Generating Conformations for Two Zinc-Binding Sites of HIV-1 Nucleocapsid Protein from Random Conformations by a Hierarchical Procedure and Polarizable Force Field

Nohad Gresh*,† and Philippe Derreumaux‡

Laboratoire de Pharmacochimie Moléculaire, CNRS FRE 2463 4, Avenue de l'Observatoire, 75006 Paris, France, and Information Génétique et Structurale, CNRS-UMR 1889, 31 Chemin Joseph Aiguier, 13402 Marseille, France

Received: December 5, 2002

As a first step toward searching for the most favorable folds of metallo-oligopeptides, we have developed a hierarchical procedure, which starting from random conformations, selects candidate conformers by a kinetic Monte Carlo approach and a potential of mean force and then postprocesses them using the polarizable SIBFA force field. The energy includes both first-order and second-order, polarization, charge-transfer, and dispersion components as well as a solvation free-energy term. In this work, we have focused on the 18-residue Zn-finger of the HIV-1 nucleocapsid protein with a CCHC core and its CCHH mutant. Whereas for the CCHC finger the simulations yield structures deviating by 3.5 Å rms from the targeted NMR conformation, rms deviations from 2.2 to 3 Å are obtained for the CCHH finger. As expected, the NMR conformation of the CCHC finger is the global minimum. Surprisingly, one predicted conformer of the mutant finger is of lower free energy than the NMR conformation. These results are, however, validated by quantum chemical computations bearing on fragments of the Zn-finger complexes encompassing up to 178 atoms. The prospects of improving our strategy for larger models are discussed.

Introduction

Attempts to predict the most stable conformations of proteins from random coils have resorted to a variety of algorithms. These include genetic-like algorithms,[1] Monte Carlo simulations[2,3] and all-atom molecular dynamics,[4] among others. Steady advances are reported on a 2-year basis at CASP meetings.[5] A promising study showed that a combination of Monte Carlo simulations from the assembly of fragments and all-atom molecular dynamics coupled with a continuum solvation model can lead to 2-Å rms (root-mean-square) deviation structure predictions for two small helical proteins with 36 and 65 residues.[6]

Surprisingly, de novo structure prediction of metalloproteins and metallopeptides by global optimization of an effective energy model remains unexplored. Reported molecular dynamics (MD) simulations of metalloproteins have sampled the conformational space starting from structures derived from experiment[7,8] or comparative modeling using homologues in the Protein Data Bank.[9] Yet, these proteins intervene in a number of essential functions including Ca^{2+}-mediated regulation, electron transport, and gene regulation by transcription activators,[10] or they are involved in misfunctions such as the prion protein that binds four Cu^{2+} ions.[11] Two reasons can be proposed to explain the lack of simulations to predict the equilibrium structure of metalloproteins from randomly chosen states.

The first reason is that metalloproteins such as iron−sulfur rubredoxin[12] and apo-plastocyanin proteins[13] do not obey the two-state folding model and fold on a much longer time scale than standard proteins. As a result, it is expected that convergence to their ground states is more difficult using current sampling procedures. The second reason relates to the very large nonadditivity effects in polycoordinated complexes of divalent cations. Energy-decomposition analyses show that these effects predominantly arise from the polarization and charge-transfer components.[14] Nonpolarizable molecular mechanics energy functions lacking such components may thus not be appropriate. This limitation holds if the cation−ligand distances are restrained because the polarization of the protein extends beyond the metal-binding core.

This study constitutes the first step toward de novo determination of the most stable conformations of metallopeptides from random coils. We have resorted to three procedures in succession. The first step is conformational sampling by a kinetic Monte Carlo approach using a simplified chain representation and a potential of mean force. This procedure has been shown to predict the equilibrium structures for several small proteins with 20−56 amino acids.[15−17] The second step is clusterization of the OPEP-MC structures generated at high temperature. The final step is the application of the polarizable molecular mechanics SIBFA force field to refine and rank the selected conformations. The polarizable SIBFA force field has successfully passed several comparisons with ab initio supermolecule computations for polycoordinated Zn^{2+} complexes,[14,19] multiply hydrogen-bonded complexes,[20] and drug-metalloenzyme complexes.[18,21,22] We focus here, among a database of 203 Zn-finger cores,[23] on the first Zn-finger of the HIV-1 nucleocapsid protein. This shortest finger determined thus far is characterized by a C X2 C X4 H X4 C pattern where C and H stand for Cys and His residues and X stands for any residue. Its conformation was determined by NMR spectroscopy within the protein[24,25] and in isolation.[26] The Cys protonation state of the CCHC Zn-finger being the subject of debate[27] and by consulting references therein, a de novo search for the best candidate conformers is also carried out for all three Cys residues deprotonated and for

* Corresponding author. E-mail: gresh@citi2.fr.
† CNRS FRE.
‡ CNRS-UMR.

10.1021/jp022527z CCC: $25.00
Published on Web 04/29/2003

each Cys residue protonated in turn. This allows us to determine whether protonation can occur at some preferred Cys position. We also extend our approach to the Cys28/His (CCHH) mutant so as to assess the efficiency of our approach versus the core sequence. The two tautomers, characterized by NMR[29] and deviating by 3.5 Å from the CCHC Zn-finger structure, are examined. In the first, denoted as "DD", both His residues bind Zn^{2+} by their Ndelta N atoms whereas in the second, denoted as "DE", His23 binds Zn^{2+} by its Ndelta and His28 binds Zn^{2+} by its Nepsilon.

To validate SIBFA calculations, we have performed density-functional theory (DFT) and Hartree−Fock (HF) computations on each of the candidate conformers using two simplified representations. The complete Zn-finger consisting of 279 atoms is not yet amenable to such quantum chemical (QM) computations. The two representations encompass the side chain−side chain plus side chain−Zn interactions (model Sc−Zn, 173 and 178 atoms for the CCHC and CCHH fingers, respectively) and the bare backbone (model M, 138 atoms). The organization of this paper is as follows. In section II, the hierarchical optimization procedure is presented, and its various components are reviewed. In section III, we give the results of the simulations for the two Zn-fingers and compare the energies of the predicted and NMR-derived structures using SIBFA and QM calculations, among others. In section IV, we summarize our results and discuss possible improvements of the method.

Methods

Hierarchical Method. The hierarchical procedure used for predicting the equilibrium conformations consists of the following three steps:

(1) Determination of the lowest-energy topologies by global optimization of a potential of mean force using a kinetic Monte Carlo (MC) procedure and a simplified chain representation.[15,16] Each zinc finger is subject to 10 simulations at 450 K for 15 000 MC steps starting from randomly chosen conformations. In this folding model, all main chain atoms are included; each side chain and ion is represented by a bead. New conformations are generated by diffusion process-controlled moves, which limit the transition time from the current state, and are subsequently accepted or rejected according to the Metropolis criterion[30] using the minimized OPEP energies of the current and new states. OPEP (optimized potential for efficient peptide structure prediction), parametrized for all 20 amino acids, is expressed by[15]

$$E_{\mathrm{OPEP}} = w_{\mathrm{H}}E_{\mathrm{HB1}} + w_{\mathrm{HH}}E_{\mathrm{HB2}} + w_{\mathrm{L}}E_{\mathrm{L}} + w_{\mathrm{SC}}E_{\mathrm{SC,SC}} + w_{\mathrm{A}}E_{\mathrm{C\alpha,C\alpha}} + E_{\mathrm{SC,M}} + E_{\mathrm{M,M}} + \sum_{20} w_{\mathrm{P}}^{\alpha}E_{\mathrm{P}}^{\alpha} + \sum_{20} w^{\beta}_{\mathrm{P}}E^{\beta}_{\mathrm{P}} + \sum_{20} w_{\mathrm{P}}{}^{\alpha_{\mathrm{L}}}E_{\mathrm{P}}{}^{\alpha_{\mathrm{L}}} \quad (1)$$

The *w*'s are the weigths of the energy components. E_{HB1} and E_{HB2} are the enthalpies of H-bond formation and the cooperative energy between H bonds, respectively. E_{L} is a quadratic term that maintains all bonds, bond angles, and improper dihedral angles near their equilibrium values. $E_{\mathrm{SC,SC}}$, the contact potential between two side chains separated by a distance r_{ij}, is represented by a 12−6 potential if the interaction is hydrophobic in character ($\epsilon_{ij} > 0$) and by a 6 potential if $\epsilon_{ij} < 0$.

$$E_{\mathrm{SC,SC}} = \epsilon_{ij}\left(\left(\frac{r^{0}_{ij}}{r_{ij}}\right)^{12} - 2\left(\frac{r^{0}_{ij}}{r_{ij}}\right)^{6}\right)H(\epsilon_{ij}) - \epsilon_{ij}\left(\frac{r^{0}_{ij}}{r_{ij}}\right)^{6}H(-\epsilon_{ij}) \quad (2)$$

$H(x) = 1$ if $x \geq 0$ and 0 if $x < 0$. $E_{\mathrm{C\alpha,C\alpha}}$ between two C_α atoms $j \geq i + 4$ is represented by a 12−6 potential. $E_{\mathrm{M,M}}$ and $E_{\mathrm{SC,M}}$ are the excluded-volume potentials of the main chain interactions and of main chain−side chain interactions, respectively. $E_{\mathrm{P}}{}^{\alpha}$, $E_{\mathrm{P}}{}^{\beta}$, and $E_{\mathrm{P}}{}^{\alpha_{\mathrm{L}}}$ are side-chain propensity terms of the 20 amino acids for the α helix, β strand, and α_{L} Ramachandran regions because of the neglect of detailed side chains.

In this work, the parameters for cysteinate residues are taken from cysteine residues.[15] The parameters (ϵ_{ij} in kcal/mol, r^{0}_{ij} in Å) are set to (5.7, 2.7) for Zn...Cys, (2.0, 3.3) for Zn...His, (5.4, 3.25) for Zn..Glu, and (1.5, $r^{0}_{ij} = (r^{0}_{i} + r^{0}_{j})/2$ with the van der Waals radii taken from ref 15 and $r^{0}_{j} = 0.2$ Å) for Zn for other side chains. Although this set of parameters is not optimized, the zinc fingers are stable at 300 K starting from their NMR structures by MC simulations.

(2) Clustering of the MC-generated conformations. Application of a rmsd threshold between conformations is not sufficient because the MC-generated states are both unfolded and folded at 450 K. It is thus necessary to consider a thermodynamic requirement. On the basis of earlier studies on small proteins, all conformations destabilized by an energy of 0−5 kcal/mol relative to the lowest-energy conformation are retained for analysis. The impact of this energy threshold is discussed in the last section. Two structures are clustered into the same ensemble if they differ by less than a 2.0-Å rms deviation from each other for the C_α atoms. As a result, five and three conformers are selected for the native and mutant zinc fingers, respectively.

(3) Postprocessing of the selected conformations by the polarizable SIBFA force field.[31,32] The intermolecular interaction energy, $\Delta E_{\mathrm{inter}}$, is computed as the sum of five terms:

$$\Delta E_{\mathrm{inter}} = E_{\mathrm{MTP}} + E_{\mathrm{rep}} + E_{\mathrm{pol}} + E_{\mathrm{ct}} + E_{\mathrm{disp}} \quad (3)$$

E_{MTP} is the electrostatic (multipolar) component, computed as a sum of multipole−multipole interaction terms. The multipoles (up to quadrupoles) are distributed on the atoms and bonds of the individual molecules or molecular fragments making up a larger molecule. They are derived from the ab initio SCF molecular wave function using the procedure developed by Vigné-Maeder and Claverie.[33]

E_{rep} is the short-range repulsion energy. To account for its anisotropic character, it is computed as a sum of bond−bond, bond−lone pair, and lone pair−lone pair interactions. The formulation of E_{rep} takes into account the explicit hybridization nature of the bonds, in addition to that of the lone pairs. The expression of E_{rep} was detailed in refs 20 and 32 for the general case and for ligand−cation interactions, respectively.

E_{pol} is the polarization energy component calculated with distributed anisotropic polarizabilities on the individual molecular fragments. The polarizabilities are distributed on the centroids of the localized orbitals (heteroatom lone pairs and bond barycenters) using the procedure of Garmer and Stevens.[34] A Gaussian screening of the polarizing field is used. (See refs 20 and 32 for details.) The fragment multipoles and polarizabilities are derived from ab initio computations using the CEP 4-31G(2d) basis set[35] and are supplemented with two diffuse 3d orbitals on heavy atoms. The ab initio computations were made with the GAMESS software.[36] E_{ct} is the charge-transfer contribution. The numerator of E_{ct} is a function of the overlap between the lone-pair hybrids of the electron donor and the electron acceptor, and the denominator takes into account the difference between the ionization potential $I_{\mathrm{L}\alpha}$ of the electron donor and the electron affinity A_{β}* of the electron acceptor. $I_{\mathrm{L}\alpha}$ is increased by the predominantly positive electrostatic potential exerted on this atom by all of the other molecules in the complex

whereas A_β* is reduced by the predominantly negative electrostatic potential due to its surrounding ligands. The dependency of the denominator upon the electrostatic potential and field effects was found to be essential in accounting for the strong nonadditive character of E_{ct} in polyligated Zn^{2+} complexes.[14] The detailed expression for E_{ct} was given in refs 20 and 32.

E_{disp} is the dispersion component developed using the formulation of Creuzet et al. and expressed as a sum of $1/R^6$ $1/R^8$, and $1/R^{10}$ terms at short distances by means of exponential damping terms.[37] An explicit exchange-dispersion term is introduced for the mutual interactions between polyatomic molecules. Directionality effects are accounted for by the explicit introduction of fictitious atoms with reduced van der Waals radii to represent the lone pairs.

We use the same method as in the original derivation of the SIBFA procedure[31] to assemble a large molecule from its constitutive fragments, whose internal geometries and distributed multipoles and polarizabilities are available in a library. Briefly, for oligopeptides, the backbone is made out of successive formamide and methyl groups having the original multipolar distribution of *N*-methylformamide; each side chain is built from methyl groups and its specific end side-chain moiety (imidazole for His, formate for Asp and Glu, methylammonium and methylguanidinium for Lys and Arg, respectively, etc.). A junctional bond such as C−C or C−N between two connecting fragments is created by overlapping their corresponding end bonds (e.g., CH−HC or CH−HN) upon imposing the appropriate C−C or C−N bond lengths. The multipoles belonging to the vanished H atoms and C−H or N−H bond barycenters are redistributed on the two atoms and the midpoint of the newly created junctional bond, thus preserving the net charge on the molecule. Standard bond lengths and bond angles are used.

The intramolecular energy (E_{intra}) is the sum of the interaction energy between the fragments (ΔE_{intra}) and the torsional (E_{tor}) and bond angle ($E_{bending}$) energies.

$$E_{intra} = \Delta E_{intra} + E_{tor} + E_{bending} \quad (4)$$

ΔE_{intra} uses the formulation of ΔE_{inter}. A 3-fold torsional term (E_{tor}) along the main-chain and side-chain C−C bonds, with a barrier of 2.3 kcal/mol so as to reproduce the rotation barrier in ethane, is used.[31] All other torsional barriers are set to 0 kcal/mol. A quadratic bending term characterized by a force constant of 50 kcal/(mol deg^2) and an equilibrium value of 109.5° is defined for the $C_\alpha-C_\beta-C_\gamma$ angle of histidine residues. This term facilitates the position of Zn^{2+} in the plane of the imidazole ring (trans Zn−N−C−N dihedral angle) during minimization.

The very strong nonadditive character of E_{pol} and E_{ct}, which was analyzed for several polycoordinated Zn(II) complexes,[14] requires a simultaneous and consistent computation of intra- and intermolecular E_{pol} and E_{ct} components. This was recently applied to the study of flexible polyligated complexes of Zn(II) and tested by comparisons with parallel ab initio computations.[22,38] In the present study, we have further elaborated on this procedure following tests on the conformational energies of the alanine tetrapeptide and validation by ab initio results (Gresh et al., manuscript in preparation). Thus, although the multipoles are redistributed at the junctions between fragments to compute E_{MTP}, the present procedure relocates the charge of the junctional H atoms on their attached C or N atoms to calculate E_{pol}. This prevents the fragments from being too close to the H atoms ($\leq$1 Å) and having a fractional charge following multipole redistribution that could locally exaggerate the polarization component. Similar to the procedure used in ref 38, the polarizabilities of the collapsed CH or NH bonds, whose magnitudes are smaller than those of the lone pairs and exposed bonds of the peptide, were set to zero to prevent a divergence of E_{pol} due to the effects of the induced dipoles. The intramolecular energies of the finger and the Zn(II)−finger interactions are computed as one global "intramolecular" interaction. Such an approach was recently tested and applied to novel organometallic complexes of Cu(II), some of which can act as HIV-1 inhibitors (Ledecq et al., submitted for publication). We note that the need for a simultaneous and consistent computation of inter- and intramolecular energies was recently underlined.[39]

The 1999 version of the SIBFA software is available from the computational chemistry list at http://www.ccl.net/cca/software/SOURCES/FORTRAN/sibfa/index.shtml. It includes the library of fragments necessary to construct large oligopeptides and oligonucleotides, along with sample input files. An updated version enabling the simultaneous computation of inter- and intramolecular interactions as a single total energy, which we also plan to release, is presently available upon request from one of us (N.G.). The present approach is deemed preferable to the previous ones. The parameters used for E_{rep}, E_{ct}, E_{pol}, and E_{disp} are given as Supporting Information. They were originally derived in ref 32.

We have proceeded along the following steps (3a−3e) for the representative MC-generated structures. Energy minimization was carried out using the Merlin package.[40]

3a. Construction of optimized models with the bond lengths and bond angles of SIBFA and all nonglycine residues replaced by alanine residues. Such models fit the OPEP-MC backbones within a rmsd of 0.2 Å.

3b. Construction of all-atom models using the base mean-field theory to optimize side-chain packing.[41]

3c. Minimization with the backbone frozen and the side-chain torsional angles and the Zn^{2+} ion free to move under five constraints. These harmonic constraints include the four Zn−ligand distances set to 2.1 and 2.3 Å for Zn−N and Zn−S, respectively (force constant of 100 kcal/(mol Å^2) and the Zn−N−C−N dihedral angle (forced to be trans, energy barrier of 25 kcal/mol). These constraints facilitate the position of Zn^{2+} with respect to the Cys/His side-chains upon passing from a simplified model to an all-atom model.

3d. Unconstrained minimization with all variables free to move. E_{ct} is the most time-consuming term of E_{intra}. Since this term has a very shallow radial dependence,[32] it was not taken into account at step 3d but rather was included at step 3e for the converged minima.

3e. Postprocessing of the previous minima by energy minimization in the presence of solvation. This is done by including the solvation free energy (ΔG_{solv}) in the intramolecular energy. ΔG_{solv} is computed using the continuum model of Langlet−Claverie.[42] A grid of 233 Korobov points around each solute atom was used. At the converged minima, a higher-resolution grid of 633 points was also used. The energy difference upon using the denser grid amounts to ca. 10 kcal/mol out of 350 kcal/mol and remains fairly constant between the various conformers. The integrated SIBFA/continuum procedure[43] was used to study the binding of inhibitors to mono- and binuclear Zn-metalloenzymes.[18,21,22] Within this procedure, ΔG_{solv} is computed as a sum of four components: cavitation, solute−solvent electrostatic interactions (E_{el}), solute polarization (E_{pol}), and dispersion−repulsion. The E_{el} and E_{pol} contributions are computed with the distributed multipoles used by E_{MTP} and E_{pol} in SIBFA, which should ensure consistency between both approaches.

TABLE 1: SIBFA Energies (kcal/mol) and Zn^{2+} Coordination (Å) of the NMR-Minimized and the Five Predicted (a−e) Conformations for the All-Atom CCHC Zn-Finger

	NMR-min	a	b	c	d	e
E_{MTP}	−922.9	−840.7	−869.2	−893.7	−888.9	−897.9
E_{rep}	1111.6	1097.3	1102.3	1147.3	1089.1	1098.9
$E_1 = E_{MTP} + E_{rep}$	188.8	256.6	233.0	253.6	200.2	201.0
E_{pol}	−148.3	−155.2	−164.4	−144.3	−144.0	−138.6
E_{ct}	−53.0	−53.1	−52.0	−52.4	−49.8	−55.1
E_{disp}	−599.5	−562.4	−581.2	−601.2	−571.3	−583.8
$E_{bending} + E_{tor}$	38.2	35.4	30.4	33.1	37.1	34.2
E_{intra}(SIBFA)	−573.9	−478.8	−534.2	−511.2	−527.8	−542.3
ΔG_{solv}	−340.9	−379.9	−365.0	−352.9	−356.0	−339.6
$\boldsymbol{E_{intra} + \Delta G_{solv}}$	**−914.8**	**−858.7**	**−899.2**	**−864.1**	**−883.8**	**−881.9**
$E_{intra} + \Delta G_{solv}$[a]	−713.5	−650.4	−682.8	−667.4	−690.0	−688.2
Zn coordination						
Cys15	2.50	2.45	2.47	2.50	2.61	2.57
Cys18	2.48	2.51	2.47	2.54	2.53	2.57
His23	2.23	2.21	2.44	2.40	2.27	2.42
Cys28	2.50			2.46	2.50	2.57
O-Gly22			2.01			

[a] E_{intra} is calculated without the nonadditive E_{pol} and E_{ct} terms.

Validation by Parallel Ab Initio Computations. Each conformer of the CCHC and CCHH fingers was subjected to DFT computations using the models Sc−Zn and M. They were carried out using the B3LYP functional.[44] The size of the investigated complexes precludes the use of the MP2 procedure to account for correlation effects. Resorting to DFT is justified by the results of our previous studies of polyligated Zn complexes[19] and hydrogen-bonded complexes,[45] which showed a close reproduction of MP2 interaction energies by DFT. Computations were run using Gaussian 98[46] and Jaguar[47] packages. The CEP 4-31G(2d)[35] and LACVP**[48] basis sets were tested on the five MC-predicted structures and the NMR-minimized structure of the CCHC finger using a simplified representation. Table 1S shows that the in vacuo SIBFA energies follow the LACVP** energies more closely than the CEP 4-31G(2d) energies, with relative errors of 3 and 5%, respectively, in agreement with the results on DL-captopril−metallo-betalactamase models.[22] As a result, the next DFT calculations will be limited to the LACVP** set. The solvation free energy of all models using the continuum reaction field was also compared to that obtained by a self-consistent reaction-field method for solving the Poisson−Boltzmann equation.[49] An important parameter in the continuum reaction field is the λ multiplicative factor of the van der Waals surface. $\lambda = 1.0$ was previously used,[18] but it exaggerates the solvation enthalpies of complexes having a Zn^{2+} cation partially exposed to the continuum (unpublished results). Table 1S shows that $\lambda = 1.2$ provides a better reproduction of ΔG_{solv} determined by DFT/Delphi than 1.0 and 1.1. $\lambda = 1.2$ will be used in the calculations discussed below.

Results and Discussion

A. Native Zn Finger. 1. SIBFA Calculations. Figure 1 superposes the predicted structure of lowest SIBFA energy (conformer b) and of lowest rmsd (conformer d) on the NMR-minimized conformation. Table 1 lists the in vacuo E_{intra} energy and its components (E_1 denoting the sum of E_{MTP} and E_{rep} and E_2 denoting the sum of E_{pol} and E_{ct}), the solvation free energy ΔG_{solv} using the Langlet−Claverie procedure, and the Zn−ligand distances of the five predicted and NMR conformations. Table 2 is the rms deviation matrix between each conformer.

As expected, the NMR-derived conformation is of the lowest free energy, but the predicted conformer (b) is destabilized by only 15 out of 900 kcal/mol. This conformer is characterized by a fourth Zn ligand with the oxygen atom of Gly22 (vs the S atom of Cys28 in the NMR structure). Conformers d, e, and c, which retain the tetrahedral Zn-binding coordination, come next

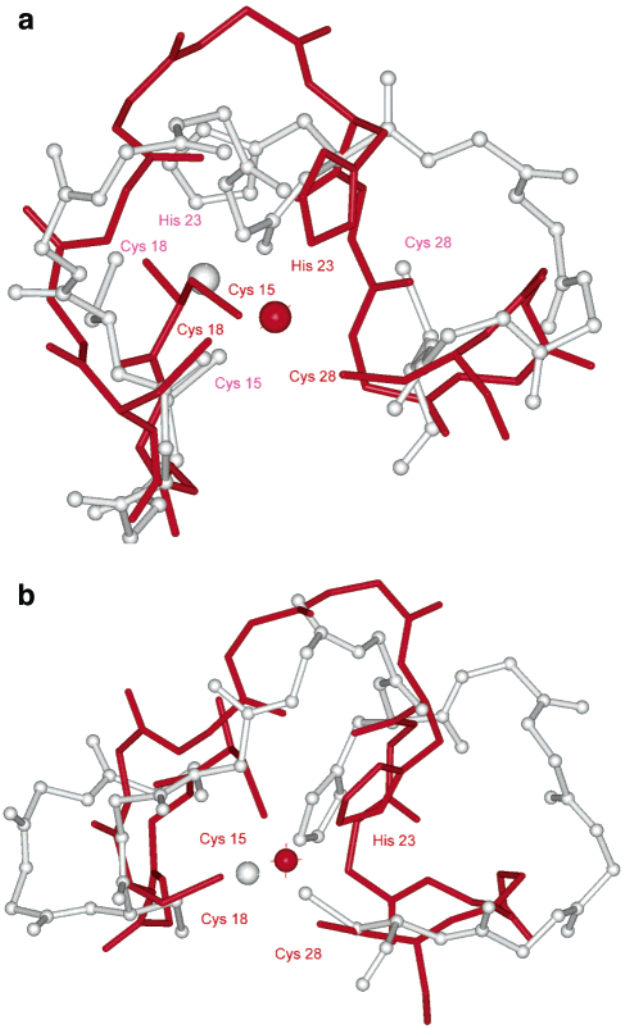

Figure 1. Native CCHC Zn-finger. Superposition of the predicted conformations with (a) the lowest SIBFA energy (conformer b) and (b) the lowest rmsd (conformer d) on the NMR structure. The Cys/His residues are represented at the all-atom level; the others are represented by their main chain atoms. The predicted structures are in gray, and the NMR is in red.

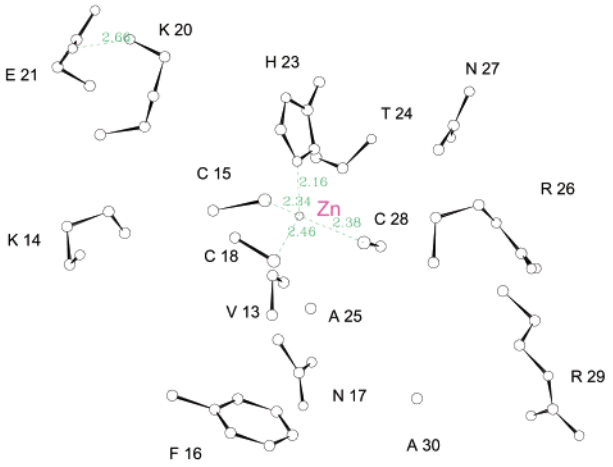

Figure 2. Native Zn-finger. The model Sc−Zn for the NMR-minimized structure used to compare DE_{inter}, DE_{intra} and ΔG_{solv} using SIBFA and DFT.

TABLE 2: CCHC Zn-Finger rms Deviation Matrix (Å) between the NMR, the NMR-Minimized, and the Five Predicted a−e Conformations[a]

	NMR	NMR-min	a	b	c	d	e
NMR	0.0	1.3	3.8	3.8	3.7	3.5	4.3
NMR-min		0.0	3.7	3.8	3.9	3.4	4.1
a			0.0	3.2	3.7	2.5	3.6
b				0.0	2.5	3.3	3.4
c					0.0	3.7	4.1
d						0.0	1.9
e							0.0

[a] The rmsd is calculated using the α-carbon atoms.

and are destabilized by 31, 33, and 51 kcal/mol, respectively. In the absence of solvation, these energy differences are enlarged, reaching 46, 32, and 63 kcal/mol. Conformer a, which has only a 3-fold Zn coordination, is highly destabilized, although it has the most favorable solvation free energy. In the absence of the nonadditive E_{pol} and E_{ct} terms, the NMR structure is still ranked first, but the ordering of the predicted structures changes: b is ranked fourth, and d is ranked second. The NMR-minimized conformation deviates by 1.3 Å rms from the NMR conformation whereas conformers a−e deviate by 3.5−4.3 Å and 3.4−4.1 Å rms from the NMR and NMR-minimized structures, respectively. Most predicted conformers are more expanded than the NMR structure. The C_α(Cys15)−C_α(Cys28) distance is 6.4 Å in the NMR conformation and conformer b, within 8.0−9.0 Å in c−e, and 13 Å in a.

To evaluate the sensitivity of the results, we have determined the free energy of an NMR-like conformer with the side chain of Lys20 perturbed from its NMR position (side chain dihedral angle X1 from g+ to t, X2 from t to g+, and X3 from g− to t after mimimization). The resulting conformation is destabilized by 4.2 kcal/mol with respect to the NMR conformation (E_{intra} = −568.8, ΔG_{solv} = −341.8). This result illustrates the difficulty of metallopeptide structure predictions because the energy of the slightly perturbed NMR conformation comes close to that of conformer b.

2. Quantum Chemical Calculations. The model Sc−Zn for the NMR-minimized structure is shown in Figure 2. The in vacuo energies of the five predicted and NMR-minimized conformations are compared using SIBFA, DFT, and HF in Table 3. The solvation free energies ΔG_{solv} are also compared to their Delphi counterparts using DFT.

We first analyze the SIBFA and QM intermolecular interaction and solvation free energies. The SIBFA ΔE_{inter} energies

TABLE 3: CCHC Zn-Finger SIBFA and DFT ΔE_{inter}, ΔG_{solv}, and E_{intra} Energies of the NMR and the Predicted Conformations Using Models Zn−Sc and M[a]

	Model Sc−Zn					
	NMR-min	a	b	c	d	e
E_1	−663.9	−646.2	−594.9	−662.7	−630.2	−653.8
E_{pol}	−88.6	−109.5	−107.3	−98.2	−93.0	−85.3
E_{ct}	−30.2	−37.9	−36.9	−31.0	−31.1	−32.6
E_{disp}	−105.8	−77.9	−75.4	−103.2	−92.1	−85.1
ΔE_{inter}(SIBFA)	**−888.6**	**−871.6**	**−814.5**	**−895.2**	**−846.5**	**−856.8**
ΔE_{inter}(DFT)	**−865.6**	**−861.6**	**−805.2**	**−880.6**	**−841.0**	**−850.6**
ΔG_{solv}(SIBFA)	−305.2	−308.6	−354.2	−283.6	−313.7	−316.3
ΔG_{solv}(DFT)	−253.9	−256.1	−289.9	−229.5	−264.8	−250.8
	Model M					
E_{intra}(DFT)	3.1	21.3	0.0	28.8	35.1	10.0
E_{intra}(SIBFA)	28.0	27.2	0.0	24.5	52.0	25.8
E_{intra}(AMBER)	15.6	22.0	0.0	14.7	33.6	10.4
E_{intra}(Cff91)	28.3	34.5	0.0	27.9	57.4	23.8
E_{intra}(HF)	0.0	21.5	0.0	28.3	34.8	9.1
E_{intra}(SIBFA/ without E_{disp})	23.1	24.3	0.0	34.0	40.3	23.8
E_1(SIBFA)	17.2	24.7	0.0	32.2	30.6	19.1
E_{MTP}(SIBFA)	28.2	14.4	0.0	−1.7	40.7	22.9
$E2$(SIBFA)	−0.9	1.1	0.0	−1.3	1.3	2.1
E_{disp}(SIBFA)	6.5	2.6	0.0	−8.8	13.6	2.5
E_{tor}(SIBFA)	5.3	−1.1	0.0	2.3	6.4	2.0

[a] The energies, using model M, are expressed with respect to conformer b.

closely follow the DFT results (within an error of 2.7%), and both methods rank the six conformers in the same order. This accuracy is encouraging, considering the size of the models (173 atoms) and the onset of differing interactions. As expected, there may be no energy correlation between these models and the corresponding all-atom zinc fingers. The NMR-minimized structure and conformer c, which have the correct CCHC Zn^{2+} coordination, are ranked second and first, respectively, versus first and fifth for the all-atom models. Conformer a, which is only 3-fold coordinated to Zn, has a more favorable ΔE_{inter} than 4-fold-coordinated conformers d and e because of its lower polarization energy. An increase in E_{pol} upon decreasing the coordination number of Zn^{2+} was previously observed.[21,14] As observed in Table 1S, the SIBFA-generated ΔG_{solv} energies are larger than those from Delphi/DFT, but the relative solvation free energies between conformers are independent of the method used, except for conformer e.

Next, we compare the SIBFA and QM intrabackbone interaction energies. As seen in Table 3, DFT and HF conformational energies are almost identical whereas SIBFA energies are overestimated. This is most evident for the NMR-minimized structure and predicted conformers d and e with respect to conformer b: 28, 52, and 25.8 kcal/mol using SIBFA versus 3.1, 35, and 10 kcal/mol using DFT, respectively. Such a finding is somewhat surprising in view of the close agreement between SIBFA and DFT results for the CCHH Zn-finger (see next section) and between SIBFA and 6-311G** DFT conformational energies for 10 representative conformers of the alanine tetrapeptide (Gresh et al., manuscript in preparation) used by Beachy et al.[50] to benchmarch molecular mechanics potentials against ab initio computations. The results from single-point computations using Discover sotfware[51] and the nonpolarizable AMBER[52] and Cff91[53] force fields show that AMBER closely follows DFT for conformers a, d, and e but overestimates the energy of the NMR conformation by 12 kcal/mol and underestimates the energy of conformer c by 14 kcal/mol. In contrast, Cff91 gives energies consistent with SIBFA and thus overestimates the energies with respect to DFT.

TABLE 4: Total Energies (kcal/mol) and Zn^{2+} Coordination (Å) of the Native Zn-Finger Protonated on Cys15, Cys18, and Cys28[a]

	CysH15	CysH18	CysH28
E_{MTP}	−709.3	−748.0	−728.1
E_1	397.8	380.5	384.6
E_{pol}	−172.8	−168.6	−169.4
E_{ct}	−50.5	−49.5	−51.0
E_{disp}	−601.6	−614.1	−607.2
$E_{bending} + E_{tor}$	34.1	32.9	31.5
E_{intra}	−392.9	−418.5	−411.8
ΔG_{solv}	−449.1	−428.0	−437.2
$\boldsymbol{E_{intra} + \Delta G_{solv}}$	**−842.0**	**−845.5**	**−849.0**
Zn coordination			
Cys15	2.96	2.42	2.48
Cys18	2.34	2.47	2.40
His23	2.19	2.16	2.15
Cys28	2.36	2.37	2.53

[a] The NMR-minimized structure is used.

Why then does SIBFA overestimate the intrabackbone energy differences? The E_{intra} energies computed at the HF level and by SIBFA in the absence of E_{disp} are shown in Table 3. Energy overestimation is still observed for the NMR and e conformers, but the absence of E_{disp} allows a better fit of conformer d. The E_{MTP}, E_1 ($E_{MTP} + E_{rep}$), E_2, and E_{tor} energies of all conformers with respect to the energies of conformer b are also reported in Table 3. We see that although E_{tor} cannot be neglected E_{MTP} is responsible for the high ΔE_{intra}(SIBFA) value of the NMR-derived and e conformers. This is consistent with an earlier study on a bispyridine entity in conformations eclipsing two successive C−N bonds.[54] The inclusion of an explicit and attractive penetration component to E_{MTP}[55] was proposed. This could be necessary for intramolecular computations in which the interacting fragments are not free to optimize their mutual orientation of approach.

B. Native Zn Finger with One Cys Protonated. The protonation state of one Cys residue in the Zn-binding Cys3His motif is a subject of debate. Maynard and Covell, using protein packing and electrostatic screening studies, have suggested extensive shielding of the first Zn-finger core of NCp7 and stabilization of a state having all three Cys residues deprotonated.[23] This deprotonation state is confirmed by quantum chemical computations[56−58] and X-ray crystallography measurements of Zn−S bond distances,[59] but recent mass spectrometric measurements are more in line with one protonated Cys residue.[60] Furthermore, QM and ΔG_{solv} computations of a Zn^{2+} cation tetrahedrally bound to four ligands ($n = 0-3$ ethaniols, $3-n$ ethylthiolates, and 1 ethylimidazole have underlined the need to consider, at least transiently, one protonated Cys residue)[27] and ab initio/continuum dielectric calculations of small Zn−Cys4 complexes have shown that a high dielectric medium favors cysteine deprotonation whereas a low one favors the protonated state.[28]

This has prompted us to determine whether protonation can be differentiated energywise. To this end, we evaluate the energy of the three isomers in which either Cys15, Cys18, or Cys28 is protonated using the NMR-derived conformation as a starting point. As expected, all three isomers deviate by 0.2 Å rms from the native NMR-minimized conformation. However, we find that the Zn−Cys15 distance increases by 0.5 Å upon Cys15 protonation. Table 4 shows that the isomers CysH18 and CysH28 have similar in vacuo energies but that CysH28 has the lowest free energy. CysH15 and CysH18 are destabilized by 7.1 and 3.4 kcal/mol, respectively, with respect to CysH28. These differences are nevertheless relatively small, and the energy change between CysH18 and CysH28 is equivalent to the cost of one side-chain positional change.

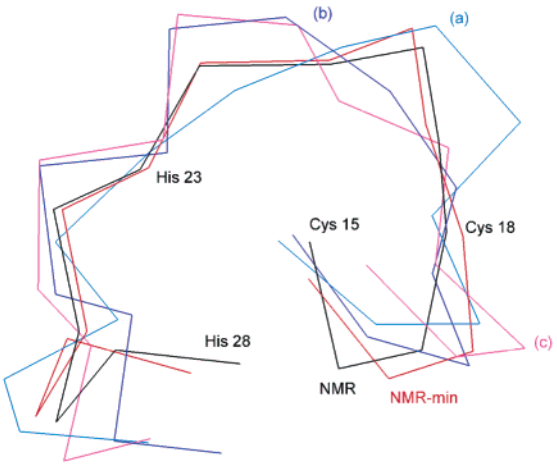

Figure 3. Mutant CCHH Zn-finger. Superposition of the three MC-generated (a−c) conformations on the NMR and NMR-minimized trace for the DD tautomer. The location of the Cα atoms of Cys/His residues in the NMR structure is shown.

Figure 4. Mutated CCHH Zn-finger. Superposition of the lowest SIBFA energy predicted conformer (grey) on the NMR (red) structure. All-atom representation for Cys/His residues and main chain representation for others.

C. Mutated Zn Finger CCHH. The three conformers (a−c) selected from the OPEP-MC procedure are superposed on the NMR conformation of the DD tautomer[62] in Figures 3 and 4. The total energy and Zn^{2+} coordination of each conformation are reported in Table 5. The rms deviations between the conformers are reported in Table 6. Conformer b retains the NMR 4-fold coordination whereas conformer a has Zn^{2+} ligated to Oe1 of Glu21 versus Ndelta of His28 by NMR, and conformer c has a Zn^{2+} 3-fold coordination.

One unexpected result emerges from the free energy of the four conformers. The NMR-minimized structure is ranked third, and predicted conformers b and c are more stable by 42 and 22 kcal/mol, respectively. Yet, the NMR-minimized structure deviates by 1 Å rms from the NMR structure (vs 1.3 Å for the CCHC finger) whereas the three predicted conformers a−c deviate by 2.4−3.1 and 2.2−3.0 Å rms from the NMR and NMR-minimized structures, respectively (vs 3.5−4.3 Å for the

TABLE 5: SIBFA Energies (kcal/mol) and Zn^{2+} Coordination (Å) of the NMR-Minimized and the Three Predicted (a−c) Conformations for the All-Atom Mutated Zn-Finger in the DD Tautomer

	NMR-min	a	b	c
E_{MTP}	−738.7	−803.3	−832.0	−769.0
E_{rep}	1110.1	1183.7	1111.7	1136.5
$E_1 = E_{MTP} + E_{rep}$	371.4	380.6	279.6	367.4
E_{pol}	−202.6	−191.0	−185.7	−207.2
E_{ct}	−41.0	−42.9	−46.5	−50.6
E_{disp}	−594.0	−604.2	−592.9	−580.3
$E_{bending} + E_{tor}$	23.6	33.2	34.3	26.8
E_{intra}	−442.7	−424.4	−511.2	−443.8
ΔG_{solv}	−421.4	−416.6	−395.2	−442.7
$\boldsymbol{E_{intra} + \Delta G_{solv}}$	**−864.1**	**−840.9**	**−906.4**	**−886.4**
$E_{intra} + \Delta G_{solv}$[a]	−620.5	−607.0	−674.2	−628.6
Zn coordination				
Cys15	2.31	2.36	2.33	2.41
Cys18	2.21	2.28	2.31	2.22
His23	2.18	2.27	2.18	2.15
His28	2.10		2.18	
Glu21		1.90		

[a] E_{intra} is calculated without the nonadditive E_{pol} and E_{ct} terms.

TABLE 6: Mutated Zn-Finger rms Deviation Matrix between the Predicted and NMR Structures for Tautomers DD and DE[a]

	NMR	NMR-min	a	(b)	(c)
NMR	0.0	1.1(1.8)	2.4(2.3)	2.8(2.9)	3.1(3.2)
NMR-min		0.0	2.2(2.6)	3.0(3.2)	2.8(2.4)
(a)			0.0	2.6(2.5)	2.9(2.8)
(b)				0.0	2.7(2.9)
(c)					0.0

[a] In parentheses.

native finger). The NMR-minimized structure is also found to be less stable than 3-fold coordinated conformer c. As seen in Table 5, the E_1 ($E_{MTP} + E_{rep}$) component is responsible for ranking conformer b first, and ΔG_{solv} is critical for ranking conformer c second. E_1 is 280 kcal/mol for b versus ∼370 kcal/mol for the others; E_{intra} is ca. −440 kcal/mol for the NMR and c conformations. The absence of the nonadditive E_{pol} and E_{ct} terms does not change the ordering of the conformers but increases the energy differences by 12 kcal/mol between b and NMR and by 25 kcal/mol between b and c.

It was therefore essential to validate our results by DFT calculations. The model Sc−Zn for the NMR-minimized structure is shown in Figure 1S. As seen in Table 7, the in vacuo E_{inter} SIBFA energies reproduce the DFT energies with a relative error of ≤3% and still rank conformer b first and the NMR-minimized structure third. The role of the E_1 component in favoring b is evident: −530 versus −495 kcal/mol for NMR. The importance of second-order components, particularly E_{pol}, is also highlighted. Although conformer b is destabilized by E_1 with respect to conformer a (−530.5 vs −541.8 kcal/mol), b is favored by E_{pol} (−105.1 vs −79.0 kcal/mol). This is consistent with recent ab initio and SIBFA studies of polycoordinated Zn^{2+} complexes that show that E_{pol} favors complexes having two neutral and two anionic ligands over complexes having one neutral and three anionic ligands.[14]

The intrabackbone interaction energies from various methods are compared in Table 7. DFT, HF, SIBFA, AMBER, and Cff91 rank the four conformers in the same order. In all cases, the NMR-minimized structure has a much higher E_{intra} energy than a−c, yet the resulting structure satisfies all PROCHECK stereochemical verifications.[61] The decomposition of E_{intra} into its components reveals the essential role of the electrostatic multipolar energy in destabilizing the NMR conformation.

TABLE 7: CCHH Zn-Finger SIBFA and DFT DE_{inter}, ΔG_{solv}, and E_{intra} Energies of the NMR and the Predicted Conformations Using Models Zn−Sc and M[a]

	NMR-min	a	b	c
Model Sc−Zn				
E_1	−495.6	−541.8	−530.5	−494.0
E_{pol}	−114.2	−79.0	−105.1	−117.1
E_{ct}	−26.6	−24.8	−29.1	−34.3
E_{disp}	−94.6	−87.6	−88.3	−79.2
$\boldsymbol{\Delta E_{inter}}$**(SIBFA)**	**−731.1**	**−733.1**	**−753.1**	**−724.7**
$\boldsymbol{\Delta E_{inter}}$**(DFT)**	**−709.8**	**−717.1**	**−727.5**	**−704.1**
Model M				
E_{intra}(DFT)	51.4	15.1	8.7	30.0
E_{intra}(SIBFA)	46.9	23.3	15.1	28.2
E_{intra}(AMBER)	33.7	20.3	10.2	30.2
E_{intra}(CFF91)	67.0	27.3	21.9	39.5
E_{intra}(HF)	56.4	16.9	6.9	31.9
E_{intra}(SIBFA/ without E_{disp}	37.0	18.3	10.3	20.6
E_1(SIBFA)	44.5	13.9	3.0	16.4
E_{MTP}(SIBFA)	69.8	12.3	11.2	11.2
E_2(SIBFA)	−3.1	2.0	3.2	2.1
E_{disp}(SIBFA)	10.0	5.0	4.8	7.7
E_{tor}(SIBFA)	−3.6	2.4	4.2	2.0

[a] The energies using model M are expressed with respect to conformer b.

To estimate the dependency of relative free energies as a function of the tautomeric state, we repeated the SIBFA calculations for the DE tautomer. To do so, we have used the same multipolar expansion on the imidazole ring as in the DD tautomer calculation. Such an approximation is not valid for computing the relative stability of DD versus DE since the C_β carbon should exert a different induction effect whether ortho to a protonated N (as in the Nepsilon tautomer) or ortho to a deprotonated N (as in the Ndelta tautomer). But this approximation is sufficient, however, to compare the energy of the NMR and predicted conformations. Because of uncertainties in the NMR structure of the DE tautomer,[62] we have retained the experimentally derived DD conformation as a reference. The energy of the three predicted and reference conformations are reported in Table S2 for the DE tautomer. Table 6 shows that the rms deviation between the predicted and experimental structures is independent of the tautomeric state; all rms deviations are ≤3.2 Å. Conformers b and c have the CCHH 4-fold Zn^{2+} coordination. Conformer a, which has a fifth coordination to Oe1(Glu21), is highly destabilized. We now observe that the NMR-minimized conformation becomes second best along the energy scale, being destabilized by 13 out of 900 kcal/mol with respect to b. The solvation free energy is critical to the reduction of the energy difference between the NMR-derived conformer and b, from 55 kcal/mol in vacuo to 15 kcal/mol in solvent.

Conclusions and Perspectives

The aim of this study was to determine a strategy that allows accurate structural and energetics predictions of metallopeptides starting from random coils. This strategy integrates two distinct procedures and force fields. In this work, we have focused on the 18-residue Zn-finger of the HIV-1 nucleocapsid protein with a CCHC core and its CCHH mutant.

As expected, the NMR-mimimized conformer of the CCHC Zn-finger is the global free-energy minimum, but the MC-generated conformer of lowest free energy is destabilized only

by 15 out of 900 kcal/mol. The SIBFA energies were validated by DFT results on simplified models with a relative error of less than 3% for the side chain−side chain and side chains−Zn interactions. ΔG_{solv} is found to be critical in reducing the in vacuo-computed energy differences between NMR and this candidate, from 37 in vacuo to 13 kcal/mol in solvent. The C_α rms deviations of conformers a−e from the NMR-minimized structure vary from 3.4 to 4.1 Å.

The protonation of each Cys residue was also considered within the Cys3His zinc finger. We find that the least stable is CysH15, followed by CysH18 and CysH28. This ordering of protonation difficulties is consistent with the positive screening potentials of the protein on the three unprotonated cysteines (CysH15 > CysH18 > CysH28) as determined by Maynard and Covell.[23] Their *kT*/e values lead to 4.8 and 1.8 kcal/mol between the successive forms whereas we find 3.7 and 3.4 kcal/mol, respectively. Although the SIBFA free-energy differences are small, these results imply that, if protonation occurs in solution, protonation would select one particular cysteine side chain rather than spending part of its time on each cysteine side chain.[23,27]

The three predicted conformations of the CCHH Zn-finger deviate by 2.3 to 3.2 Å from the NMR-minimized structure. Independently of the tautomeric state, predicted conformer b retaining the standard Zn^{2+} 4-fold coordination is found to have a greater stability than the NMR conformation. However, this unexpected result was validated by DFT calculations. On account of this finding, the NMR observables are being reexamined, and conformer b is being used as an additional starting point for restrained MD.

Current generation force fields do not treat polarization and charge-transfer effects explicitly, although they are believed to be important for understanding ion selectivity in channels and improving the accuracy of structure-based drug design calculations. We have shown that the absence of the nonadditive E_{pol} and E_{ct} terms changes the ranking of the selected conformers and increases the energy difference between the candidates for both Zn-fingers. The neglect of explicit polarization and charge-transfer interactions is also found to increase the relative error between the SIBFA and DFT ΔE_{inter} energies of model Zn−Sc from 2 to 3% to ∼15%. These results show that standard potentials of mean force (such as OPEP) cannot accurately reproduce the energetics of the two ionic−metallic systems studied here.

The present OPEP-MC/SIBFA procedure can be improved in many respects. Efficient sampling of conformational space using a simulated annealing protocol[63] and SIBFA force field is desirable to locate the lowest-energy conformations in the vicinity of each representative candidate. The OPEP parameters and notably the interactions involving Zn and the parameters used for clusterization, namely, the energy gap and the rmsd threshold, must be revisited on the structures of several Zn-finger folding motifs. The OPEP-MC method works reasonably well for the mutant finger (the lowest-energy structures deviate by 2−3 Å from NMR) but fails for the native finger. Two reasons can be proposed: (1) The energy-gap threshold is not optimal. The OPEP-MC runs of the native-finger-generated structures deviate by 2.0 Å from NMR, but these were destabilized by 10 kcal/mol with respect to the lowest-energy conformations. Although SIBFA postprocessing of all OPEP/MC representative conformers within 10 kcal/mol is feasible for 18-residue fingers (in this work, this number is 50), it becomes problematic for larger Zn-fingers. (2) The balance between the various OPEP interactions is not optimal. The low-energy conformations, shown in Figure 1, lack the β turn at Ala25-Arg26 and prefer more extended conformations in the region of Ala24-Arg28.

Two improvements of the SIBFA force field are in progress. The first is to consider the dipoles induced by the solvent reaction field and to vary the λ multiplicative factor of the van der Waals surface as a function of the nature of the fragments to increase the accuracy of the solvation energies for ionic, polar, and apolar molecules. The second improvement involves the explicit introduction of an exponentially varying penetration component to the electrostatic E_{MTP} term.[55] Taken together, such improvements should help us to investigate a much wider range of Zn-finger folding motifs as well as small proteins encompassing two successive Zn-fingers (40 residues).

Acknowledgment. We thank Drs. B.P. Roques and M.C. Fournié-Zaluski for instructive discussions and Dr. Nelly Morellet for providing us with the coordinates of the mutated CCHH Zn-finger (DD tautomer). The SIBFA computations were performed at the Centre d'Informatique National de l'Enseignement Superieur (CINES), Montpellier, France. The quantum chemical computations were performed at the Institut de Developpement en Ressources Informatiques (IDRIS), Orsay, France. We thank CINES and IDRIS for support.

Supporting Information Available: Parameters used for the E_{rep}, E_{pol}, E_{ct}, and E_{disp} energy components. Intermolecular interaction energies, ΔE_{inter}, and solvation free energies ΔG_{solv} (in kcal/mol) of the NMR and five predicted conformations for the native Zn-finger using a simplified representation. SIBFA energies (kcal/mol) and Zn^{2+} coordination (Å) of the NMR-minimized (NMR-min) and the three predicted (a−c) conformations for the all-atom mutated Zn-finger (DE tautomer). Model Sc−Zn for the NMR-minimized structure of the CCHH Zn-finger used to compare ΔE_{inter}, ΔE_{intra} and ΔG_{solv} using SIBFA and DFT. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Pillardy, J.; Czaplewski, C.; Liwo, A.; Lee, J.; Ripoll, D. R.; Kazmierkiewicz, R.; Oldziej, S.; Wedemeyer, W. J.; Gibson, K. D.; Arnautova, Y. A.; Saunders: J.; Ye, Y. J.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 2329−2333.
(2) Liu, Y.; Beveridge, D. L. *Proteins* **2002**, *46*, 128−146.
(3) Simons, K. T.; Strauss, C.; Baker, D. *J. Mol. Biol.* **2001**, *306*, 1191−1199.
(4) Duan, Y.; Kollman, P. A. *Science (Washington, D.C.)* **1998**, *282*, 740−744.
(5) Zemla, A.; Venclovas, Moult, J.; Fidelis, K. *Proteins* **2001**, *45-(Suppl 5*) 13−21.
(6) Lee, M. R.; Baker, D.; Kollman, P. A. *J. Am. Chem. Soc.* **2001**, *123*, 1040−1046.
(7) Diaz, N.; Suarez, D.; Merz, KM. Jr. *J. Am. Chem. Soc.* **2001**, *123*, 9867−9879.
(8) Berweger, C. D.; Thiel, W.; van Gunsteren, W. F. *Proteins* **2000**, *41*, 299−315.
(9) Bienstock., R. J.; Darden, T.; Wiseman, R.; Pedersen, L.; Barrett, J. C. *Cancer Res.* **1996**, *56*, 2539−2545.
(10) Peterson, M. J.; Morris, J. F. *Gene* **2000**, *254*, 105−118.
(11) Viles, J. H.; Cohen, F. E.; Prusiner, S. B.; Goodin, D. B.; Wright, P. E.; Dyson, J. H. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 2042−2047.
(12) Cavagnero, S.; Zhou, Z. H.; Adams, M. W.; Chan, S. I. *Biochemistry* **1998**, *37*, 3377−3385.
(13) Bai, Y.; Chung, J.; Dyson, H. J.; Wright, P. E. *Protein Sci.* **2001**, *10*, 1056−1066.
(14) Tiraboschi, G.; Gresh, N.; Giessner-Prettre, C.; Pedersen, L. G.; Deerfield, D. W. *J. Comput. Chem.* **2000**, *21*, 1011−1039.
(15) Derreumaux, P. *J. Chem. Phys.* **1999**, *111*, 2301−2310.
(16) Derreumaux, P. *Phys. Rev. Lett.* **2000**, *85*, 206−209.
(17) Derreumaux, P. *J. Chem. Phys.* **2002**, *117*, 3499−3503.
(18) Gresh, N.; Roques, B. P. *Biopolymers* **1997**, *41*, 145−164.
(19) Gresh, N.; Sponer, J. *J. Phys Chem. B* **1999**, *104*, 11415−11427.

(20) (a) Gresh, N. *J. Phys. Chem. A* **1997**, *101*, 8680−8694. (b) Gresh, N.; Guo, H.; Salahub, D. R.; Roques, B. P.; Kafafi, S. A. *J. Am. Chem. Soc.* **1999**, *121*, 7885−7894. (c) Guo, H.; Gresh, N.; Roques, B. P.; Salahub, D. R. *J. Phys. Chem. B* **2000**, *105*, 9746−9754.

(21) Garmer, D. R.; Gresh, N.; Roques, B. P. *Proteins* **1998**, *31*, 42−60.

(22) Antony, J.; Gresh, N.; Olsen, L.; Hemmingsen, L.; Schofield, C.; Bauer, R. *J. Comput. Chem.* **2002**, *23*, 1281−1296.

(23) Maynard, A. T.; Covell, D. G. *J. Am. Chem. Soc.* **2001**, *123*, 1047−1058.

(24) Morellet, N.; Jullian, N.; De Rocquigny, H.; Maigret, B.; Darlix, J. L.; Roques, B. P. *EMBO J.* **1992**, *11*, 3059−3065.

(25) Summers, M. F.; Henderson, L. E.; Chance, M. R.; Bess, J. W.; South, T. L.; Blake, P. R.; Sagi, I.; Perez-Alvarado, G.; Sowder, R. C. I.; Hare, D. R.; Arthur, L. O. *Protein Sci.* **1992**, *1*, 563−574.

(26) Julian, N.; Demene, H.; Morellet, N.; Maigret, B.; Roques, B. P. *FEBS Lett.* **1993**, *331*, 43−48.

(27) Simonson, T.; Calimet, N. *Proteins* **2002**, *49*, 37−48.

(28) Dudev, T.; Lim, C. *J. Am. Chem. Soc.* **2002**, *124*, 6759−6756.

(29) Ramboarina, S.; Morellet, N.; Fournié-Zaluski, M.-C.; Roques, B. P. *Biochemistry* **1999**, *38*, 9600−9607.

(30) Metropolis, N. S.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. *J. Chem. Phys.* **1953**, *21*, 1087−1092.

(31) Gresh, N.; Claverie, P.; Pullman, A. *Theor. Chim. Acta* **1984**, *66*, 1−20.

(32) Gresh, N. *J. Comput. Chem.* **1995**, *16*, 856−882.

(33) Vigné-Maeder, F.; Claverie, P. *J. Chem. Phys.* **1988**, *88*, 4934−4948.

(34) Garmer, D. R.; Stevens, W. J. *J. Phys. Chem.* **1989**, *93*, 8263−8270.

(35) Stevens, W. J.; Basch, H.; Krauss, M. *J. Chem. Phys.* **1984**, *81*, 6026−6033.

(36) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A., Jr. *J. Comput. Chem.* **1993**, *14*, 1347−1363.

(37) Creuzet, S.; Langlet, J.; Gresh, N. *J. Chim.-Phys (Paris)* **1991**, *88*, 2399−2409.

(38) (a) Rogalewicz, F.; Ohanessian, G.; Gresh, N. *J. Comput. Chem.* **2000**, *21*, 963−973. (b) Tiraboschi, G.; Fournie-Zaluski, M. C.; Roques, B. P.; Gresh, N. *J. Comput. Chem.* **2001**, *22*, 1038−1047.

(39) Ren, P.; Ponder, J. W. *J. Comput. Chem.* **2002**, *23*, 1497−1506.

(40) Evangelakis, G. A.; Rizos, J. P.; Lagaris, I. E.; Demetropoulos, I. N. *Comput. Phys. Commun.* **1987**, *46*, 401−415.

(41) Koehl, P.; Delarue, M. *J. Mol. Biol.* **1994**, *239*, 249−275.

(42) Langlet, J.; Claverie, P.; Caillet, J.; Pullman, A. *J. Phys. Chem.* **1988**, *92*, 1631−1643.

(43) Langlet, J.; Gresh, N.; Giessner-Prettre, C. *Biopolymers* **1995**, *36*, 765−780.

(44) Becke, A. *J. Chem. Phys.* **1993**, *98*, 5648−5652.

(45) Gresh, N.; Leboeuf, M.; Salahub, D. R. In *Modeling the Hydrogen Bond*; Smith, D. A. Ed.; ACS Symposium Series 569; American Chemical Society: Washington, DC, 1994; pp 82−112.

(46) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, revision A.9; Gaussian, Inc.: Pittsburgh, PA, 1998.

(47) *Jaguar* 4.0. Schrodinger Inc.: Portland, OR, 2000.

(48) Hay, P. J.; Wadt, W. R. *J. Chem. Phys.* **1985**, *82*, 299−310.

(49) Tannor, D. J.; Marten, B.; Murphy, R.; Friesner, R. A.; Sitkoff, D.; Nicholls, A.; Ringnalda, M.; Goddard, W. A., III; Honig, B. *J. Am. Chem. Soc.* **1994**, *116*, 11875−11882.

(50) Beachy, M. D.; Chassman, D.; Murphy, R. B.; Halgren, T. A.; Friesner, R. A. *J. Am. Chem. Soc.* **1997**, *119*, 5908−5920.

(51) *Discover* software. Accelrys Inc.: San Diego, CA.

(52) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230−252.

(53) Maple, J. R.; Thacher, T. S.; Dinur, U.; Hagler, A. T. *Chem. Des. Auto. News* **1990**, *5*, 5−10.

(54) Gresh, N.; Policar, C.; Giessner-Prettre, C. *J. Phys. Chem. A* **2002**, *106*, 5660−5671.

(55) Day, P. N.; Jensen, H. H.; Gordon, M. S.; Webb, S. P.; Stevens, W. J.; Krauss, M.; Garmer, D.; Cohen, D. *J. Chem. Phys.* **1996**, *105*, 1968−1986.

(56) Topol, I. A.; Casas-Finet, J. R.; Gussio, R.; Burt, S. K.; Erickson, J. W. *J. Mol. Struct.: THEOCHEM* **1998**, *423*, 13−28.

(57) Ryde, U. *Eur. Biophys. J.* **1996**, *24*, 213−221.

(58) Ryde, U. *Int. J. Quantum Chem.* **1994**, *52*, 1229−1243.

(59) Dauter, Z.; Wilson, K.; Sieker, L.; Moulis, J.; Meyer, J. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 8836−8840.

(60) Fabris, D.; Hathout, Y.; Fenselau, C. *Inorg. Chem.* **1999**, *38*, 1322−1325.

(61) Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. *J. Appl. Crystallogr.* **1993**, *26*, 283−291.

(62) Morellet, private communication.

(63) Moret, M. A.; Pascutti, P. G.; Bisch, P. M.; Mundim, K. C. *J. Comput. Chem.* **1998**, *19*, 647−657.