Metal-Assembled Modular Proteins: Toward Functional Protein Design

MARTIN A. CASE*,† AND
GEORGE L. MCLENDON‡

Department of Chemistry, The University of Vermont, Burlington, Vermont 05405, and Department of Chemistry, Princeton University, Princeton, New Jersey 08544

Received November 4, 2003

ABSTRACT

Metal-assembled parallel helix-bundle proteins have been used to investigate electron transfer through α-helical structures. Fermi Golden Rule distance dependence of electron transfer rates was established in a family of designed metalloproteins, and the contribution of intrahelical hydrogen bonding to the matrix tunneling element was explored. The first steps toward the design of functional proteins using dynamic combinatorial assembly of α-helical structural elements are described.

Introduction

Protein design, while no longer in its infancy, has entered a period of awkward adolescence, with the attendant mixture of boundless optimism and uncertain goals. Pioneering work from many labs has shown that it is possible to rationally design certain motifs with good predictive success.[1−10] Particular success has been achieved for helix-bundle motifs of two,[11−15] three,[16−21] four helices,[21−25] and even higher oligomers.[3,26−27] These studies have employed a number of thoughtful approaches, from the clever combinatorial "binary coding" of Hecht[28−30] to the most recent elegant rational design strategies of DeGrado et al.[31−33]

In contrast to the (relative) success of structural design, functional design remains less developed and more challenging. Among the many approaches that could be adapted for functional design, we have been particularly charmed by metal-ion assembly, as first realized by Ghadiri and Sasaki.[17−19,24] (Other templated assembly approaches have been pursued elsewhere.[34−37])

In this Account, we outline our recent and continuing work on metal-assembled modular proteins (MAMPs) in our continuing quest to develop robust design principles not only for structure but also for function.

Bifunctional Redox Proteins: The Simplest Paradigm?

In the original MAMP approach, a bidentate 2,2′-bipyridine ligand was used to bind an octahedral metal ion (Figure 1). In this way, both the peptide/metal stoichiometry and the parallel three-helix topology are dictated by metal assembly.

For one of us (G.L.M.) steeped in redox photochemistry, the idea of a protein based on the ubiquitous "Ru bipy" motif proved irresistible. One can readily imagine appending an acceptor motif to such a construct and creating a "donor−acceptor" protein with defined distance and driving force for intramolecular electron transfer and thus, perhaps, a predictable rate. This seemed to us a good test bed for protein design. We hoped to use the insights developed through many studies of (fixed-distance) election transfer to address a simple challenge. Could one design, from first principles, a system that gave not only a predictable structure but also a predictable reaction rate? Since electron transfer rates depend dramatically and predictably on designable parameters such as distance and driving force, redox protein design offered an acceptable challenge for a first generation of functional design.

Whenever a molecule is designed to probe function, questions regarding the detailed molecular structure are invariably (and rightly) raised. In the case of the parallel three-helix MAMPs, such questions have proved difficult to answer. Leaving aside the special problems of X-ray crystallography, multidimensional ^{1}H NMR spectra have proved difficult to interpret because of the amino acid sequence degeneracy, which causes resonances due to residues distant in sequence space to have coincident chemical shifts. Also, the 3-fold symmetry of the system precludes observation of cross-strand nuclear Overhauser enhancement effects (NOE), observations that are necessary if the interhelical angles are to be determined (is the structure parallel, as suggested, or splayed?). To resolve these issues, we used paramagnetic metal ions in the assembly process to allow the resonance degeneracies to be lifted by exploiting the well-known orientational dependence of paramagnetic pseudocontact shifts.[38−44] Three N-bipyridylated peptides can coordinate to a single paramagnetic Co^{2+} (or Ni^{2+}) ion to create an octahedral coordination complex for which large pseudocontact shifts and minimal dipolar line broadening would be expected. In addition, pseudocontact shift analysis permits the determination of the relative arrangement of individual helices within the bundle. Solution NMR studies on Ni^{2+} and Co^{2+} complexes of the peptide αpL (bpy-GELAQKLEQALQKLEQALQK−NH_2) demonstrated that the overall structural properties are the same for these two metal ions.

The trimeric complexes $[Co(\alpha pL)_3]^{2+}$ and $[Ni(\alpha pL)_3]^{2+}$ exist in two detectable isomeric forms, which involve a structural reorganization at the N-terminus (the bipyridyl end). Inter- and intrastrand NOEs were unobtainable for this region, as the relaxation time of the high-spin Ni^{2+} system is of the same order as the rate of structural interconversion. However, evidence of helical structure for

George McLendon is Russell Wellman Moore Professor of Chemistry and Chair of the Department of Chemistry at Princeton University.

Martin Case is an assistant professor in the Department of Chemistry at the University of Vermont.

* To whom correspondence should be addressed. E-mail: martin.case@uvm.edu.
† The University of Vermont.
‡ Princeton University.

10.1021/ar960245+ CCC: $27.50
Published on Web 08/12/2004

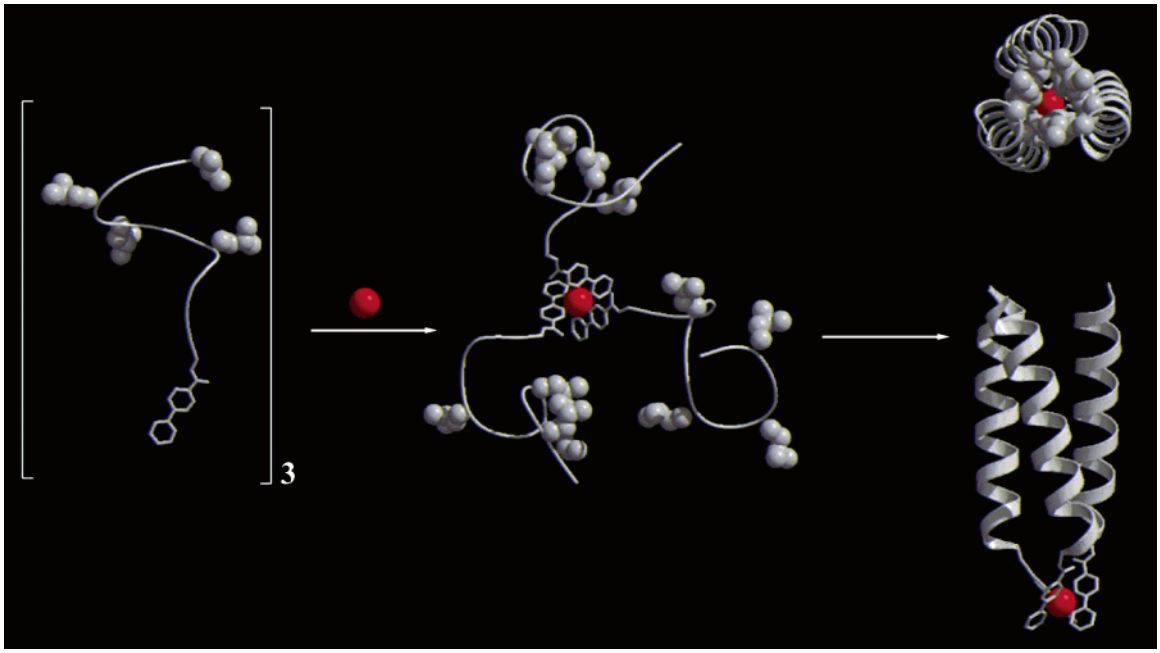


FIGURE 1. MAMP assembly. An octahedral, six-coordinate transition metal ion sequesters three bidentate 2,2′-bipyridyl ligands (left). The resulting complex (center) undergoes hydrophobic collapse to the folded parallel three-helix bundle (right). Note that complex formation is in reality stepwise, proceeding via a folded two-helix intermediate rather than the unfolded tripeptide shown.

this region was obtained from the observed paramagnetic shifts. Paramagnetic shift calculations for a variant peptide αpL(Q5H), in which histidine was introduced as an NMR marker, put the imidazole ring protons between 12 and 14 Å of the metal ion. This is in accordance with a helical backbone structure at the N-terminus. Thus the N-terminal region of the bundle appears to be dynamic but not disordered. The properties of the NMR spectrum of the Co^{2+} complex of αpL(Q5H) provide some insight into the nature of the conformational fluctuation. Detailed analysis implicates the bundle structure as causing a distortion of the ideal octahedral geometry of a tris-bipyridyl metal complex. In the final structures obtained by simulated annealing, deviations from ideal backbone geometry are restricted to the configuration around the bipyridyl−glycine amide bond. The selection of the linkage bpy-GEL- to couple the metal coordination to the helix bundle is thus probably not optimal. The leu3 side chain is bulky and may be the cause of the two distinct conformations as it flips from a hydropathically stable but sterically overcrowded core environment to a more open but destabilizing solvent exposed position.

In contrast, the C-termini of the complexes are surprisingly well structured for such short peptides. Good dispersion of NMR resonances and a clear helical pattern for the NOEs are observed. Discrete NOE contacts, including the leucine residues leu11, leu14, and leu18, are observed. Taken together, these results may indicate that there is in fact a discrete structure (at least over the 10 C-terminal amino acids) and that the bundle is well folded (Figure 2). Circular dichroism studies confirm the high degree of helical structure, although they are not sensitive to the dynamic properties seen in the NMR.

Our first attempts, albeit rudimentary, demonstrated that we could indeed design three helical bundles with bifunctional redox assemblies and directly observe electron transfer within the three-helix bundle. However, the

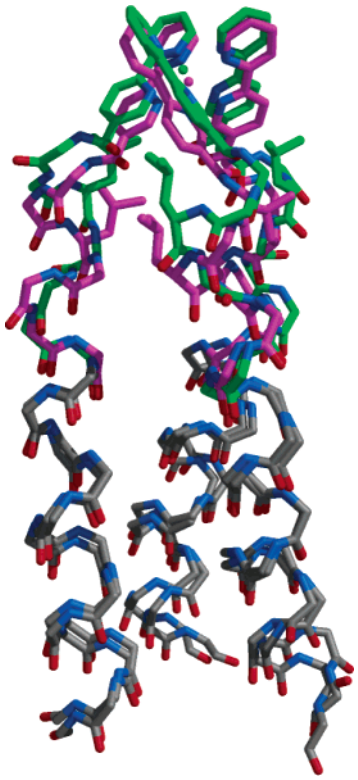


FIGURE 2. Superposition of conformers arising from pseudocontact shift analysis and molecular dynamics simulations of $[Co(\alpha pL\text{-}(Q5H))_3]^{2+}$. The two structures were superimposed from residues 13−20, in accordance with experimental observation of a unique structure for this region. Only the backbone, bipyridyl group, metal ion, and leu3 side chain are shown. Residues 1−12 for the two principal conformers are shown in magenta and green.

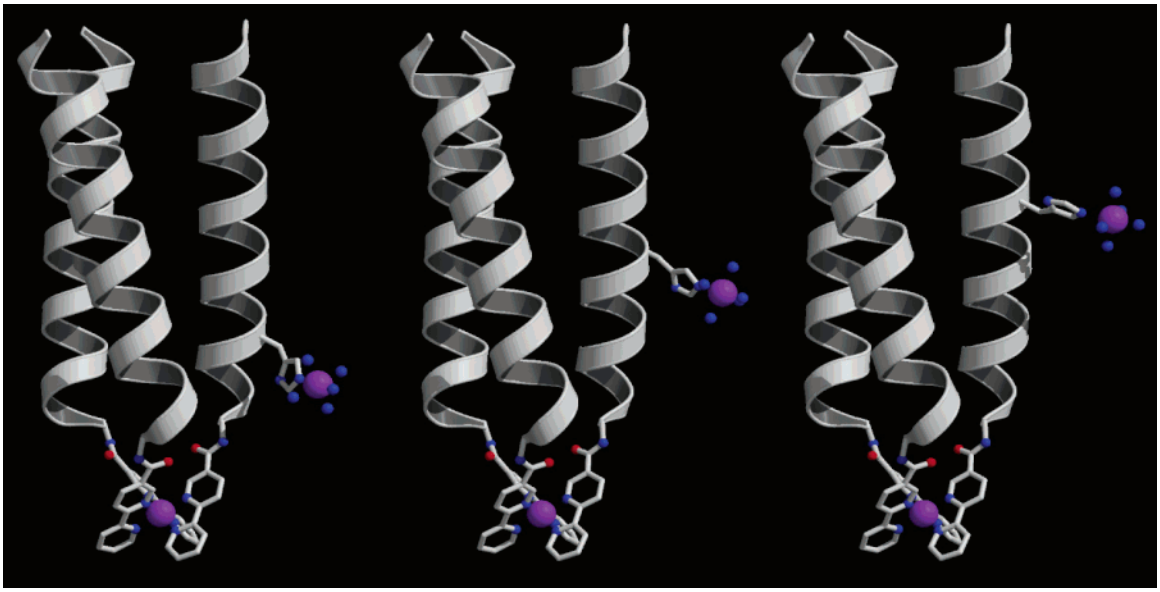


FIGURE 3. The family of MAMPs constructed to explore the distance dependence of ET rates. The $Ru(bpy)_3$ electron donor and Ru(imidazolyl)-$(NH_3)_5$ electron acceptor groups are shown. The constructs (left to right), are $[Ru(\alpha pL)_2\alpha pL(Q5H)]^{2+}$, $[Ru(\alpha pL)_2\alpha pL(Q9H)]^{2+}$, and $[Ru(\alpha pL)_2\alpha pL(Q12H)]^{2+}$. The "line-of-sight" donor−acceptor distances are 15, 20, and 25 Å, and the ET rates are 3×10^6, 3×10^3, and 1×10^2 per second, respectively. Note that in each case only one of the three helices has been elaborated with an acceptor moiety.

second generation of such molecules provided a more stringent test. We created a *family* of proteins (Figure 3) in which the donor acceptor distance was offset by one helical turn.[45]

Since electron transfer rates depend exponentially on distance (or the number of functional bond couplings), each 6-Å axial displacement along the helix is expected to produce an 800-fold change in rate. To span a nearly million-fold dynamic range, we utilized pulse radiolytic techniques which had proven useful in earlier studies of natural proteins. The results are (remarkably) gratifying. It indeed proved possible to quantitatively predict how electron transfer (ET) rates depend on protein structure, providing the first such benchmark for functional protein design.

Finally, as protein ET theory had helped advance protein design, it seemed appropriate to try and return the favor and inquire "can protein design advance ET theory?" To this end, we chose to investigate the *pathway* for electron transfer within the MAMP framework. The concept of an electron transfer pathway was first clearly defined by Beratan et al.[46−50] They noted that, within a protein, the closest (line-of-sight) donor acceptor distance might not correspond to the optimal path for conduction. Direct covalent connections, they reasoned, will maximize the electronic coupling necessary for electron transfer while a "vacuum" gap over the closest approach distance would obviously minimize such coupling. By analogy, lightning does not proceed in a straight line to ground but takes the path of optimal conductivity.

As a first approach, they suggested that a parametric "counting scheme" might allow one to predict (relative) rates via different pathways. This approach had notable successes, particularly in cataloging the beautiful studies of Gray and co-workers.[51−55] However, it proved tantalizingly difficult to directly test the assumptions of the pathways models, particularly with respect to the role of hydrogen bonds in mediating electron transfer. In the pathways models, hydrogen bonds are assigned a special weight, with ostensibly better coupling than "through-space" van der Waals contacts but weaker coupling than a covalent bond. Previous attempts to test the role of H-bonds in natural proteins proved vexing. Site-directed mutagenesis allowed the perturbation of hydrogen-bonded networks within natural redox proteins, provided such networks involved H-bonding amino acid *side chains*. Unfortunately, such perturbed networks can subsequently rearrange to create a new, equally facile pathway!

In the helical geometry of the MAMP, the H-bond assumes particular importance by proving a "short circuit" along the helix axis (Figure 4). The purely synthetic construction of the MAMP offers an alternate strategy in which a *backbone* hydrogen bond is synthetically excised by replacing the amide H-bond donor with an isosteric ester depsipeptide linkage.[56−57]

The results were instructive to us. While "greenpath" computations[58] predict up to 3-fold changes in the rate as a function of hydrogen-bond deletion for three different H-bond-deleted variants (Figure 4), we found experimentally that all the data for all four variants was best fit to a single coupling parameter, which did *not* depend on H-bond deletion (Table 1). Note that our results still imply a major role for H-bonds as mediators of electron transfer. If this was not the case, and the only conduit was the covalent pathway, the predicted ET rates would be slower by many orders of magnitude.[59−63] Rather, our experimental data suggests that the multiple pathways made available by including H-bonds in the tunneling matrix element provide a mechanism for compensating for the deletion of one of their number. The conformational flexibility of the maquette and recognition of a possible role for solvent molecules made it impossible for us to

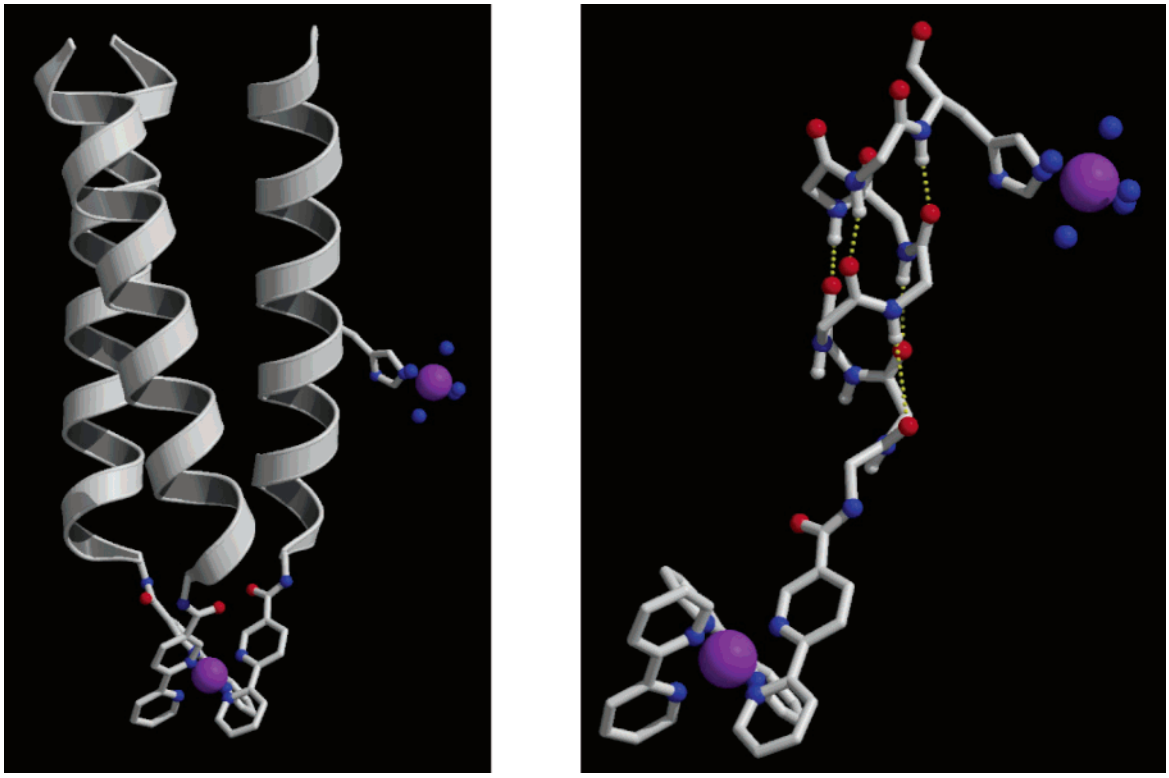


FIGURE 4. The overall topology of the three-helix construct used in the pathways explorations (left) and a detailed exposition of the hydrogen-bond network in the donor−acceptor helix (right). Four systems were measured: all H bonds intact and deletion of the H bonds between his9 and gln5, gln8 and ala4, and leu7 and leu3.

Table 1. Unimolecular Electron Transfer Rates Determined from Pulse Radiolysis Experiments

	k_{obs} (s^{-1})	k_{int} (s^{-1})	$k_{et} = k_{obs} - k_{int}$
$Ru(\alpha pL)_2\alpha pL(Q9H)$	900 (300)	200 (100)	700 (300)
$Ru(\alpha pL)_2\alpha pL(Q9H)\varphi(Q_8H_9)$[a]	800 (400)	200 (100)	600 (400)
$Ru(\alpha pL)_2\alpha pL(Q9H)\varphi(L_7Q_8)$[a]	1500 (200)	200 (100)	1300 (200)
$Ru(\alpha pL)_2\alpha pL(Q9H)\varphi(K_6L_7)$[a]	1000 (500)	200 (100)	800 (500)

[a] Depsipeptide ester linkages are between the residues X_i and Y_{i+1} in $Ru(\alpha pL)_2\alpha pL(Q9H)\varphi(X_iY_{i+1})$.

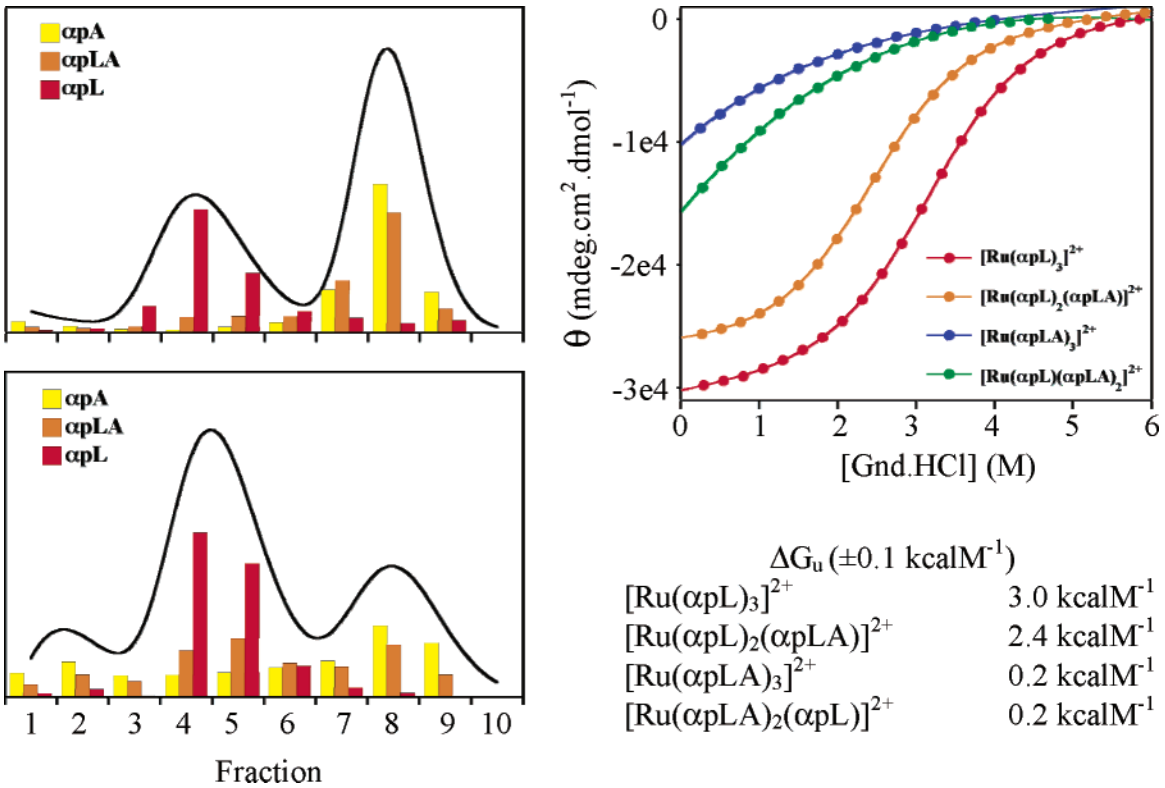
determine the absolute magnitude of the contribution of individual H-bonds to the collective tunneling pathways.

Protein Packing and Molecular Recognition

A triumph of protein evolution has been the creation of binding sites that recognize substrate size, shape, and functionality with exquisite specificity. This specificity derives from the protein's ability to spontaneously pack the protein core with extraordinary precision. The challenge to match this by design and craft substrate binding and activation sites into de novo proteins was one to which we did not feel equal.

Recognizing this dismal state of affairs, we cheerfully chipped away at the ET paradigm until a happy accident refocused our attention. A coordination chemist, viewing the trisbipyridyl chelate which defines the MAMP topology, will recognize that it is chiral, with Λ and Δ forms (a second class of isomer, facial, and meridional also exists for this geometry). Since the MAMP is built from chiral amino acids, these enantiomers become separable diastereomers when bound to an inert metal ion like ruthenium. Upon isolating the unique diastereomers, we found that the Λ diastereomer differed in one crucial way from its Δ homolog: Λ is more stably folded by ca. 0.5 kcal M^{-1}. This was striking because earlier work had shown that, in the presence of more labile metal ions such as Fe^{2+} and Ni^{2+}, the Λ diastereoisomer dominated the equilibrium distribution with a diastereomeric excess of some 40%, a free-energy difference of ca. 0.5 kcal M^{-1}! This coincidence suggested that binding of the metal ion might be thermodynamically coupled to packing of the appended peptide(s).[64] A better packed interior could more effectively compete for the available metal ion. This hypothesis suggested an interesting experiment. If better packed interiors indeed bind metal more strongly, what would happen if a variety of proteins had to compete for (limiting) metal ion? Presumably, only the most stable packing solutions would bind the metal at equilibrium. Thus the peptide modules would constitute a "dynamic library" for the assembly of proteins, and in a struggle for (metal) resources, only the fittest would survive.

We explored this idea with a very limited library (more of a bookshelf), of peptide modules whose hydrophobic core residues were comprised of all leucine (αpL), all



ΔG_u (±0.1 kcalM^{-1})

$[Ru(\alpha pL)_3]^{2+}$	3.0 kcalM^{-1}
$[Ru(\alpha pL)_2(\alpha pLA)]^{2+}$	2.4 kcalM^{-1}
$[Ru(\alpha pLA)_3]^{2+}$	0.2 kcalM^{-1}
$[Ru(\alpha pLA)_2(\alpha pL)]^{2+}$	0.2 kcalM^{-1}

FIGURE 5. Left: size-exclusion chromatograms (black line) of library selected at $^1/_9$ (top) and $^2/_9$ (bottom) equivalents Fe^{2+}. The bars show the relative amounts of peptide in each fraction. Right: chemical denaturation of Ru^{2+} complexes followed by CD spectroscopy at 222 nm. Calculated unfolding free energies are shown below.

alanine (αpA), or alternating leucine and alanine (αpLA) in the hydrophobic a and d positions of a three heptad α-helix.[65] The consensus sequence of all three peptides is that of the αpL peptide described in the ET studies in the first part of this Account.

These three peptides can produce 10 different three-helix products. These products were independently synthesized and purified under *nonequilibrating* conditions (as their Ru^{2+} adducts). The folding free energies of the individual library members were then used to predict the distribution of equilibrated Fe^{2+} products (Figure 5).

We found that the equilibrium distribution of exchange-labile products was indeed predicted by the stabilities of the isolated members of the ensemble. Modular peptide assembly, under thermodynamic conditions, creates a dynamic library. This appeared to us to create a new paradigm for exploring protein packing, in which one could perform an "analogue computation", wherein the proteins explore *all* options to find the best solutions. What practical limits might we encounter in such a scheme?

Two key issues occurred to us early on. First, we have to separate the winners from the losers. In principle, this is easy: the assembled winners (three helix bundles) are three times as large as those monomers left behind. In practice, separation requires that the species not re-equilibrate during the separation process (a few minutes of chromatography). On the other hand, to fully explore all possible binding motifs, exchange must be relatively fast (seconds). This apparent conundrum is resolved by the exchange mechanism. We naively assumed that exchange would occur by a (slow) dissociation mechanism. However, experiments with labeled peptides[65] showed that exchange proceeds associatively, as the three-helix bundles transiently add a fourth helix. Thus, exchange can occur rapidly in more concentrated solutions and more slowly in the dilution of a chromatography column.

How large a library might be investigated in this way? To answer this question, we randomly varied the helix bundle core using five amino acids (ala, val, ile, leu, phe) at four positions producing $5^4 = 625$ peptide modules. Metal-assisted assembly could then (in principle) produce $(n^3 + 4n^2 + n)/6 \cong 4 \times 10^7$ possible products! This leads to the issue of how we might follow the selection process. Recent advances in mass spectrometry appeared to be particularly suited to the task, and in a productive ongoing collaboration with Alan Marshall, we revisited our "combinatorial bookshelf" using electrospray ionization (ESI)[66−69] Fourier transform ion cyclotron resonance (FT-ICR)[70,71] mass spectrometry. ESI is a gentle ionization technique that allows the generation of intact noncovalent complexes in the gas phase.[72,73] FT-ICR mass spectrometry provides high mass resolution and mass accuracy making it ideal for analyzing complex mixtures such as combinatorial libraries, and the technique has previously been applied to the analysis of dynamic libraries.[74,75] By use of these techniques, we were able to directly observe those iron(II)-assembled three-helix bundles that had been previously shown to adopt a stable fold.[76] In fact, by limiting the concentration of iron(II), we were able to

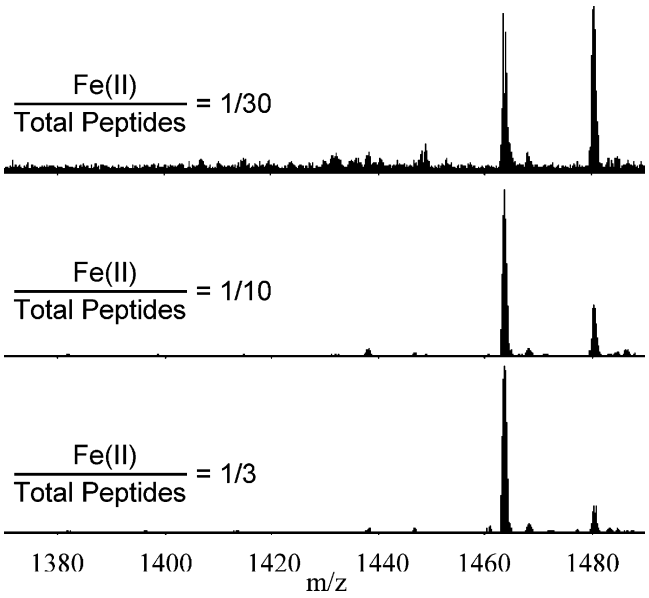


FIGURE 6. Mass spectral segments showing 5+ charge state of Fe-bound trimers obtained from samples containing the peptides αpL (10 μM), αpLA (10 μM), and αpA (10 μM) and (top) 1 μM Fe(II); (middle) 3 μM Fe(II); (bottom) 10 μM Fe(II). The leftmost peak is $[Fe(\alpha pL)_2(\alpha pLA)]^{5+}$, and the peak to the right is $[Fe(\alpha pL)_3]^{5+}$.

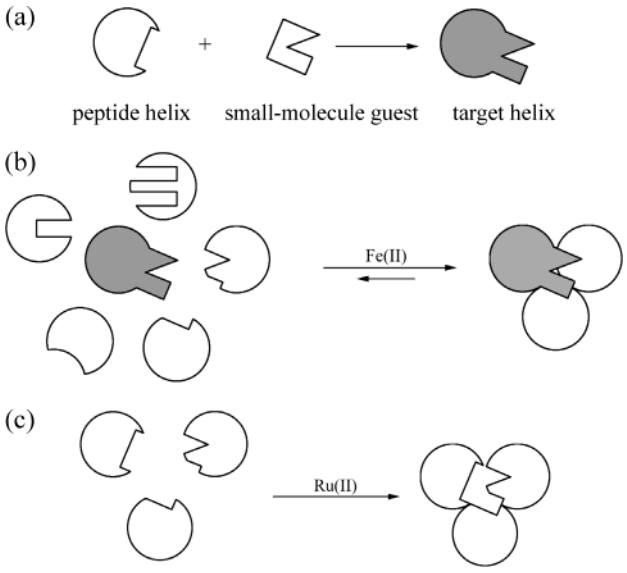


FIGURE 7. (a) Schematic synthesis of a "target" peptide helix. Actual synthesis incorporates the small-molecule target as the side chain of an unnatural amino acid in a solid-phase peptide synthesis protocol. (b) Equilibrium selection of the optimally stable 3-helix bundle incorporating the template helix. (c) Host synthesis from the peptide helix in Figure 6a and the two optimal helices from the selection in Figure 7b. The host contains a cavity "imprinted" by the small molecule target.

ascertain their relative stabilities. Circular dichroism studies of the exchange-inert ruthenium analogues of our library members suggested that only $[Ru(\alpha pL)_3]^{2+}$ and $[Ru(\alpha pL)_2(\alpha pLA)]^{2+}$ adopt the coiled-coil fold. The remaining trimers showed little or no α-helicity. Clearly, at high Fe(II) concentration, chelation of peptides that *do not* assume three-helix folds is possible. This raised the interesting question of whether unstable, nonhelical trimers could be observed in the mass spectra. The results of this study revealed an unexpected advantage in the application of this technique to screening of dynamic libraries in that unfolded structures were effectively filtered out of the analysis (Figure 6).

We also performed infrared multiphoton dissociation (IRMPD)[77,78] and electron capture dissociation (ECD)[79,80] of several trimeric species. These experiments were undertaken with a view to fragmenting the three-helix complex to learn about the peptide subunit amino acid sequence. Electron capture dissociation of peptide monomers provided extensive sequence information. However, ECD of the Fe-bound complexes yielded little sequence information, presumably due to the retention of secondary structure in the gas phase. Further evidence for gas-phase secondary structure comes from the IRMPD of the Fe-bound heterotrimers. IRMPD of Fe-bound heterotrimers also provides previously unavailable information about the Fe-bound dimers in these systems. Our results to date suggest that a hybrid approach, whereby the three-helix bundles are dissociated to dimers and monomers using IRMPD and the unfolded monomers subsequently sequenced using ECD, might present an optimal approach. Such studies are being pursued.

In summary, we learned that ESI FT-ICR mass spectrometry not only could resolve the bookshelf experiment but, in a single step, could in principle be used to identify the trimers, dimers, and monomers and sequence them! We were and remain impressed. Thus, in ongoing experiments, we now have to resolve the 4×10^7 member library down to unique sequences.

Functional Design: Molecular Recognition

After this lengthy aside, we finally return to the issue of molecular recognition. Our suggestion is that recognition of a "guest" molecule in a suitably sculpted core is no different, in principle, than is recognition of an amino acid side chain in a protein. Thus we can envision a scenario (Figure 7) in which a "target" helix selects partners by dynamic exchange.

Successful sequences may be isolated, sequenced, and resynthesized with the "target" element excised. One anticipates thereby creating a host that has been thermodynamically selected to bind a specific, complementary guest.

In a more "rational" study, we were inspired by the elegant work of Alber on the two-stranded coiled coil yeast transcription factor GCN4.[81] The parallel, two-stranded oligomerization of this protein was shown by Kim to be dependent on a buried asparagine residue, N16.[82−85] When this residue is replaced by alanine, oligomerization specificity is lost and a cavity is created. However, in the presence of benzene, the system re-equilibrates, favoring a parallel three-helix bundle that was shown by X-ray crystallography to bind benzene in the A16 cavity. It is not difficult to see how our metal-assembled three-helix systems might be brought to bear in a more rigorous investigation of host−guest specificity.

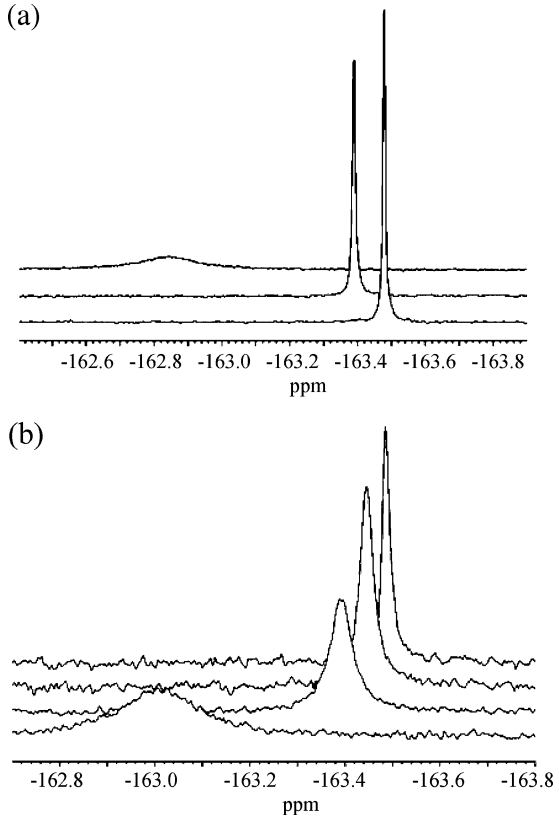


FIGURE 8. (a) Hexafluorobenzene in buffer (bottom) with $[Fe(bpyGCN4\text{-}p1(N16V))_3]^{2+}$ (center) and with $[Fe(bpyGCN4\text{-}p1(N16A))_3]^{2+}$ (top). (b) Competition titration with toluene. Relative concentration of toluene increases bottom to top.

Two 25-residue peptides were synthesized:

bpyGCN4-p1(N16A): bpy-GDKVEELLSK**A**YHLENKVAAR-LKKLV$-NH_2$

bpyGCN4-p1(N16V): bpy-GDKVEELLSK**V**YHLENKVAAR-LKKLV$-NH_2$

Homotrimers were assembled upon addition of Fe^{2+}. The $[Fe(bpyGCN4\text{-}p1(N16V))_3]^{2+}$ helix bundle serves as a control. The valine residues are expected to pack the core without creating a cavity. Binding of hexafluorobenzene to the two systems was followed using ^{19}F NMR.[86,87] The resonance of the ^{19}F nucleus is more environment sensitive than that of the proton, and significantly larger chemical shift changes may result upon binding than in ^1H NMR. Additionally, the titration spectra are easy to interpret, as there is only one set of resonances.

The hexafluorobenzene signal shifts significantly downfield in the presence of $[Fe(bpyGCN4\text{-}p1(N16A))_3]^{2+}$ to a maximum of about -162.4 ppm (when [protein] $\gg$ [ligand]) compared to -163.4 ppm in the presence of the control complex $[Fe(bpyGCN4\text{-}p1(N16V))_3]^{2+}$. The line width also significantly broadens, indicative of a bound/free ligand exchange process. Typical spectra are shown in Figure 8a. The dissociation constant, K_D, was thus determined to be $6(4) \times 10^{-4}$ M, and this was confirmed by pulsed-field gradient diffusion experiments ($K_D = 1.1(9) \times 10^{-4}$ M).

Several small organic molecules were tested to determine the specificity of the binding site for recognition of a given ligand.[88] The hexafluorobenzene ^{19}F NMR signal reports competitive displacement from the protein cavity as shown in Figure 8b.

Table 2. Inhibition Constants K_i for the Binding of Small Molecules to $[Fe(bpyGCN4\text{-}p1(N16A))_3]^{2+}$ Determined by Competition Titrations with Hexafluorobenzene

ligand	K_D (M)
toluene	3×10^{-6}
cyclohexane	8×10^{-6}
benzene	4×10^{-5}
1,3,5-trimethylbenzene	5×10^{-5}
m-xylene	1×10^{-4}
hexafluorobenzene	1×10^{-4}
1,3,5-trihydroxybenzene	1×10^{-4}
phenol	1×10^{-4}
3,5-dimethylphenol	4×10^{-4}
tetrahydropyran	1×10^{-3}
pyridine	4×10^{-3}
1,4-dioxane	3×10^{-2}

Calculated K_i values for several small guest molecules are reported in Table 2. The hydrophobic molecules toluene, cyclohexane, 1,3,5-trimethylbenzene, and benzene occupy the designed cavity with high affinity, and they bind more strongly than hexafluorobenzene. Benzene is the weakest binding of the nonpolar ligands. The data thus allows the size of the cavity to be mapped. Polar isosteres (phenol, pyridine, 1,4-dioxane, and 1,3,5-trihydroxybenzene) also bind, but with much lower affinity as might be expected for a hydrophobic cavity.

The results thus validate the design hypothesis. Work is now underway to engineer different shaped cavities, and cavities with hydrogen-bonding donors/acceptors, into metal-assembled modular proteins. Time will tell whether the combinatorial strategy for building guest-specific host proteins will prove best or whether, armed with sufficient experimental feedback, the heuristic design methods will permit fully rational design to be implemented in the laboratory.

In either case, we are confident that the modular assembly approach will continue to provide an enjoyable vehicle for exploring key issues in protein design.

This work was supported by the National Science Foundation (Grant CHE-0106342).

References

(1) Bolon, D. N.; Voigt, C. A.; Mayo, S. L. De Novo Design of Biocatalysts. *Curr. Opin. Chem. Biol.* **2002**, *6*, 125−129.
(2) Lazar, G. A.; Handel, T. M. Hydrophobic Core Packing and Protein Design. *Curr. Opin. Chem. Biol.* **1998**, *2*, 675−679.
(3) Micklatcher, C.; Chmielewski, J. Helical Peptide and Protein Design. *Curr. Opin. Chem. Biol.* **1999**, *3*, 724−729.
(4) Kraemer-Pecore, C. M.; Wollacott, A. M.; Desjarlais, J. R. Computational Protein Design. *Curr. Opin. Chem. Biol.* **2001**, *5*, 690−695.
(5) Penning, T. M.; Jez, J. M. Enzyme Redesign. *Chem. Rev.* **2001**, *101*, 3027−3046.
(6) Kennedy, M. L.; Gibney, B. R. Metalloprotein and Redox Protein Design. *Curr. Opin. Struct. Biol.* **2001**, *11*, 485−490.
(7) Pokala, N.; Handel, T. M. Review: Protein Design - Where We Were, Where We Are, Where We're Going. *J. Struct. Biol.* **2001**, *134*, 269−281.
(8) Baltzer, L.; Nilsson, J. Emerging Principles of De Novo Catalyst Design. *Curr. Opin. Biotech.* **2001**, *12*, 355−360.
(9) Bornscheuer, U. T.; Pohl, M. Improved Biocatalysts by Directed Evolution and Rational Protein Design. *Curr. Opin. Chem. Biol.* **2001**, *5*, 137−143.

(10) Hill, R. B.; Raleigh, D. P.; Lombardi, A.; DeGrado, W. F. De Novo Design of Helical Bundles as Models for Understanding Protein Folding and Function. *Acc. Chem. Res.* **2000**, *33*, 745−754.

(11) Kohn, W. D.; Hodges, R. S. De Novo Design of Alpha-Helical Coiled Coils and Bundles: Models for the Development of Protein Design Principles. *Trends Biotechnol.* **1998**, *16*, 379−389.

(12) Gurnon, D. G.; Whitaker, J. A.; Oakley, M. G. Design and Characterization of a Homodimeric Antiparallel Coiled Coil. *J. Am. Chem. Soc.* **2003**, *125*, 7518−7519.

(13) Kretsinger, J. K.; Schneider, J. P. Design and Application of Basic Amino Acids Displaying Enhanced Hydrophobicity. *J. Am. Chem. Soc.* **2003**, *125*, 7907−7913.

(14) Kwok, S. C.; Hodges, R. S. Clustering of Large Hydrophobes in the Hydrophobic Core of Two-Stranded Alpha-Helical Coiled-Coils Controls Protein Folding and Stability. *J. Biol. Chem.* **2003**, *278*, 35248−35254.

(15) Litowski, J. R.; Hodges, R. S. Designing Heterodimeric Two-Stranded Alpha-Helical Coiled-Coils: The Effect of Chain Length on Protein Folding, Stability and Specificity. *J. Pept. Res.* **2001**, *58*, 477−492.

(16) Bryson, J. W.; Desjarlais, J. R.; Handel, T. M.; DeGrado, W. F. From Coiled Coils to Small Globular Proteins: Design of a Nativelike Three-Helix Bundle. *Protein Sci.* **1998**, *7*, 1404−1414.

(17) Ghadiri, M. R.; Soares, C.; Choi, C. A Convergent Approach to Protein Design. Metal-Ion Assisted Spontaneous Self-Assembly of a Polypeptide into a Triple Helix Bundle Protein. *J. Am. Chem. Soc.* **1992**, *114*, 825−831.

(18) Lieberman, M.; Sasaki, T. Iron(II) Organizes a Synthetic Peptide into 3-Helix Bundles. *J. Am. Chem. Soc.* **1991**, *113*, 1470−1471.

(19) Lieberman, M.; Tabet, M.; Sasaki, T. Dynamic Structure and Potential-Energy Surface of a 3-Helix Bundle Protein. *J. Am. Chem. Soc.* **1994**, *116*, 5035−5044.

(20) Schneider, J. P.; Lombardi, A.; DeGrado, W. F. Analysis and Design of Three-Stranded Coiled Coils and Three-Helix Bundles. *Folding and Design* **1998**, *3*, R29−R40.

(21) Ghadiri, M. R.; Soares, C.; Choi, C. Design of an Artificial 4-Helix Bundle Metalloprotein via a Novel Ruthenium(II)-Assisted Self-Assembly Process. *J. Am. Chem. Soc.* **1992**, *114*, 4000−4002.

(22) Huang, S. S.; Gibney, B. R.; Stayrook, S. E.; Dutton, P. L.; Lewis, M. X-ray Structure of a Maquette Scaffold. *J. Mol. Biol.* **2003**, *326*, 1219−1225.

(23) Emberly, E. G.; Wingreen, N. S.; Tang, C. Designability of Alpha-Helical Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 11163−11168.

(24) Andersson, L. K.; Dolphin, G. T.; Baltzer, L. Multifunctional Folded Polypeptides from Peptide Synthesis and Site-Selective Self-Functionalization - Practical Scaffolds in Aqueous Solution. *Chembiochem* **2002**, *3*, 741−751.

(25) Vu, C.; Robblee, J.; Fairman, R. Effects of Charged Amino Acids at b and c Heptad Positions on Specificity and Stability of Four-Chain Coiled Coils. *Protein Sci.* **2001**, *10*, 631−637.

(26) De, E.; Chaloin, L.; Heitz, A.; Mery, J.; Molle, G.; Heitz, F. Conformation and Ion Channel Properties of a Five-Helix Bundle Protein. *J. Pept. Sci.* **2001**, *7*, 41−49.

(27) Ghirlanda, G.; Lear, J. D.; Ogihara, N. L.; Eisenberg, D.; DeGrado, W. F. A Hierarchic Approach to the Design of Hexameric Helical Barrels. *J. Mol. Biol.* **2002**, *319*, 243−253.

(28) Beasley, J. R.; Hecht, M. H. Protein Design: The Choice of De Novo Sequences. *J. Biol. Chem.* **1997**, *272*, 2031−2034.

(29) Roy, S.; Ratnaswamy, G.; Boice, J. A.; Fairman, R.; McLendon, G.; Hecht, M. H. A Protein Designed by Binary Patterning of Polar and Nonpolar Amino Acids Displays Nativelike Properties. *J. Am. Chem. Soc.* **1997**, *119*, 5302−5306.

(30) Wei, Y.; Kim, S.; Fela, D.; Baum, J.; Hecht, M. H. Letter to the Editor: H-1, C-13 and N-15 Resonance Assignments of S-824, a De Novo Protein from a Designed Combinatorial Library. *J. Biomol. NMR* **2003**, *27*, 395−396.

(31) Summa, C. M.; Rosenblatt, M. M.; Hong, J. K.; Lear, J. D.; DeGrado, W. F. Computational De Novo Design, and Characterization of an A_2B_2 Diiron Protein. *J. Mol. Biol.* **2002**, *321*, 923−938.

(32) Marsh, E. N. G.; DeGrado, W. F. Noncovalent Self-Assembly of a Heterotetrameric Diiron Protein. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 5150−5154.

(33) Maglio, O.; Nastri, F.; Pavone, V.; Lombardi, A.; DeGrado, W. F. Preorganization of Molecular Binding Sites in Designed Diiron Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 3772−3777.

(34) Tuchscherer, G.; Mutter, M. Templates in Protein De Novo Design. *J. Biotechnol.* **1995**, *41*, 197−210.

(35) Schneider, J. P.; Kelly, J. W. Templates that Induce Alpha-Helical, Beta-Sheet and Loop Conformations. *Chem. Rev.* **1995**, *95*, 2169−2187.

(36) Mutter, M.; Tuchscherer, G. Non-Native Architectures in Protein Design and Mimicry. *Cell. Mol. Life Sci.* **1997**, *53*, 851−863.

(37) Kwak, J.; De Capua, A.; Locardi, E.; Goodman, M. TREN (tris(2-aminoethyl)amine): An Effective Scaffold for the Assembly of Triple Helical Collagen Mimetic Structures. *J. Am. Chem. Soc.* **2002**, *124*, 14085−14091.

(38) Gochin, M.; Khorosheva, V.; Case, M. A. Structural Characterization of a Paramagnetic Metal-Ion-Assembled Three-Stranded α-Helical Coiled Coil. *J. Am. Chem. Soc.* **2002**, *124*, 11018−11028.

(39) Bertini, I.; Luchinat, C.; Rosato, A. The Solution Structure of Paramagnetic Metalloproteins. *Prog. Biophys. Mol. Biol.* **1996**, *66*, 43−80.

(40) Tu, K.; Gochin, M. Structure Determination by Restrained Molecular Dynamics Using NMR Pseudocontact Shifts as Experimentally Determined Constraints. *J. Am. Chem. Soc.* **1999**, *121*, 9276−9285.

(41) Crans, D. C.; Yang, L. Q.; Gaidamauskas, E.; Khan, R.; An, W. Z.; Simonis, U. Applications of Paramagnetic NMR Spectroscopy for Monitoring Transition Metal Complex Stoichiometry and Speciation. *ACS Symp. Ser.* **2003**, *858*, 304−326.

(42) Abel, E. W.; Coston, T. J. P.; Orrell, K. G.; Sik, V.; Stephenson, D. Two-Dimensional NMR Exchange Spectroscopy. Qualitative Treatment of Multisite Exchanging Systems. *J. Magn. Reson.* **1986**, *70*, 34−53.

(43) Bertini, I.; Luchinat, C. NMR of Paramagnetic Molecules in Biological Systems. *Physical Bioinorganic Chemistry Series*; Lever, A. B. P.; Gray, H. B., Eds.; Benjamin/Cummins Publishing Co. Inc.: Reading, MA, 1986; Vol. 3.

(44) Gochin, M.; Roder, H. Protein-Structure Refinement Based on Paramagnetic NMR shifts - Applications to Wild-Type and Mutant Forms of Cytochrome *c*. *Protein Sci.* **1995**, *4*, 296−305.

(45) Mutz, M. M.; Case, M. A.; Wishart, J. F.; Ghadiri, M. R.; McLendon, G. L. De Novo Design of Protein Function: Predictable Structure-Function Relationships in Synthetic Redox Proteins. *J. Am. Chem. Soc.* **1999**, *121*, 858−859.

(46) Beratan, D. N.; Betts, J. N.; Onuchic, J. N. Protein Electron-Transfer Rates Set by the Bridging Secondary and Tertiary Structure. *Science* **1991**, *252*, 1285−1288.

(47) Onuchic, J. N.; Beratan, D. N.; Winkler, J. R.; Gray, H. B. Pathway Analysis of Protein Electron-Transfer Reactions. *Annu. Rev. Biophys. Biomol. Struct.* **1992**, *21*, 349−377.

(48) Curry, W. B.; Grabe, M. D.; Kurnikov, I. V.; Skourtis, S. S.; Beratan, D. N.; Regan, J. J.; Aquino, A. J. A.; Beroza, P.; Onuchic, J. N. Pathways, Pathway Tubes, Pathway Docking and Propagators in Electron-Transfer Proteins. *J. Bioenerg. Biomembr.* **1995**, *27*, 285−293.

(49) Skourtis, S. S.; Beratan, D. N. High and Low Resolution Theories of Protein Electron Transfer. *J. Biol. Inorg. Chem.* **1997**, *2*, 378−386.

(50) Skourtis, S. S.; Beratan, D. N. Theories of Structure−Function Relationships for Bridge-Mediated Electron-Transfer Reactions. *Adv. Chem. Phys* **1999**, *106*, 377−452.

(51) Wuttke, D. S.; Bjerrum, M. J.; Winkler, J. R.; Gray, H. B. Electron-Tunneling Pathways in Cytochrome *c*. *Science* **1992**, *256*, 1007−1009.

(52) Regan, J. J.; DiBilio, A. J.; Langen, R.; Skov, L. K.; Winkler, J. R.; Gray, H. B.; Onuchic, J. N. Electron-Tunneling in Azurin − The Coupling Across a Beta-Sheet. *Chem. Biol.* **1995**, *2*, 489−496.

(53) Winkler J. R.; Gray, H. B. Electron Tunneling in Proteins: Role of the Intervening Medium. *J. Biol. Inorg. Chem.* **1997**, *2*, 399−404.

(54) Langen, R.; Colon, J. L.; Casimiro, D. R.; Karpishin, T. B.; Winkler, J. R.; Gray, H. B. Electron Tunneling in Proteins: Role of the Intervening Medium. *J. Biol. Inorg. Chem.* **1996**, *1*, 221−225.

(55) Gray, H. B.; Winkler, J. R. Electron Transfer in Proteins. *Annu. Rev. Biochem.* **1996**, *65*, 537−561.

(56) Zhou, J.; Case, M. A.; Wishart, J. F.; McLendon, G. L. Thermodynamic and Structural Effects of a Single Backbone Hydrogen Bond Deletion in a Metal-Assembled Helical Bundle Protein. *J. Phys. Chem.* **1998**, *102*, 9975−9980.

(57) Zheng, Y.; Case, M. A.; Wishart, J. F.; McLendon, G. L. Do Main Chain Hydrogen Bonds Create Dominant Electron-Transfer Pathways? An Investigation in Designed Proteins. *J. Phys. Chem. B* **2003**, *107*, 7288−7292.

(58) Regan, J. J. "Greenpath Software 0.95". San Diego, 1993.

(59) Moser, C. C.; Page, C. C.; Farid, R.; Dutton, P. L. Biological Electron-Transfer. *J. Bioenerg. Biomembr.* **1995**, *27*, 263−274.

(60) Moser, C. C.; Page, C. C.; Chen, X.; Dutton, P. L. Biological Electron Tunneling through Native Protein Media. *J. Biol. Inorg. Chem.* **1997**, *2*, 393−398.

(61) Kobayashi, C.; Baldridge, K.; Onuchic, J. N. Multiple versus Single Pathways in Electron Transfer in Proteins: Influence of Protein Dynamics and Hydrogen Bonds. *J. Chem. Phys.* **2003**, *119*, 3550−3558.

(62) Cukier, R. I.; Nocera, D. G. Proton-Coupled Electron Transfer. *Annu. Rev. Phys. Chem.* **1998**, *49*, 337−369.

(63) *Protein Electron Transfer*; Bendall, D. S., Ed.; BIOS Scientific Publishers: Oxford, 1996.
(64) Case, M. A.; Ghadiri, M. R.; Mutz, M. M.; McLendon, G. L. Stereoselection in Designed Three-Helix Bundle Metalloproteins. *Chirality* **1998**, *10*, 35−40.
(65) Case, M. A.; McLendon, G. L. A Virtual Library Approach to Investigate Protein Folding and Internal Packing. *J. Am. Chem. Soc.* **2000**, *122*, 8089−8090.
(66) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. Electrospray Ionization for Mass Spectrometry of Large Biomolecules. *Science* **1989**, *246*, 64−71.
(67) Smith, R. D.; Loo, J. A.; Ogorzalek Loo, R. R.; Busman, M.; Udseth, H. R. Electrospray MS Review. *Mass Spectrom. Rev.* **1991**, *10*, 359−451.
(68) Hendrickson, C. L.; Emmett, M. R. Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Annu. Rev. Phys. Chem.* **1999**, *50*, 517−536.
(69) Lorenz, S. A.; Maziarz, E. P. I.; Wood, T. D. Electrospray Ionization Fourier Transform Mass Spectrometry of Macromolecules: The First Decade. *Appl. Spectrosc.* **1999**, *53*, 18A−36A.
(70) Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S. Fourier Transform Ion Cyclotron Resonance Mass Spectrometry: A Primer. *Mass Spectrom. Rev.* **1998**, *17*, 1−35.
(71) Comisarow, M. B.; Marshall, A. G. Fourier Transform Ion Cyclotron Resonance Spectroscopy. *Chem. Phys. Lett.* **1974**, *25*, 282−283.
(72) Loo, J. A. Studying Noncovalent Protein Complexes by Electrospray Ionization Mass Spectrometry. *Mass Spectrom. Rev.* **1997**, *16*, 1−23.
(73) Smith, R. D.; Light-Wahl, K. J. The Observation of Noncovalent Interactions in Solution by Electrospray Ionization Mass Spectrometry − Promise, Pitfalls and Prognosis. *Biol. Mass Spectrom.* **1993**, *22*, 493−501.
(74) Poulsen, S. A.; Gates, P. G.; Cousins, G. R. L.; Sanders, J. K. M. Electrospray Ionisation FTICR MS of Dynamic Combinatorial Libraries. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 44−48.
(75) Otto, S.; Furlan, R. L. E.; Sanders, J. K. M. Dynamic Combinatorial Libraries of Macrocyclic Disulfides in Water. *J. Am. Chem. Soc.* **2000**, *122*, 12063−12064.
(76) Cooper, H. J.; Case, M. A.; McLendon, G. L.; Marshall, A. G. Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometric Analysis of Metal-Ion Selected Dynamic Protein Libraries. *J. Am. Chem. Soc.* **2003**, *125*, 5331−5339.
(77) Little, D. P.; Speir, J. P.; Senko, M. W.; O'Connor, P. B.; McLafferty, F. W. Infrared Multiphoton Dissociation of Large Multiply Charged Ions for Biomolecule Sequencing. *Anal. Chem.* **1994**, *66*, 2809−2815.
(78) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. Electron Capture Dissociation of Multiply Charged Protein Cations. A Nonergodic Process. *J. Am. Chem. Soc.* **1998**, *120*, 3265−3266.
(79) McLafferty, F. W.; Horn, D. M.; Breuker, K.; Ge, Y.; Lewis, M. A.; Cerda, B.; Zubarev, R. A.; Carpenter, B. K. Electron Capture Dissociation of Gaseous Multiply Charged Ions by Fourier Transform Ion Cyclotron Resonance. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 245−249.
(80) Senko, M. W.; Hendrickson, C. L.; Pasa-Tolic, L.; Marto, J. A.; White, F. M.; Guan, S.; Marshall, A. G. Electrospray Ionization FT-ICR Mass Spectrometry at 9.4 T. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 1824−1828.
(81) Gonzalez, L.; Plecs, J. J.; Alber, T. An Engineered Allosteric Switch in Leucine-Zipper Oligomerization. *Nature Struct. Biol.* **1996**, *3*, 510−515.
(82) O'Shea, E. K.; Klemm, J. D.; Kim, P. S.; Alber, T. X-ray Structure of the GCN4 Leucine Zipper, a 2-Stranded, Parallel Coiled Coil. *Science* **1991**, *254*, 539−544.
(83) Lumb, K. J.; Carr, C. M.; Kim, P. S. Subdomain Folding of the Coiled-Coil Leucine-Zipper from the Bzip Transcriptional Activator GCN4. *Biochemistry* **1994**, *33*, 7361−7367.
(84) Gonzalez, L. J.; Woolfson, D. N.; Alber, T. Buried Polar Residues and Structural Specificity in the GCN4 Leucine Zipper. *Nat. Struct. Biol.* **1996**, *3*, 1011−1018.
(85) Gonzalez, L. J.; Brown, R. A.; Richardson, D.; Alber, T. Crystal Structures of a Single Coiled-Coil Peptide in Two Oligomeric States Reveal the Basis for Structural Polymorphism. *Nat. Struct. Biol.* **1996**, *3*, 1002−1009.
(86) Dalvit, C.; Flocco, M.; Veronesi, M.; Stockman, B. J. Fluorine-NMR Competition Binding Experiments for High-Throughput Screening of Large Compound Mixtures. *Comb. Chem. High Throughput Screen* **2002**, *5*, 605−611.
(87) Gerig, J. T. Fluorine Nuclear-Magnetic-Resonance of Fluorinated Ligands. *Methods Enzymol.* **1989**, *177*, 3−23.
(88) Doerr, A. J.; Case, M. A.; Pelczer, I.; McLendon, G. L. Design of a Functional Protein for Molecular Recognition: Specificity of Ligand Binding in a Metal-Assembled Protein Cavity Probed by ^{19}F NMR. *J. Am. Chem. Soc.* **2004**, *126*, 4192−4198.

AR960245+