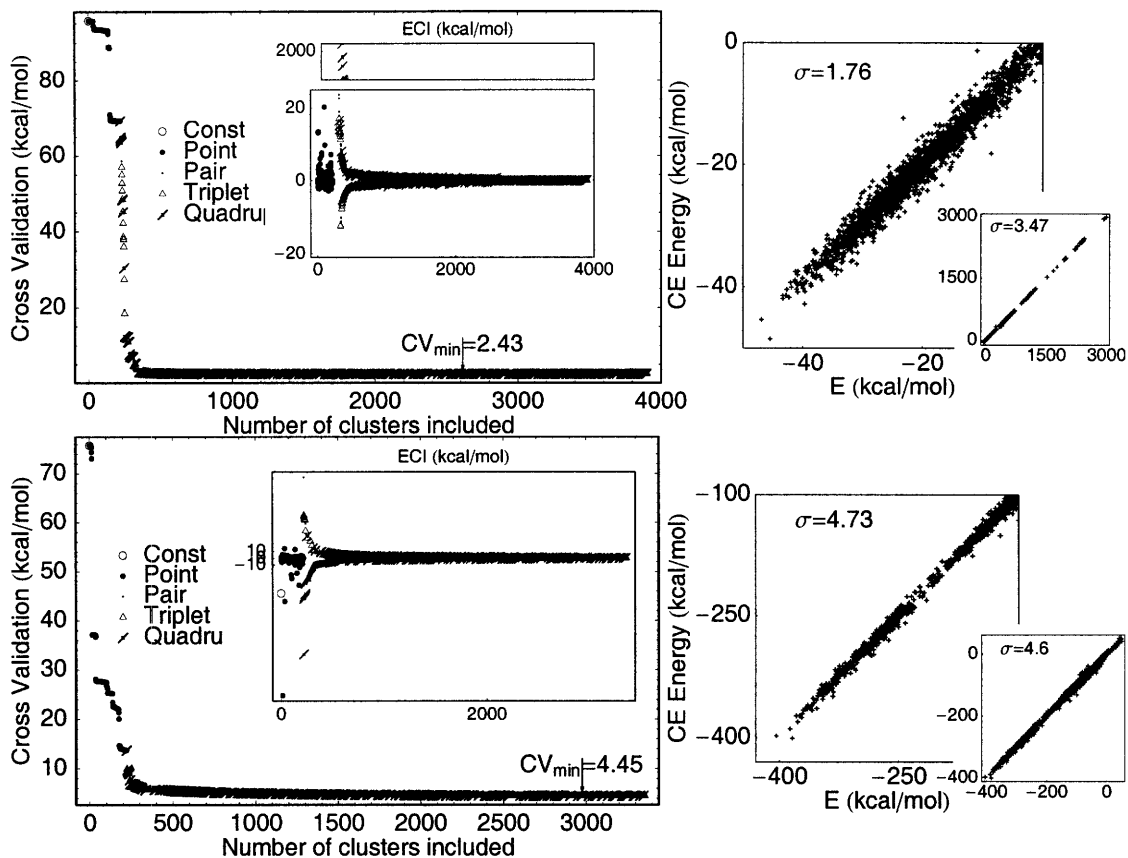
Figure 5-9: Cluster expansion of WW domain energies from direct repacking (upper panel) and from sidechain relaxation (lower panel).

two tryptophans spaced 20-22 residues apart. We used the structure of human PIN1 WW domain (PDB ID 1PIN) to define backbone coordinates. The choices of amino acids at each position covered at least 90% of all naturally occurring residues.

We explicitly computed structures for $\sim 42,700$ random sequences and estimated their folding energies. Fig. 5-9 shows the fitting and prediction of expanding repacked energy (upper panel) and the energy with minimization (lower panel) for the WW domain. Similar to the Zn finger, we found that higher order terms were necessary for good agreement. However, unlike for the zinc finger, for the WW domain there are two types of triplet clusters. One consists of structurally compact sites, and CFs arising from these clusters are mostly positive and correspond to large amino acids (see Fig. 5-8a for an example). In the other, sites are more structurally dispersed and combinations of residues producing significant CFs consist mostly of charged and

System	Relaxation	Number of optimized clusters/cluster functions					Prediction error	
		constant	point	pair	triplet	quadruplet	All	low-E
Coiled coil	no	1/1	4/72	36/2697			1.96	1.08
	yes	1/1	4/72	36/2992			2.08	1.96
Zinc finger	no	1/1	26/273	86/1727	20/934	1/1	15.25	2.51
	yes	1/1	26/273	78/2851	3/78	1/1	4.83	4.61
WW domain	no	1/1	32/219	43/474	12/1514	1/406	3.47	1.76
	yes	1/1	32/219	42/770	11/1581	1/405	4.6	4.73

Table 5.1: Summary of the three cluster expansion case studies.

polar amino acids (see Fig. 5-8b). These two types of clusters roughly correspond to the two main classes of interactions we model: van der Waals (short-range) and electrostatics (which can be long-range). Additionally, there is one quadruplet cluster that seems to be important for overall accuracy. It is shown in Fig. 5-8c. The set of amino acids at this cluster that give rise to large CFs is diverse and it does not have a clear structural or energetic interpretation. The error of the fit, ~ 4 kcal/mol, is higher than before but, considering the energy range of over 300 kcal/mol, this is sufficiently accurate to be very useful. Here CE reduces a sequence-structure space of 2.6×10^{65} to 1.1×10^{27} sequences.

5.3.5 Discussions and conclusions

Table 5.1 summarizes the three case studies. An obvious trend is that as the 3d structure of the protein becomes more complex, so do the effective interactions. A trend seen in all three systems is that the accuracy of the CE fit is worse after relaxing the sidechains and evaluating the energy with a non-pairwise decomposable energy function. This indicates that the energy resulting from this procedure is a more complicated function of sequence. Most importantly, fairly accurate energy prediction is achieved, resulting in enormous speedup.

In summary, we have demonstrated how the energetics of a protein with pre-

defined backbone can be coarse-grained to a function of sequence only with the cluster expansion method. We have successfully applied the method to three distinct families of proteins, and found that interactions are closely related to the protein structures. The accuracy of the CE predictions, which can be systematically improved, implies that this much simpler expression can be used in place of traditional Hamiltonians, dramatically improving computational efficiency.

5.4 Nanotube-binding peptide

5.4.1 Introduction

Since their discovery in 1991, carbon nanotubes (CNTs) have been associated with a dazzling array of applications [60, 24, 175, 19], e.g.

- as field emitters in (among others) lighting elements and flat panel displays,

- for x-ray production for biological applications such as endoscopy,

- as supercapacitors for energy storage and electro-mechanical actuators (akin to piezoelectrics),

- as additives in storage devices for hydrogen and lithium in fuel cells and batteries,

- as strengtheners in polymer composites for space and other advanced materials,

- as conductive fillers for charge dissipation in automotive parts and gas lines,

- as probes and sensors in the study of biochemical reactions, and

- as tips in atomic force microscopes and scanning tunneling microscopes

In these applications, CNTs offer improvements over more conventional materials, e.g. more moderate vacuum in field emission, more precise energy range in x-ray production, lower operating voltage as electro-mechanical actuators, lower weight as fillers, etc. In addition, the possibilities for applications such as DNA sequencing

[107], as nanoscale field-effect transistors, and as thermoelectric materials are under active investigation.

The richness of applications owe its origin to the wide range of electronic, thermal, thermoelectric, magneto-electric, optical and mechanical properties displayed by CNTs. These properties are in turn influenced by each CNT's structure and topology — most notably, whether it is single-walled or multi-walled, its chirality (n, m) (the existence and "pitch" of the helices of carbon hexagons around the tube), and its diameter d. A prime example of how structure and topology affect CNT properties and hence function is electrical conductivity, the versatility of which underlies much of the utility of CNTs. One of the first computational studies of CNTs [76] established that single-walled CNTs can be either metallic $(n - m =$ multiples of 3) or semiconducting with a band gap inversely proportional to the diameter d. Applications such as the fabrication of a diode by grafting of a metallic nanotube to a semiconducting one depends on the ability to select CNTs by chirality and diameter.

Mechanically, chiral nanotubes (those with helices of carbon hexagons) are found to respond differently to torsion in opposite directions, in contrast with non-chiral ones [183]. The Young modulus and tensile strength of a CNT also depends on its diameter. Further, the band gap of single-walled tubes changes with uniaxial strain, so the diameter affects electro-mechanical behavior of CNTs as well. In such applications the ability to select CNTs based on chirality (which affects conductivity and band gap) as well as diameter (which affects how band gap changes with mechanical environment) is vital.

5.4.2 Carbon nanotube separation

To fully use the unique potential of CNTs in applications one either needs to produce them with high selectivity for diameter and chirality, or separate them on the basis of structure after synthesis. Most synthesis techniques (e.g. carbon arc discharge [64] or laser ablation [182]) produce bundles or "ropes" of carbon nanotubes in various chiralities and diameters. The separation of nanotubes in such ropes by desired properties is an important first step towards fabrication of useful devices and manufacturing of

composites from CNTs. Such separation is complicated by the dispersion in chirality and diameter in these bundles and the poor solubility of CNTs in both aqueous and non-aqueous solutions.

Separation by chirality

It was shown that DNA can be used to separate metallic and semi-conducting carbon nanotubes through ion-exchange chromatography [215, 214]. Peptides are also known for their selective affinity to CNTs [198, 57, 45]. In both cases the bio-molecules non-covalently bind to the surface of CNTs and help their dispersion in solution. Using peptides is advantageous because they are less expensive than DNA and, with flexible structural conformations and 20 types of amino acids compared to 4 types of nucleic acids for DNA, offer more design options. However, the vast design options for peptides requires the use of computational modeling methods to rapidly search through all the possible amino acid sequences for specific target properties. In this section the desired property (e.g. binding energy to a particular CNT) is the target of cluster expansion

An optimal peptide sequence for the purpose of carbon nanotube separation satisfies several design criteria: Strong binding energy to CNTs, weak binding to the same sequences (so that they do not coagulate in solution), strong preference for unique pitch when it binds around the CNT (so that the CNT + peptide hybrid has constant charge density). The latter property is important for the separation of the hybrids by ion-exchange chromatography.

5.4.3   Method

Target properties

The space of possible peptides is very large. Even for a short peptide with 10 amino acids there are 20^{10} possible sequences. While some screening can be done with intuitive rules, it is not possible to capture all the possible sequences and complex interactions between peptides and CNTs through intuition, making computational
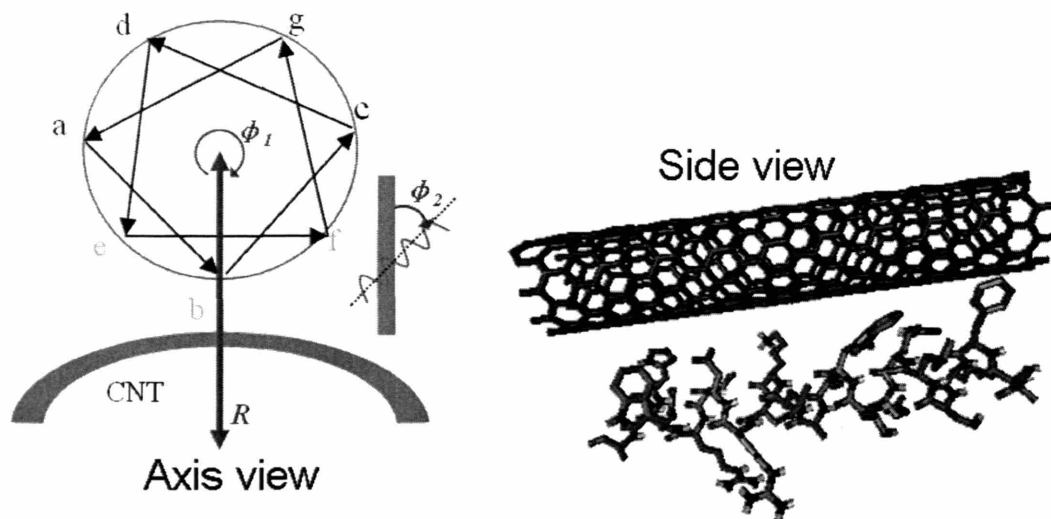
Figure 5-10: Schematic of α-helix binding to a nanotube. Parameters R, ϕ_1, and ϕ_2 determines how close the peptide and the nanotube are, which the amino acid comes into contact with the CNT, and the binding pitch, respectively.

screening so valuable. Such a situation is ideal for cluster expansion, which is capable of coarse-graining the protein energetics and evaluating sequences extremely fast. According to the design requirements, we need to expand the CNT binding energy E_{bind} as well as the sequence dimerization energy E_{dimer}.

Binding motif

Two important questions arise when we attempt to apply the established cluster expansion method to design *in silico* nanotube-binding peptide sequences that disperse and separate CNTs.

1. What is an appropriate binding motif that is relatively easy to manipulate with modeling methods while providing enough design flexibility?

2. Since the peptide solvation effect, i.e. the difference between hydrophobic and hydrophilic amino acids, is a major factor determining the binding strength, how do we model it accurately?

We have found from the literature [57, 45] that the amphiphilic α-helical peptides are especially promising. The relatively simple α-helix is very well studied and un-

126

derstood by biologists and ample computer algorithms that model its stability exist (e.g. AGADIR [3], PROF [148] and SSPRO [172]). As shown in Fig.5-10, one side of the amphiphilic α-helix is dominated by hydrophilic amino acids (sites a, c, d, g), providing solubility in aqueous solution. The other side (sites b, e, f) is mostly occupied by hydrophobic amino acids that prefer to stay close to the hydrophobic CNT. Three important parameters that determine the peptide-nanotube binding geometry are also labeled in Fig.5-10, namely the peptide-CNT distance R, the angle at which the closest amino acid faces the CNT ϕ_1, and the angle between the axes of the peptide and the CNT ϕ_2.

An additional advantage of the α-helix motif is that we are able to reuse results obtained in the coiled coiled system (section 5.3.2) to approximate the motif self-binding energy E_{dimer}. This is because the two strands in the coiled coil are both α-helices.

Energy model

The second question is the energy model to use. Because the hydrophobic effect is a main driving force for the peptide to bind to CNT, accurate modeling of the solvation energy is crucial in our modeling. We have applied the same energy evaluation procedure as performed in the peptide case studies. After rotamer repacking and structural minimization of side-chains, non-pairwisable electrostatics calculation are performed with the Generalized Born/surface area (GB/SA) method [174]:

$$\Delta G^{\text{solvation}} = \sum_i \left(\Delta E_i^{\text{self}} - \frac{1}{2} \sum_{i \neq j} E_{ij}^{\text{screening}} + \Delta E_i^{\text{nonpolar}} \right) \qquad (5.2)$$

where $\Delta E_i^{\text{nonpolar}}$ is a surface area-dependent approximation to the hydrophobic solvation term, the screening term $E_{ij}^{\text{screening}}$ depends on the structural environment of the surrounding groups and E_i^{self} is obtained using an approximation to the integral of the energy density of the electric field over space.
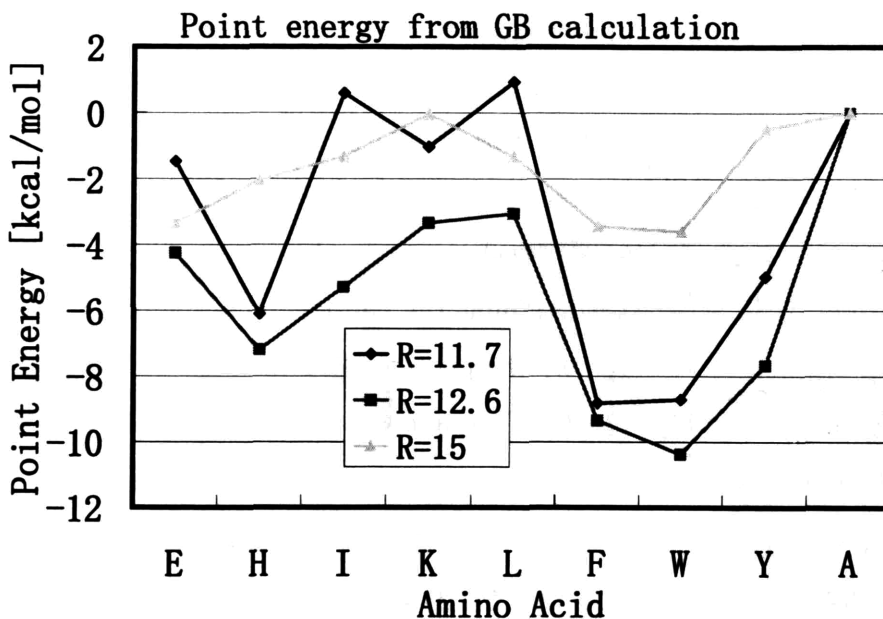
Figure 5-11: Point binding energy of different amino acids with the Generalized Born (GB) solvation model at R=11.7, 12.6 and 15 Å, respectively

Point energy and amino acid space

To determine which of the amino acids occupy the "binding surface" and which are in contact with the solution and therefore should be hydrophilic, we start from a helix conformation with reasonable $\phi_{1,2}$ and cluster expand the binding energy:

$$E_{bind} = E(\text{peptide} + \text{CNT}) - E(\text{peptide}) - E(\text{CNT}) \qquad (5.3)$$

Not surprisingly, the leading ECIs are point energies, which can be interpreted as an indication of amino acid affinity to CNT. The calculated energies are not very sensitive to the type of CNT as long as the radius of the tube is not too small, i.e. the CNTs all act like a graphene sheet when peptides bind to it.

Fig. 5-11 lists the point energy on site b (the site closest to CNT in Fig.5-10) calculated with the GB/SA model for a (10,0) CNT. Since alanine is the reference amino acid, its binding energy is always zero. All AAs has the smallest binding energy when CNT-peptide distance R=12.6 Å. AAs with aromatic sidechains H, F, W and Y have the the most negative binding energy of all, while control species like I, L and

128

K have $E_{\text{bind}} \sim 0$. Therefore H, F, W and Y are included on sites b, e and f, while hydrophilic AAs are included on sites a, c, d and g.

5.4.4 Design procedure

The sequence optimization procedure is as follows:

1. 219 binding geometries g with different R, ϕ_1 or ϕ_2 are sampled with 100 "good sequences", of which 47 geometries with lowest average binding energy are selected. The good sequences have relatively low E_{bind} with high frequency of amino acids H, F, W and Y.

2. Cluster expansion is constructed for each geometry with enough training sequences to obtain geometry-specific ECIs.

3. The optimal binding energy for a given sequence is approximated by the sequence's lowest binding energy when it adopts each of the 47 selected geometries, or $\min_g[E_{\text{bind}}(g)]$. We optimize for low $\min_g[E_{\text{bind}}(g)]$ over sequence space to find strong CNT binders. Here we use a limited search space for sequences of the form $(\text{abcdefg})_n$. About 10^7 sequences have been screened with fast CE evaluation.

4. Optimization for high E_{dimer} into the coiled coil (using parameters from the case study) to avoid self-aggregation, because we want the peptides to bind to CNTs, not to themselves.

5. Optimization for high helix content with AGADIR [3] and secondary structure prediction with PROF [148], SSPRO [172], to ensure α-helical structure in solution.

6. 100 ns MD simulation with explicit solvent molecules starting from the α-helix structure. Except for a few sequences where the ends went off, the predicted structures mostly retain their α-helical structure.

7. Optimization for high E_{bind}(2nd optimal geometry)- E_{bind}(optimal geometry), which means the binding is preferential, to ensure a stable binding and constant charge distribution. The latter property is important for the separation of the hybrids by chromatography.

Out of about 10^7 possible sequences, we have finally come up with 3 designed sequences (AWAAYHK)n, (RWAAYRS)n, (RWADYRK)n. It is up to future experiment to test their affinity with carbon nanotubes.

5.5   Conclusions

We have demonstrated the practical predictive power of the cluser expansion in several case studies of small peptides, as well as in the design of short peptides for sequence specificity. What is shown in the previous and this chapter is just the beginning. Since the method is new to the computational biology community, algorithms more sophisticated than what we currently use that incorporate more physical intuition are easy to imagine. Additionally, given the interpretability of CE, cluster expansions of many closely related structures may reveal key structure determinants.

It is found that the successful application of cluster expansion in peptides requires a stable protein topology. Through the nanotube-binder design example, we show that this strict requirement can be loosened by expanding energies for several variants of the same structure. Once expansions are complete, evaluation of a sequence, or of all sequences in a proteome, on each of the backbones is extremely fast.

The CE methodology can be coupled with any energy model, e.g. more accurate Hamiltonians or experimentally determined energies, and properties other than energy are potentially expandable. Thus, it can be extended to treat any multi-species search problem for which an appropriate scoring scheme can be generated. In structural biology, this includes modeling not only protein stability, but protein interaction specificity, DNA and RNA structure, protein-DNA interactions, and potentially the interactions of small-molecule pharmaceuticals. We are optimistic that the method will find a wide range of practical applications in biology research.

Chapter 6

Conclusions

In this thesis, we have dealt with two computational modeling problems. In the first part, the aim is to accurately predict properties of transition metal oxides from first principles DFT calculations. The strong electronic correlations are explicitly treated with the "$+U$" extension of the LDA or GGA approximation. We have obtained not only the expected improvement in the electronic structure property predictions, e.g. band gap and magnetization, but also very accurate results in total energy calculations, including redox potential of transition metal ions and low temperature phase stability. Given these good results on ground state properties, we have also explored the finite temperature thermodynamics of mixed-valence oxides. The most exciting finding is associated with Li_xFePO_4, where configurational entropy of localized d-electrons leads to an unexpected solid solution phase. This result should encourage further research on a wide range of materials investigating the effects of electronic entropy on their temperature-dependent phase diagram.

In the second part, we have successfully applied the cluster expansion method, traditionally used in crystalline solids, to protein systems. The method is able to simplify expensive energy calculations for protein sequences to simple energy functions in terms of sequence variables, reducing the order of the problem by many orders of magnitude. By using a heuristic algorithm, we are able to adopt the model to the specific structure of the protein and reach good prediction accuracy. Additionally, physically meaningful interaction parameters are extracted that can help our quanti-

tative understanding of the studied protein. A "real world" protein design problem exemplifies the power of the cluster expansion in greatly expediting the design process. Since the method has just been recently introduced to biological systems, we are optimistic that it can be further refined and will find even more applications in the biology research.

Bibliography

[1] A. Acharya, S. B. Ruvinov, J. Gal, J. R. Moll, and C. Vinson. A heterodimerizing leucine zipper coiled coil system for examining the specificity of a position interactions: Amino acids I, V, L, N, A, and K. *Biochemistry*, 41(48):14122–14131, 2002.

[2] S. L. Adler. Quantum theory of the dielectric constant in real solids. *Phys. Rev.*, 126:413–420, 1962.

[3] AGADIR. http://www.embl-heidelberg.de/services/serrano/agadir/.

[4] Mayssam H. Ali, Christina M. Taylor, Gevorg Grigoryan, Karen N. Allen, Barbara Imperiali, and Amy E. Keating. Design of a Heterospecific, Tetrameric, 21-Residue Miniprotein with Mixed a/b Structure. *Structure*, 13(2):225–234, 2005.

[5] M. Allert, S. S. Rizk, L. L. Looger, and H. W. Hellinga. Computational design of receptors for an organophosphate surrogate of the nerve agent soman. *Proc. Natl. Acad. Sci.*, 101(21):7907–7912, 2004.

[6] J. D. Althoff, D. D. Johnson, and F. J. Pinski. Commensurate and incommensurate ordering tendencies in the ternary fcc Cu-Ni-Zn system. *Phys. Rev. Lett.*, 74(1):138–141, 1995.

[7] K. Amine, H. Yasuda, and M. Yamachi. Olivine $LiCoPO_4$ as 4.8 V electrode material for lithium batteries. *Electrochem. Solid State Lett.*, 3(4):178–179, 2000.

[8] A. S. Andersson, B. Kalska, L. Haggstrom, and J. O. Thomas. Lithium extraction/insertion in $LiFePO_4$: an X-ray diffraction and Mossbauer spectroscopy study. *Solid State Ionics*, 130(1-2):41–52, 2000.

[9] A. S. Andersson, J. O. Thomas, B. Kalska, and L. Haggstrom. Thermal stability of $LiFePO_4$-based cathodes. *Electrochem. Solid State Lett.*, 3(2):66–68, 2000.

[10] V. I. Anisimov. *Strong coulomb correlations in electronic structure calculations : beyond the local density approximation.* Advances in condensed matter science ; v. 1. Gordon and Breach, Amsterdam, Netherlands, 2000.

[11] V. I. Anisimov, F. Aryasetiawan, and A. I. Lichtenstein. First-principles calculations of the electronic structure and spectra of strongly correlated systems: The LDA+U method. *J. Phys.-Condens. Mat.*, 9(4):767–808, 1997.

[12] V. I. Anisimov, I. V. Solovyev, M. A. Korotin, M. T. Czyzyk, and G. A. Sawatzky. Density-Functional Theory and NiO Photoemission Spectra. *Phys. Rev. B*, 48(23):16929–16934, 1993.

[13] V. I. Anisimov, J. Zaanen, and O. K. Andersen. Band theory and Mott insulators - Hubbard-U instead of Stoner-I. *Phys. Rev. B*, 44(3):943–954, 1991.

[14] M. Asta, R. McCormack, and D. Defontaine. Theoretical-study of alloy phase-stability in the Cd-Mg system. *Phys. Rev. B*, 48(2):748–766, 1993.

[15] M. Asta, V. Ozolins, and C. Woodward. A first-principles approach to modeling alloy phase equilibria. *Jom-J. Min. Met. Mat. Soc.*, 53(9):16–19, 2001.

[16] A. Augustsson, G. V. Zhuang, S. M. Butorin, J. M. Osorio-Guillen, C. L. Dong, R. Ahuja, C. L. Chang, P. N. Ross, J. Nordgren, and J. H. Guo. Electronic structure of phospho-olivines Li_xFePO_4 (x=0,1) from soft-x-ray-absorption and -emission spectroscopies. *J. Chem. Phys.*, 123(18):184717, 2005.

[17] M. K. Aydinol, A. F. Kohan, and G. Ceder. Ab initio calculation of the intercalation voltage of lithium transition metal oxide electrodes for rechargeable batteries. *J. Power Sources*, 68(2):664–668, 1997.

[18] M. K. Aydinol, A. F. Kohan, G. Ceder, K. Cho, and J. Joannopoulos. Ab initio study of lithium intercalation in metal oxides and metal dichalcogenides. *Phys. Rev. B*, 56(3):1354–1365, 1997.

[19] R. H. Baughman, A. A. Zakhidov, and W. A. de Heer. Carbon nanotubes - the route toward applications. *Science*, 297(5582):787–792, 2002.

[20] A. D. Becke. Density-Functional Exchange-Energy Approximation with Correct Asymptotic-Behavior. *Phys Rev A*, 38(6):3098–3100, 1988.

[21] R. Benedek, M. M. Thackeray, and L. H. Yang. First-principles calculation of atomic structure and electrochemical potential of Li1+xV3O8. *J. Power Sources*, 82:487–490, 1999.

[22] P. Beroza and D. R. Fredkin. Calculation of amino acid pK(a)s in a protein from a continuum electrostatic model: Method and sensitivity analysis. *J. Comput. Chem.*, 17(10):1229–1244, 1996.

[23] P. E. Blochl. Projector Augmented-Wave Method. *Phys. Rev. B*, 50(24):17953–17979, 1994.

[24] O. Breuer and U. Sundararaj. Big returns from small fibers: A review of polymer/carbon nanotube composites. *Polym Composite*, 25(6):630–645, 2004.

[25] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus. Charmm - a program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.*, 4(2):187–217, 1983.

[26] Peter G. Bruce. *Solid state electrochemistry*. Chemistry of solid state materials ; [5]. Cambridge University Press, Cambridge ; New York, NY, USA, 1995.

[27] C. W. Carter, B. C. LeFebvre, S. A. Cammer, A. Tropsha, and M. H. Edgell. Four-body potentials reveal protein-specific correlations to stability changes caused by hydrophobic core mutations. *J. Mol. Biol.*, 311(4):625–638, 2001.

135

[28] G. Ceder. A derivation of the Ising model for the computation of phase diagrams. *Comp. Mater. Sci.*, 1(2):144–150, 1993.

[29] G. Ceder. A Computational Study of Oxygen Ordering in YBa_2Cu_3Oz and Its Relation to Superconductivity. *Mol. Simulat.*, 12(2):141–153, 1994.

[30] G. Ceder. Computational materials science - Predicting properties from scratch. *Science*, 280(5366):1099–1100, 1998.

[31] G. Ceder, M. K. Aydinol, and A. F. Kohan. Application of first-principles calculations to the design of rechargeable Li-batteries. *Comp. Mater. Sci.*, 8(1-2):161–169, 1997.

[32] G. Ceder, Y. M. Chiang, D. R. Sadoway, M. K. Aydinol, Y. I. Jang, and B. Huang. Identification of cathode materials for lithium batteries guided by first-principles calculations. *Nature*, 392(6677):694–696, 1998.

[33] G. Ceder and A. Van der Ven. Phase diagrams of lithium transition metal oxides: investigations from first principles. *Electrochim. Acta*, 45(1-2):131–150, 1999.

[34] G. Y. Chen, X. Y. Song, and T. J. Richardson. Electron microscopy study of the LiFePO4 to FePO4 phase transition. *Electrochem. Solid State Lett.*, 9(6):A295–a298, 2006.

[35] Z. H. Chen and J. R. Dahn. Reducing carbon in LiFePO4/C composite electrodes to maximize specific energy, volumetric energy, and tap density. *J. Electrochem. Soc.*, 149(9):A1184–a1189, 2002.

[36] S. Choi and A. Manthiram. Synthesis and electrochemical properties of LiCo2O4 spinel cathodes. *J. Electrochem. Soc.*, 149(2):A162–a166, 2002.

[37] S. Y. Chung, J. T. Bloking, and Y. M. Chiang. Electronically conductive phospho-olivines as lithium storage electrodes. *Nature Mater.*, 1(2):123–128, 2002.

[38] S. Y. Chung and Y. M. Chiang. Microscale measurements of the electrical conductivity of doped LiFePO4. *Electrochem. Solid State Lett.*, 6(12):A278–a281, 2003.

[39] M. Cococcioni and S. de Gironcoli. Linear response approach to the calculation of the effective interaction parameters in the LDA+U method. *Phys. Rev. B*, 71(3):035105, 2005.

[40] J. W. D. Connolly and A. R. Williams. Density-Functional Theory Applied to Phase-Transformations in Transition-Metal Alloys. *Phys. Rev. B*, 27(8):5169–5172, 1983.

[41] T. P. Creamer and G. D. Rose. Side-chain entropy opposes alpha-helix formation but rationalizes experimentally determined helix-forming Propensities. *Proc. Natl. Acad. Sci.*, 89(13):5937–5941, 1992.

[42] F. H. Crick. The packing of α-helices: simple coiled-coils. *Acta. Cryst.*, 6:689, 1953.

[43] B. I. Dahiyat and S. L. Mayo. De novo protein design: Fully automated sequence selection. *Science*, 278(5335):82–87, 1997.

[44] D. Dai, M. H. Whangbo, H. J. Koo, X. Rocquefelte, S. Jobic, and A. Villesuzanne. Analysis of the spin exchange interactions and the ordered magnetic structures of lithium transition metal phosphates LiMPO4 (M = Mn, Fe, Co, Ni) with the olivine structure. *Inorg. Chem.*, 44(7):2407–2413, 2005.

[45] A. B. Dalton, A. Ortiz-Acevedo, V. Zorbas, E. Brunner, W. M. Sampson, L. Collins, J. M. Razal, M. M. Yoshida, R. H. Baughman, R. K. Draper, I. H. Musselman, M. Jose-Yacaman, and G. R. Dieckmann. Hierarchical self-assembly of peptide-coated carbon nanotubes. *Adv Funct Mater*, 14(12):1147–1151, 2004.

[46] M. E. A. Y. de Dompablo, U. Amador, and F. Garcia-Alvarado. An experimental and computational study of the electrode material olivine-LiCoAsO4. *J. Electrochem. Soc.*, 153(4):A673–a678, 2006.

[47] M. E. A. Y. de Dompablo, J. M. Gallardo-Amores, and U. Amador. Lithium insertion in the high-pressure polymorph of FePO4 - Computational predictions and experimental findings. *Electrochem. Solid State Lett.*, 8(11):A564–a569, 2005.

[48] M. E. A. Y. de Dompablo, A. Van der Ven, and G. Ceder. First-principles calculations of lithium ordering and phase stability on Li_xNiO_2. *Phys. Rev. B*, 66(6):064112, 2002.

[49] D. de Fontaine. In H. Ehrenreich and D. Turnbull, editors, *Solid State Physics*, volume 47, page 33. Academic Press, 1994.

[50] C. Delacourt, L. Laffont, R. Bouchet, C. Wurm, J. B. Leriche, M. Morcrette, J. M. Tarascon, and C. Masquelier. Toward understanding of electrical limitations (electronic, ionic) in $LiMPO_4$ (M = Fe, Mn) electrode materials. *J. Electrochem. Soc.*, 152(5):A913–a921, 2005.

[51] C. Delacourt, P. Poizot, J. M. Tarascon, and C. Masquelier. The existence of a temperature-driven solid solution in Li_xFePO_4 for $0 <= x <= 1$. *Nature Mater.*, 4(3):254–260, 2005.

[52] C. Delacourt, J. Rodriguez-Carvajal, B. Schmitt, J. M. Tarascon, and C. Masquelier. Crystal chemistry of the olivine-type Li_xFePO_4 system $(0 <= x <= 1)$ between 25 and 370 degrees C. *Solid State Sci.*, 7(12):1506–1516, 2005.

[53] C. Delacourt, C. Wurm, P. Reale, M. Morcrette, and C. Masquelier. Low temperature preparation of optimized phosphates for Li-battery applications. *Solid State Ionics*, 173(1-4):113–118, 2004.

[54] C. Delmas, M. Menetrier, L. Croguennec, S. Levasseur, J. P. Peres, C. Pouillerie, G. Prado, L. Fournes, and F. Weill. Lithium batteries: a new tool in solid state chemistry. *Int. J. Inorg. Mater.*, 1(1):11–19, 1999.

[55] P. Deniard, A. M. Dulac, X. Rocquefelte, V. Grigorova, O. Le Bacq, A. Pasturel, and S. Jobic. High potential positive materials for lithium-ion batteries: transition metal phosphates. *J. Phys. Chem. Solids.*, 65(2-3):229–233, 2004.

[56] J. Desmet, M. Demaeyer, B. Hazes, and I. Lasters. The Dead-End Elimination Theorem and Its Use in Protein Side-Chain Positioning. *Nature*, 356(6369):539–542, 1992.

[57] G. R. Dieckmann, A. B. Dalton, P. A. Johnson, J. Razal, J. Chen, G. M. Giordano, E. Munoz, I. H. Musselman, R. H. Baughman, and R. K. Draper. Controlled assembly of carbon nanotubes by designed amphiphilic peptide helices. *J. Am. Chem. Soc.*, 125(7):1770–1777, 2003.

[58] J. L. Dodd, R. Yazami, and B. Fultz. Phase diagram of Li_4FePO_4. *Electrochem. Solid State Lett.*, 9(3):A151–a155, 2006.

[59] M. L. Doublet, F. Lemoigno, F. Gillot, and L. Monconduit. The Li(x)VPn(4) ternary phases (Pn = P, As): Rigid networks for lithium intercalation/deintercalation. *Chem. Mater.*, 14(10):4126–4133, 2002.

[60] M. S. Dresselhaus, G. Dresselhaus, and Phaedon Avouris. *Carbon nanotubes : synthesis, structure, properties, and applications*. Topics in applied physics, v. 80. Springer, Berlin ; New York, 2001.

[61] S. L. Dudarev, G. A. Botton, S. Y. Savrasov, C. J. Humphreys, and A. P. Sutton. Electron-energy-loss spectra and the structural stability of nickel oxide: An LSDA+U study. *Phys. Rev. B*, 57(3):1505–1509, 1998.

[62] R. L. Dunbrack. Rotamer libraries in the 21(st) century. *Curr. Opin. Struc. Biol.*, 12(4):431–440, 2002.

[63] M. A. Dwyer, L. L. Looger, and H. W. Hellinga. Computational design of a biologically active enzyme. *Science*, 304(5679):1967–1971, 2004.

[64] T. W. Ebbesen and P. M. Ajayan. Large-Scale Synthesis of Carbon Nanotubes. *Nature*, 358(6383):220–222, 1992.

[65] J. L. G. Fierro. *Metal oxides : chemistry and applications*. CRC Taylor & Francis, Boca Raton, FL, 2006.

[66] J. H. Fong, A. E. Keating, and M. Singh. Predicting specificity in bZIP coiled-coil protein interactions. *Genome Biol.*, 5(2):R11, 2004.

[67] G. D. Garbulsky and G. Ceder. Linear-programming method for obtaining effective cluster interactions in alloys from total-energy calculations - application to the fcc Pd-V system. *Phys. Rev. B*, 51(1):67–72, 1995.

[68] O. Garcia-Moreno, M. Alvarez-Vega, F. Garcia-Alvarado, J. Garcia-Jaca, J. M. Gallardo-Amores, M. L. Sanjuan, and U. Amador. Influence of the structure on the electrochemical performance of lithium transition metal phosphates as cathodic materials in rechargeable lithium batteries: A new high-pressure form of LiMPO4 (M = Fe and Ni). *Chem. Mater.*, 13(5):1570–1576, 2001.

[69] A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg. Dynamical mean-field theory of strongly correlated fermion systems and the limit of infinite dimensions. *Rev. Mod. Phys.*, 68(1):13–125, 1996.

[70] D. Gilis and M. Rooman. PoPMuSiC, an algorithm for predicting protein mutant stability changes. Application to prion proteins. *Protein Eng.*, 13(12):849–856, 2000.

[71] R. F. Goldstein. Efficient Rotamer Elimination Applied to Protein Side-Chains and Related Spin-Glasses. *Biophys. J.*, 66(5):1335–1340, 1994.

[72] D. B. Gordon, S. A. Marshall, and S. L. Mayo. Energy functions for protein design. *Curr. Opin. Struc. Biol.*, 9(4):509–513, 1999.

[73] D. X. Gouveia, V. Lemos, J. A. C. de Paiva, A. G. Souza, J. Mendes, S. M. Lala, L. A. Montoro, and J. M. Rosolen. Spectroscopic studies of Li_xFePO_4 and $Li_xM_{0.03}Fe_{0.97}PO_4$ (M=Cr,Cu,Al,Ti). *Phys. Rev. B*, 72(2):024105, 2005.

[74] G. Grigoryan and A. E. Keating. Structure-based prediction of bZIP partnering specificity. *J. Mol. Biol.*, 355(5):1125–1142, 2006.

[75] Gevorg Grigoryan, Fei Zhou, Steve R. Lustig, Gerbrand Ceder, Dane Morgan, and Amy E. Keating. Ultra-fast evaluation of protein energies directly from sequence. *PLoS Computational Biology*, 2(6):e63, 2006.

[76] N. Hamada, S. Sawada, and A. Oshiyama. New One-Dimensional Conductors - Graphitic Microtubules. *Phys. Rev. Lett.*, 68(10):1579–1581, 1992.

[77] Lars Hedin. New method for calculating the one-particle Green's function with application to the electron-gas problem. *Phys. Rev.*, 139(3A):A796A823, 1965.

[78] P. S. Herle, B. Ellis, N. Coombs, and L. F. Nazar. Nano-network electronic conduction in iron and nickel olivine phosphates. *Nature Mater.*, 3(3):147–152, 2004.

[79] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136:864–871, 1964.

[80] B. Honig and A. Nicholls. Classical Electrostatics in Biology and Chemistry. *Science*, 268(5214):1144–1149, 1995.

[81] R. Hoppe, G. Brachtel, and M. Jansen. Knowledge of Oxomanganates(Iii) - Limno2 and Beta-Namno2. *Z. Anorg. Allg. Chem.*, 417(1):1–10, 1975.

[82] H. Huang, S. C. Yin, T. Kerr, N. Taylor, and L. F. Nazar. Nanostructured composites: a high capacity, fast rate $Li_3V_2(PO_4)_3$/carbon cathode for rechargeable lithium batteries. *Adv. Mater.*, 14(21):1525, 2002.

[83] H. Huang, S. C. Yin, and L. F. Nazar. Approaching theoretical capacity of LiFePO$_4$ at room temperature at high rates. *Electrochem. Solid State Lett.*, 4(10):A170–A172, 2001.

[84] B. J. Hwang, Y. W. Tsai, D. Carlier, and G. Ceder. A combined computational/experimental study on LiNi$_{1/3}$Co$_{1/3}$Mn$_{1/3}$O$_2$. *Chem. Mater.*, 15(19):3676–3682, 2003.

[85] D. D. Johnson, A. V. Smirnov, J. B. Staunton, F. J. Pinski, and W. A. Shelton. Temperature-induced configurational excitations for predicting thermodynamic and mechanical properties of alloys. *Phys. Rev. B*, 62(18):R11917–R11920, 2000.

[86] R. O. Jones and O. Gunnarsson. The Density Functional Formalism, Its Applications and Prospects. *Rev. Mod. Phys.*, 61(3):689–746, 1989.

[87] K. S. Kang, Y. S. Meng, J. Breger, C. P. Grey, and G. Ceder. Electrodes with high power and high capacity for rechargeable lithium batteries. *Science*, 311(5763):977–980, 2006.

[88] K. R. Kganyago, P. E. Ngoepe, and C. R. A. Catlow. Ab initio calculation of the voltage profile for LiC6. *Solid State Ionics*, 159(1-2):21–23, 2003.

[89] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140:A1133–A1138, 1965.

[90] Y. Koyama, I. Tanaka, and H. Adachi. Electronic structures and electrode potentials of layered lithium insertion electrodes by the first principles calculation. *Adv. Quantum. Chem.*, 42:145–161, 2003.

[91] C. M. Kraemer-Pecore, A. M. Wollacott, and J. R. Desjarlais. Computational protein design. *Curr. Opin. Chem. Biol.*, 5(6):690–695, 2001.

[92] G. Kresse and J. Furthmuller. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comp. Mater. Sci.*, 6(1):15–50, 1996.

142

[93] G. Kresse and J. Furthmuller. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B*, 54(16):11169–11186, 1996.

[94] G. Kresse and D. Joubert. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B*, 59(3):1758–1775, 1999.

[95] D. Krylov, J. Barchi, and C. Vinson. Inter-helical interactions in the leucine zipper coiled coil dimer: pH and salt dependence of coupling energy between charged amino acids. *J. Mol. Biol.*, 279(4):959–972, 1998.

[96] B. Kuhlman, G. Dantas, G. C. Ireton, G. Varani, B. L. Stoddard, and D. Baker. Design of a novel globular protein fold with atomic-level accuracy. *Science*, 302(5649):1364–1368, 2003.

[97] M. H. R. Lankhorst, H. J. M. Bouwmeester, and H. Verweij. Importance of electronic band structure to nonstoichiometric behaviour of La0.8Sr0.2CoO3-delta. *Solid State Ionics*, 96(1-2):21–27, 1997.

[98] K. F. Lau and K. A. Dill. A Lattice Statistical-Mechanics Model of the Conformational and Sequence-Spaces of Proteins. *Macromolecules*, 22(10):3986–3997, 1989.

[99] M. Launay, F. Boucher, P. Gressier, and G. Ouvrard. DFT study of lithium battery materials: application to the beta-VOXO4 systems (X = P, As, S). *J. Solid State Chem.*, 176(2):556–566, 2003.

[100] G. A. Lazar, S. A. Marshall, J. J. Plecs, S. L. Mayo, and J. R. Desjarlais. Designing proteins for therapeutic applications. *Curr. Opin. Struc. Biol.*, 13(4):513–518, 2003.

[101] T. Lazaridis and M. Karplus. Effective energy function for proteins in solution. *Proteins*, 35(2):133–152, 1999.

[102] T. Lazaridis and M. Karplus. Effective energy functions for protein structure prediction. *Curr. Opin. Struc. Biol.*, 10(2):139–145, 2000.

[103] O. Le Bacq, A. Pasturel, and O. Bengone. Impact on electronic correlations on the structural stability, magnetism, and voltage of $LiCoPO_4$ battery. *Phys. Rev. B*, 69(24):245107, 2004.

[104] V. Lemos, S. Guerini, J. Mendes, S. M. Lala, L. A. Montoro, and J. M. Rosolen. A new insight into the LiFePO4 delithiation process. *Solid State Ionics*, 177(11-12):1021–1025, 2006.

[105] G. H. Li, H. Azuma, and M. Tohda. LiMnPO4 as the cathode for lithium batteries. *Electrochem. Solid State Lett.*, 5(6):A135–A137, 2002.

[106] A. I. Liechtenstein, V. I. Anisimov, and J. Zaanen. Density-functional theory and strong-interactions - Orbital ordering in Mott-Hubbard insulators. *Phys. Rev. B*, 52(8):R5467–R5470, 1995.

[107] G. Lu, P. Maragakis, and E. Kaxiras. Carbon nanotube interaction with DNA. *Nano Lett*, 5(5):897–900, 2005.

[108] Steven R. Lustig and Anand Jagota. Selectivity of polypeptide binding to nanoscale substrates. *Mat. Res. Soc. Symp. Proc.*, 724:N4.6, 2002.

[109] S. Maekawa. *Physics of transition metal oxides.* Springer series in solid-state sciences, 144. Springer, Berlin ; New York, 2004.

[110] C. A. Marianetti, D. Morgan, and G. Ceder. First-principles investigation of the cooperative Jahn-Teller effect for octahedrally coordinated transition-metal ions. *Phys. Rev. B*, 6322(22):224304, 2001.

[111] C. Masquelier, A. K. Padhi, K. S. Nanjundaswamy, and J. B. Goodenough. New cathode materials for rechargeable lithium batteries: The 3-D framework structures Li3Fe2(XO4)(3) (X = P, As). *J. Solid State Chem.*, 135(2):228–234, 1998.

[112] T. Maxisch, F. Zhou, and G. Ceder. Ab initio study of the migration of small polarons in olivine Li_xFePO_4 and their association with lithium ions and vacancies. *Phys. Rev. B*, 73(10):104301, 2006.

[113] John M. Mays. Nuclear magnetic resonances and Mn-O-P-O-Mn superexchange linkages in paramagnetic and antiferromagnetic $LiMnPO_4$. *Phys. Rev.*, 131:38–53, 1963.

[114] R. McCormack, D. de Fontaine, C. Wolverton, and G. Ceder. Nonempirical phase-equilibria in the W-Mo-Cr system. *Phys. Rev. B*, 51(22):15808–15822, 1995.

[115] R. McCormack and D. deFontaine. First-principles study of multiple order-disorder transitions in Cd2AgAu Heusler alloys. *Phys. Rev. B*, 54(14):9746–9755, 1996.

[116] M. Menetrier, I. Saadoune, S. Levasseur, and C. Delmas. The insulator-metal transition upon lithium deintercalation from $LiCoO_2$: electronic properties and Li-7 NMR study. *J. Mater. Chem.*, 9(5):1135–1140, 1999.

[117] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *J. Chem. Phys.*, 21(6):1087–1092, 1953.

[118] L. A. Mirny and E. I. Shakhnovich. How to derive a protein folding potential? A new approach to an old problem. *J. Mol. Biol.*, 264(5):1164–1179, 1996.

[119] S. K. Mishra and G. Ceder. Structural stability of lithium manganese oxides. *Phys. Rev. B*, 59(9):6120–6130, 1999.

[120] S. Miyazawa and R. L. Jernigan. Estimation of Effective Interresidue Contact Energies from Protein Crystal-Structures - Quasi-Chemical Approximation. *Macromolecules*, 18(3):534–552, 1985.

[121] S. Miyazawa and R. L. Jernigan. Residue-residue potentials with a favorable contact pair term and an unfavorable high packing density term, for simulation and threading. *J. Mol. Biol.*, 256(3):623–644, 1996.

[122] D. Morgan, G. Ceder, M. Y. Saidi, J. Barker, J. Swoyer, H. Huang, and G. Adamson. Experimental and computational study of the structure and electrochemical properties of $Li_xM_2(PO_4)_3$ compounds with the monoclinic and rhombohedral structure. *Chem. Mater.*, 14(11):4684–4693, 2002.

[123] D. Morgan, G. Ceder, M. Y. Saidi, J. Barker, J. Swoyer, H. Huang, and G. Adamson. Experimental and computational study of the structure and electrochemical properties of monoclinic $Li_xM_2(PO_4)_3$ compounds. *J. Power Sources*, 119:755–759, 2003.

[124] D. Morgan, A. Van der Ven, and G. Ceder. Li conductivity in Li_xMPO_4 (M = Mn, Fe, Co, Ni) olivine materials. *Electrochem. Solid State Lett.*, 7(2):A30–A32, 2004.

[125] Gholamabbas Nazri and G. Pistoia. *Lithium batteries : science and technology*. Kluwer Academic Publishers, Boston, 2004.

[126] D. M. C. Nicholson, G. M. Stocks, Y. Wang, W. A. Shelton, Z. Szotek, and W. M. Temmerman. Stationary nature of the density-functional free-energy - application to accelerated multiple-scattering Calculations. *Phys. Rev. B*, 50(19):14686–14689, 1994.

[127] T. Ohzuku and A. Ueda. Why transition-metal (di) oxides are the most attractive materials for batteries. *Solid State Ionics*, 69(3-4):201–211, 1994.

[128] S. Okada, S. Sawa, M. Egashira, J. Yamaki, M. Tabuchi, H. Kageyama, T. Konishi, and A. Yoshino. Cathode properties of phospho-olivine $LiMPO_4$ for lithium secondary batteries. *J. Power Sources*, 97-8:430–432, 2001.

[129] A. Onufriev, D. A. Case, and D. Bashford. Effective Born radii in the generalized Born approximation: the importance of being perfect. *J. Comput. Chem.*, 23(14):1297–1304, 2002.

[130] E. K. Oshea, R. Rutkowski, and P. S. Kim. Mechanism of Specificity in the Fos-Jun Oncoprotein Heterodimer. *Cell*, 68(4):699–708, 1992.

[131] A. K. Padhi, V. Manivannan, and J. B. Goodenough. Tuning the position of the redox couples in materials with NASICON structure by anionic substitution. *J. Electrochem. Soc.*, 145(5):1518–1520, 1998.

[132] A. K. Padhi, K. S. Nanjundaswamy, and J. B. Goodenough. Phospho-olivines as positive-electrode materials for rechargeable lithium batteries. *J. Electrochem. Soc.*, 144(4):1188–1194, 1997.

[133] A. K. Padhi, K. S. Nanjundaswamy, C. Masquelier, and J. B. Goodenough. Mapping of transition metal redox energies in phosphates with NASICON structure by lithium intercalation. *J. Electrochem. Soc.*, 144(8):2581–2586, 1997.

[134] A. K. Padhi, K. S. Nanjundaswamy, C. Masquelier, S. Okada, and J. B. Goodenough. Effect of structure on the Fe3+/Fe2+ redox couple in iron phosphates. *J. Electrochem. Soc.*, 144(5):1609–1613, 1997.

[135] Robert G. Parr and Weitao Yang. *Density-functional theory of atoms and molecules.* International series of monographs on chemistry ; 16. Oxford University Press ; Clarendon Press, New York Oxford [England], 1989.

[136] J. P. Perdew. Size-consistency, self-interaction correction, and derivative discontinuity in density functional theory. *Adv. Quantum Chem.*, 21:113–134, 1990.

[137] J. P. Perdew, K. Burke, and M. Ernzerhof. Generalized gradient approximation made simple. *Phys. Rev. Lett.*, 77(18):3865–3868, 1996.

[138] J. P. Perdew and M. Levy. Physical content of the exact Kohn-Sham orbital energies - band-gaps and derivative discontinuities. *Phys. Rev. Lett.*, 51(20):1884–1887, 1983.

[139] J. P. Perdew and M. Levy. Many-body phenomena at surfaces. pages xiv, 578 p. Academic Press, Orlando, 1984.

[140] J. P. Perdew, R. G. Parr, M. Levy, and J. L. Balduz. Density-functional theory for fractional particle number - derivative discontinuities of the energy. *Phys. Rev. Lett.*, 49(23):1691–1694, 1982.

[141] J. P. Perdew and Y. Wang. Accurate and Simple Analytic Representation of the Electron-Gas Correlation-Energy. *Phys. Rev. B*, 45(23):13244–13249, 1992.

[142] J. P. Perdew and A. Zunger. Self-interaction correction to density-functional approximations for many-electron systems. *Phys. Rev. B*, 23(10):5048–5079, 1981.

[143] A. G. Petukhov, I. I. Mazin, L. Chioncel, and A. I. Lichtenstein. Correlated metals and the LDA+U method. *Phys. Rev. B*, 67(15):153106, 2003.

[144] W. E. Pickett. Electronic-structure of the high-temperature oxide superconductors. *Rev. Mod. Phys.*, 61(2):433–512, 1989.

[145] W. E. Pickett, S. C. Erwin, and E. C. Ethridge. Reformulation of the LDA+U method for a local-orbital basis. *Phys. Rev. B*, 58(3):1201–1209, 1998.

[146] N. Pokala and T. M. Handel. Review: Protein design - Where we were, where we are, where we're going. *J. Struct. Biol.*, 134(2-3):269–281, 2001.

[147] R. Prasad, R. Benedek, and M. M. Thackeray. Effect of Co on the magnetism and phase stability of lithiated manganese oxides. *Bull. Mater. Sci.*, 26(1):147–150, 2003.

[148] PROF. http://www.aber.ac.uk/ phiwww/prof/.

[149] P. P. Prosini, D. Zane, and M. Pasquali. Improved electrochemical performance of a $LiFePO_4$-based composite cathode. *Electrochim. Acta*, 46(23):3517–3523, 2001.

[150] C. N. R. Rao and B. Raveau. *Transition metal oxides : structure, properties, and synthesis of ceramic oxides.* Wiley-VCH, New York, 2nd edition, 1998.

[151] J. Reed and G. Ceder. Charge, potential, and phase stability of layered Li(Ni0.5Mn0.5)O-2. *Electrochem. Solid State Lett.*, 5(7):A145–A148, 2002.

[152] J. Reed, G. Ceder, and A. Van der Ven. Layered-to-spinel phase transition in $Li_x MnO_2$. *Electrochem. Solid State Lett.*, 4(6):A78–A81, 2001.

[153] X. Rocquefelte, F. Boucher, P. Gressier, and G. Ouvrard. First-principle study of the intercalation process in the LixV2O5 system. *Chem. Mater.*, 15(9):1812–1819, 2003.

[154] A. Rossi, C. Micheletti, F. Seno, and A. Maritan. A self-consistent knowledge-based approach to protein design. *Biophys. J.*, 80(1):480–490, 2001.

[155] A. Rougier, C. Delmas, and A. V. Chadwick. Noncooperative Jahn-Teller effect in $LiNiO_2$ - an EXAFS study. *Solid State Commun.*, 94(2):123–127, 1995.

[156] G. Rousse, J. Rodriguez-Carvajal, S. Patoux, and C. Masquelier. Magnetic structures of the triphylite $LiFePO_4$ and of its delithiated form $FePO_4$. *Chem. Mater.*, 15(21):4082–4090, 2003.

[157] W. P. Russ and R. Ranganathan. Knowledge-based potential functions in protein design. *Curr. Opin. Struc. Biol.*, 12(4):447–452, 2002.

[158] M. Y. Saidi, J. Barker, H. Huang, J. L. Swoyer, and G. Adamson. Electrochemical properties of lithium vanadium phosphate as a cathode material for lithium-ion batteries. *Electrochem. Solid State Lett.*, 5(7):A149–A151, 2002.

[159] J. M. Sanchez, F. Ducastelle, and D. Gratias. Generalized cluster description of multicomponent systems. *Physica. A*, 128(1-2):334–350, 1984.

[160] J. M. Sanchez, J. P. Stark, and V. L. Moruzzi. 1st-principles calculation of the Ag-Cu phase-diagram. *Phys. Rev. B*, 44(11):5411–5418, 1991.

[161] R. P. Santoro and R. E. Newman. Antiferromagnetism in $LiFePO_4$. *Acta Cryst.*, 22:344–347, 1967.

[162] R. P. Santoro, D. J. Segal, and R. E. Newman. Magnetic properties of $LiCoPO_4$ and $LiNiPO_4$. *J. Phys. Chem. Solids.*, 27(6-7):1192–1193, 1966.

[163] S. Y. Savrasov, G. Kotliar, and E. Abrahams. Correlated electrons in delta-plutonium within a dynamical mean-field picture. *Nature*, 410(6830):793–795, 2001.

[164] Walter A. van Schalkwijk and Bruno Scrosati. *Advances in lithium-ion batteries.* Kluwer Academic/Plenum Publishers, New York, NY, 2002.

[165] P. Schleger, W. N. Hardy, and H. Casalta. Model for the High-Temperature Oxygen-Ordering Thermodynamics in Yba2Cu3O6+X - Inclusion of Electron-Spin and Charge Degrees of Freedom. *Phys. Rev. B*, 49(1):514–523, 1994.

[166] S. Q. Shi, L. J. Liu, C. Y. Ouyang, D. S. Wang, Z. X. Wang, L. Q. Chen, and X. J. Huang. Enhancement of electronic conductivity of $LiFePO_4$ by Cr doping and its identification by first-principles calculations. *Phys. Rev. B*, 68(19):195108, 2003.

[167] A. B. Shick, V. Drchal, and L. Havela. Coulomb-U and magnetic-moment collapse in delta-Pu. *Europhys. Lett.*, 69(4):588–594, 2005.

[168] K. T. Simons, I. Ruczinski, C. Kooperberg, B. A. Fox, C. Bystroff, and D. Baker. Improved recognition of native-like protein structures using a combination of sequence-dependent and sequence-independent features of proteins. *Proteins*, 34(1):82–95, 1999.

[169] M. Sluiter, P. E. A. Turchi, F. J. Pinski, and G. M. Stocks. A 1st-principles study of phase-stability in Ni-Al and Ni-Ti alloys. *Mat Sci Eng a-Struct*, 152(1-2):1–8, 1992.

150

[170] M. H. F. Sluiter and Y. Kawazoe. Invariance of truncated cluster expansions for first-principles alloy thermodynamics. *Phys. Rev. B*, 71(21):212201, 2005.

[171] V. Srinivasan and J. Newman. Existence of path-dependence in the $LiFePO_4$ electrode. *Electrochem. Solid State Lett.*, 9(3):A110–a114, 2006.

[172] SSPRO. http://www.igb.uci.edu/tools/scratch/.

[173] M. Stadele, J. A. Majewski, P. Vogl, and A. Gorling. Exact Kohn-Sham exchange potential in semiconductors. *Phys. Rev. Lett.*, 79(11):2089–2092, 1997.

[174] W. C. Still, A. Tempczyk, R. C. Hawley, and T. Hendrickson. Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.*, 112(16):6127–6129, 1990.

[175] Y. P. Sun, K. F. Fu, Y. Lin, and W. J. Huang. Functionalized carbon nanotubes: Properties and applications. *Accounts Chem Res*, 35(12):1096–1104, 2002.

[176] A. Svane and O. Gunnarsson. Transition-metal oxides in the self-interaction corrected density-functional formalism. *Phys. Rev. Lett.*, 65(9):1148–1151, 1990.

[177] P. Tang and N. A. W. Holzwarth. Electronic structure of $FePO_4$, $LiFePO_4$, and related materials. *Phys. Rev. B*, 68(16):165107, 2003.

[178] P. D. Tepesch, G. D. Garbulsky, and G. Ceder. Model for configurational thermodynamics in ionic systems. *Phys. Rev. Lett.*, 74(12):2272–2275, 1995.

[179] I. Terasaki, Y. Sasago, and K. Uchinokura. Large thermoelectric power in NaCo2O4 single crystals. *Phys. Rev. B*, 56(20):R12685–R12687, 1997.

[180] R. Tetot, V. Pagot, and C. Picard. Thermodynamics of YBa2Cu3O6+2x: Predictions of the asymmetric next-nearest-neighbor Ising model versus experimental data. *Phys. Rev. B*, 59(22):14748–14752, 1999.

[181] M. M. Thackeray. Structural considerations of layered and spinel lithiated oxides for lithium ion batteries. *J. Electrochem. Soc.*, 142(8):2558–2563, 1995.

[182] A. Thess, R. Lee, P. Nikolaev, H. J. Dai, P. Petit, J. Robert, C. H. Xu, Y. H. Lee, S. G. Kim, A. G. Rinzler, D. T. Colbert, G. E. Scuseria, D. Tomanek, J. E. Fischer, and R. E. Smalley. Crystalline ropes of metallic carbon nanotubes. *Science*, 273(5274):483–487, 1996.

[183] L. Vaccarini, C. Goze, L. Henrard, E. Hernandez, P. Bernier, and A. Rubio. Mechanical and electronic properties of carbon and boron-nitride nanotubes. *Carbon*, 38(11-12):1681–1690, 2000.

[184] A. van de Walle and M. Asta. Self-driven lattice-model Monte Carlo simulations of alloy thermodynamic properties and phase diagrams. *Model. Simul. Mater. Sc.*, 10(5):521–538, 2002.

[185] A. van de Walle, M. Asta, and G. Ceder. The Alloy Theoretic Automated Toolkit: A user guide. *Calphad*, 26(4):539–553, 2002.

[186] A. van de Walle and G. Ceder. Automating first-principles phase diagram calculations. *J. Phase Equilib.*, 23(4):348–359, 2002.

[187] A. van de Walle and G. Ceder. The effect of lattice vibrations on substitutional alloy thermodynamics. *Rev. Mod. Phys.*, 74(1):11–45, 2002.

[188] A. Van der Ven, M. K. Aydinol, and G. Ceder. First-principles evidence for stage ordering in Li_xCoO_2. *J. Electrochem. Soc.*, 145(6):2149–2155, 1998.

[189] A. Van der Ven, M. K. Aydinol, G. Ceder, G. Kresse, and J. Hafner. First-principles investigation of phase stability in Li_xCoO_2. *Phys. Rev. B*, 58(6):2975–2987, 1998.

[190] A. Van der Ven and G. Ceder. Electrochemical properties of spinel Li_xCoO_2: A first-principles investigation. *Phys. Rev. B*, 59(2):742–749, 1999.

[191] A. Van der Ven and G. Ceder. Lithium diffusion in layered Li_xCoO_2. *Electrochem. Solid State Lett.*, 3(7):301–304, 2000.

[192] A. Van der Ven and G. Ceder. Lithium diffusion mechanisms in layered intercalation compounds. *J. Power Sources*, 97-8:529–531, 2001.

[193] A. Van der Ven and G. Ceder. First principles calculation of the interdiffusion coefficient in binary alloys. *Phys. Rev. Lett.*, 94(4):–, 2005.

[194] Anton Van der Ven, Massachusetts Institute of Technology. Dept. of Materials Science, and Engineering. *First principles investigation of the thermodynamic and kinetic properties of lithium transition metal oxides.* PhD thesis, 2000.

[195] M. Vendruscolo and E. Domany. Pairwise contact potentials are unsuitable for protein folding. *J. Chem. Phys.*, 109(24):11101–11108, 1998.

[196] M. Vendruscolo and E. Paci. Protein folding: bringing theory and experiment closer together. *Curr. Opin. Struc. Biol.*, 13(1):82–87, 2003.

[197] Axel van de Walle, Massachusetts Institute of Technology. Dept. of Materials Science, and Engineering. *The effect of lattice vibrations on substitutional alloy thermodynamics.* PhD thesis, 2000.

[198] S. Q. Wang, E. S. Humphreys, S. Y. Chung, D. F. Delduco, S. R. Lustig, H. Wang, K. N. Parker, N. W. Rizzo, S. Subramoney, Y. M. Chiang, and A. Jagota. Peptides with selective affinity for carbon nanotubes. *Nature Mater.*, 2(3):196–200, 2003.

[199] J. A. White and D. M. Bird. Implementation of Gradient-Corrected Exchange-Correlation Potentials in Car-Parrinello Total-Energy Calculations. *Phys. Rev. B*, 50(7):4954–4957, 1994.

[200] Nathan Wiser. Dielectric constant with local field effects included. *Phys. Rev.*, 129:62–69, 1963.

[201] J. Wolfenstine and J. Allen. Ni3+/Ni2+ redox potential in $LiNiPO_4$. *J. Power Sources*, 142(1-2):389–390, 2005.

[202] C. Wolverton and V. Ozolins. Entropically favored ordering: The metallurgy of Al_2Cu revisited. *Phys. Rev. Lett.*, 86(24):5518–5521, 2001.

[203] C. Wolverton and A. Zunger. First-principles theory of short-range order, electronic excitations and spin polarization in Ni-V and Pd-V alloys. *Phys. Rev. B*, 52(12):8813–8828, 1995.

[204] C. Wolverton and A. Zunger. Cation and vacancy ordering in Li_xCoO_2. *Phys. Rev. B*, 57(4):2242–2252, 1998.

[205] C. Wolverton and A. Zunger. Prediction of Li intercalation and battery voltages in layered vs. cubic Li_xCoO_2. *J. Electrochem. Soc.*, 145(7):2424–2431, 1998.

[206] Y. N. Xu, S. Y. Chung, J. T. Bloking, Y. M. Chiang, and W. Y. Ching. Electronic structure and electrical conductivity of undoped $LiFePO_4$. *Electrochem. Solid State Lett.*, 7(6):A131–a134, 2004.

[207] A. Yamada and S. C. Chung. Crystal chemistry of the olivine-type $Li(Mn_yFe_{1-y})PO_4$ and $(Mn_yFe_{1-y})PO_4$ as possible 4 V cathode materials for lithium batteries. *J. Electrochem. Soc.*, 148(8):A960–a967, 2001.

[208] A. Yamada, S. C. Chung, and K. Hinokuma. Optimized $LiFePO_4$ for lithium battery cathodes. *J. Electrochem. Soc.*, 148(3):A224–a229, 2001.

[209] A. Yamada, M. Hosoya, S. C. Chung, Y. Kudo, K. Hinokuma, K. Y. Liu, and Y. Nishi. Olivine-type cathodes achievements and problems. *J. Power Sources*, 119:232–238, 2003.

[210] A. Yamada, H. Koizumi, S. I. Nishimura, N. Sonoyama, R. Kanno, M. Yonemura, T. Nakamura, and Y. Kobayashi. Room-temperature miscibility gap in Li_xFePO_4. *Nature Mater.*, 5(5):357–360, 2006.

[211] A. Yamada, H. Koizumi, N. Sonoyama, and R. Kanno. Phase change in Li_xFePO_4. *Electrochem. Solid State Lett.*, 8(8):A409–a413, 2005.

[212] N. A. Zarkevich and D. D. Johnson. Reliable first-principles alloy thermodynamics via truncated cluster expansions. *Phys. Rev. Lett.*, 92(25):255702, 2004.

[213] C. Zhang, S. Liu, Q. Q. Zhu, and Y. Q. Zhou. A knowledge-based energy function for protein-ligand, protein-protein, and protein-DNA complexes. *J. Med. Chem.*, 48(7):2325–2335, 2005.

[214] M. Zheng, A. Jagota, E. D. Semke, B. A. Diner, R. S. Mclean, S. R. Lustig, R. E. Richardson, and N. G. Tassi. DNA-assisted dispersion and separation of carbon nanotubes. *Nature Mater.*, 2(5):338–342, 2003.

[215] M. Zheng, A. Jagota, M. S. Strano, A. P. Santos, P. Barone, S. G. Chou, B. A. Diner, M. S. Dresselhaus, R. S. McLean, G. B. Onoa, G. G. Samsonidze, E. D. Semke, M. Usrey, and D. J. Walls. Structure-based carbon nanotube sorting by sequence-dependent DNA assembly. *Science*, 302(5650):1545–1548, 2003.

[216] F. Zhou, M. Cococcioni, K. Kang, and G. Ceder. The Li intercalation potential of $LiMPO_4$ and $LiMSiO_4$ olivines with M = Fe, Mn, Co, Ni. *Electrochem. Commun.*, 6(11):1144–1148, 2004.

[217] F. Zhou, M. Cococcioni, C. A. Marianetti, D. Morgan, and G. Ceder. First-principles prediction of redox potentials in transition-metal compounds with LDA + U. *Phys. Rev. B*, 70(23):235121, 2004.

[218] F. Zhou, G. Grigoryan, S. R. Lustig, A. E. Keating, G. Ceder, and D. Morgan. Coarse-graining protein energetics in sequence variables. *Phys. Rev. Lett.*, 95(14):148103, 2005.

[219] F. Zhou, K. S. Kang, T. Maxisch, G. Ceder, and D. Morgan. The electronic structure and band gap of $LiFePO_4$ and $LiMnPO_4$. *Solid State Commun.*, 132(3-4):181–186, 2004.

[220] F. Zhou, C. A. Marianetti, M. Cococcioni, D. Morgan, and G. Ceder. Phase separation in Li_4FePO_4 induced by correlation effects. *Phys. Rev. B*, 69(20):201101–201104, 2004.

[221] Fei Zhou, Thomas Maxisch, and Gerbrand Ceder. Configurational electronic entropy and the phase diagram of mixed-valence oxides: the case of Li_xFePO_4. *submitted*, 2006.

[222] A. Zunger. In P. E. Turchi and A. Gonis, editors, *NATO ASI on Statics and Dynamics of Alloy Phase Transformation*, volume 319, page 361. Plenum Press, New York, 1994.